ISSN 2587-814X (print), ISSN 2587-8158 (online)
Russian version: ISSN 1998-0663 (print), ISSN 2587-8166 (online)

Vol. 18 No. 3 – 2024

BUSINESS INFORMATICS

HSE Scientific Journal

CONTENTS

Publisher:
HSE University

The journal is published quarterly

The journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Russian Federation

Address:
26-28, Shabolovka Street,
Moscow 119049, Russia

Tel./fax: +7 (495) 772-9590 *28509
http://bijournal.hse.ru
E-mail: bijournal@hse.ru

Circulation:
English version – 100 copies,
Russian version – 100 copies,
online versions in English and Russian – open access

Printed in HSE Printing House
44, build.2, Izmaylovskoye Shosse,
Moscow, Russia

ABOUT THE JOURNAL

Business Informatics is a peer reviewed interdisciplinary academic journal published since 2007 by HSE University, Moscow, Russian Federation. The journal is administered by HSE Graduate School of Business. The journal is issued quarterly, in English and Russian.

The mission of the journal is to develop business informatics as a new field within both information technologies and management. It provides dissemination of latest technical and methodological developments, promotes new competences and provides a framework for discussion in the field of application of modern IT solutions in business, management and economics.

The journal publishes papers in the following areas: modeling of social and economic systems, digital transformation of business, innovation management, information systems and technologies in business, data analysis and business intelligence systems, mathematical methods and algorithms of business informatics, business processes modeling and analysis, decision support in management.

The journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Russian Federation.

The journal is included into Scopus, Web of Science Emerging Sources Citation Index (WoS ESCI), Russian Science Citation Index on the Web of Science platform (RSCI), EBSCO.

The journal is distributed both in printed and electronic forms.

EDITORIAL BOARD

ABOUT HSE UNIVERSITY

Consistently ranked as one of Russia's top universities, HSE University is a leader in Russian education and one of the preeminent economics and social sciences universities in Eastern Europe and Eurasia.

Having rapidly grown into a well-renowned research university over two decades, HSE sets itself apart with its international presence and cooperation.

Our faculty, researchers, and students represent over 50 countries, and are dedicated to maintaining the highest academic standards. Our newly adopted structural reforms support both HSE's drive to internationalize and the groundbreaking research of our faculty, researchers, and students.

Now a dynamic university with four campuses, HSE is a leader in combining Russian educational traditions with the best international teaching and research practices. HSE offers outstanding educational programs from secondary school to doctoral studies, with top departments and research centers in a number of international fields.

Since 2013, HSE has been a member of the 5-100 Russian Academic Excellence Project, a highly selective government program aimed at boosting the international competitiveness of Russian universities.

ABOUT HSE GRADUATE SCHOOL OF BUSINESS

HSE Graduate School of Business was created on September 1, 2020. The School will become a priority partner for leading Russian companies in the development of their personnel and management technologies.

The world-leading model of a 'university business school' has been chosen for the Graduate School of Business. This foresees an integrated portfolio of programmes, ranging from Bachelor's to EMBA programmes, communities of experts and a vast network of research centres and laboratories for advanced management studies. Furthermore, HSE University's integrative approach will allow the Graduate School of Business to develop as an interdisciplinary institution. The advancement of the Graduate School of Business through synergies with other faculties and institutes will serve as a key source of its competitive advantage. Moreover, the evolution and development of the Business School's faculty involves the active engagement of three professional tracks at our University: research, practice-oriented and methodological.

What sets the Graduate School of Business apart is its focus on educating and developing globally competitive and socially responsible business leaders for Russia's emerging digital economy.

The School's educational model will focus on a project approach and other dynamic methods for skills training, integration of online and other digital technologies, as well as systematic internationalization of educational processes.

At its start, the Graduate School of Business will offer 22 Bachelor programmes (three of which will be fully taught in English) and over 200 retraining and continuing professional development programmes, serving over 9,000 students. In future, the integrated portfolio of academic and professional programmes will continue to expand with a particular emphasis on graduate programmes, which is in line with the principles guiding top business schools around the world. In addition, the School's top quality and all-encompassing Bachelor degrees will continue to make valuable contributions to the achievement of the Business School's goals and the development t of its business model.

The School's plans include the establishment of a National Resource Center, which will offer case studies based on the experience of Russian companies. In addition, the Business School will assist in the provision of up-to-date management training at other Russian universities. Furthermore, the Graduate School of Business will become one of the leaders in promoting Russian education.

The Graduate School of Business's unique ecosystem will be created through partnerships with leading global business schools, as well as in-depth cooperation with firms and companies during the entire life cycle of the school's programmes. The success criteria for the Business School include professional recognition thanks to the stellar careers of its graduates, its international programmes and institutional accreditations, as well as its presence on global business school rankings.

DOI: 10.17323/2587-814X.2024.3.7.23

Constructing a model to identify the determinants of successful software import substitution

Svetlana V. Begicheva [a]
E-mail: begichevas@mail.ru

Antonina K. Begicheva [b]
E-mail: abegicheva@hse.ru

Dmitry M. Nazarov [a]
E-mail: slup2005@mail.ru

[a] Ural State University of Economics, Yekaterinburg, Russia
[b] HSE University, Moscow, Russia

Abstract

In the process of import substitution, higher educational institutions face several challenges in transitioning from the predominant use of foreign software to domestic alternatives. These challenges include a lack of user experience with domestic digital solutions, difficulty in transferring data between systems and other issues. The difficulties associated with the transition period create resistance to the digital transformation process. Research on import substitution in universities has identified three main themes: the challenges and risks associated with switching to domestic software, exploring the feasibility of a complete transition to Russian software and providing recommendations for selecting Russian solutions. This study aims to identify the factors that influence the adoption of import substitution software products in higher education. The article proposes a structural model to identify the factors that contribute to successful software import substitution. The model is based on the theories of innovation diffusion and technology adoption, and it was developed using SmartPLS software. The model is based on data collected from a survey of professors and staff at the Ural State University of Economics. The results of the study indicate that the attitude towards adopting import substitution software depends on several factors, including the personal characteristics and innovative features of the software. The most significant determinants of a positive attitude towards transitioning to domestic software include user involvement and self-efficacy. In addition, a positive perception of the need for import substitution can influence individual acceptance of transitioning to Russian software and recognizing import substitution as an economic policy of the country. The theoretical significance of the study lies in its proposal of an original model for identifying the determinants of successful software import substitution that differentiates between individual acceptance and public recognition of software import substitution. The findings of the study could be useful to university management in planning and implementing measures for an import substitution strategy.

Keywords: software import substitution, technological innovations, resistance to innovations, theoretical approaches to technology acceptance, diffusion of innovations theory, structural equation modeling

Citation: Begicheva S.V., Begicheva A.K., Nazarov D.M. (2024) Constructing a model to identify the determinants of successful software import substitution. *Business Informatics*, vol. 18, no. 3, pp. 7–23. DOI: 10.17323/2587-814X.2024.3.7.23

Introduction

The Concept of Russia's Technological Development, adopted on May 20, 2023, emphasizes that by 2030 the share of domestic high-tech products, including telecommunications equipment and software, in total consumption should be at least 75%. Development of own technologies to ensure long-term competitiveness and training of qualified specialists skilled in working with Russian software should help to solve the problem of import-independence.

The Ministry of Science and Higher Education of the Russian Federation developed and approved methodological recommendations in 2022 in order to organize the effective transition of educational organizations of higher education to the predominant use of domestic software in 2022–2024 [1]. The website of the Ministry has a section "Import substitution of IT in science and higher education" with registers of hardware and software solutions for educational and research organizations [2].

Researchers have observed challenges and obstacles in implementing digital transformation in universities, particularly when transitioning to domestic software [3–8]. According to Burnyashov [3, 4], the problem of import substitution of software products used in educational programs is multifaceted: partial lack of Russian analogues of software, problems of financing the transition of universities to new software products, lack of incentives for managers and teachers of universities, the need to develop new methodological support for the educational process. The study [5] supplements this list with the lack of necessary time and human resources to transfer the IT infrastructure of universities to Russian software, as well as the reluctance of teachers of the age to retrain, which is aggravated by the shortage of young teachers of IT disciplines. Work [6] notes the problem of resistance of university teachers to the replacement of foreign software with domestic developments.

The authors of the articles [3–9] offer recommendations on leveling the problems related to import substitution of software products in universities. However, as far as we know, at the moment there are no studies that include statistical analysis of determinants of acceptance of the need for import substitution by teachers and staff of educational organizations for successful full-fledged transition to Russian software.

The article aims to analyze statistically the factors influencing the acceptance of switching to Russian software by university staff in the context of achieving technological sovereignty.

1. Materials and methods

1.1. Theoretical basis of the study

Let us consider import substitution in education as "the process of development and penetration of innovations into widespread practice" [10], in other words, as a process of involvement of university faculty and staff in the adoption of technological innovation. The authors [11] point out that there are two traditional approaches to the study of factors influencing the adoption and implementation of new technologies in the activities of organizations. The first approach is based on the *theory of diffusion of innovations* [12–13], the second – on the *technology acceptance model* [14–15].

The *theory of diffusion of innovation* explains how new products, technologies, practices, ideas, etc. spread among consumers, and defines innovation as an idea, action or object that is perceived by members of a social system (organization, settlement, society, etc.) as new [12, 13]. Within the framework of this theory, the problem of diffusion and acceptance of technological innovation is considering the characteristics of the technology being introduced. Rogers in his study [13] explained that certain characteristics of an innovation can promote or hinder its adoption by different users, and planned five key factors affecting the perception of the innovation:

- advantage of the innovation over previously used technologies;
- compatibility of the innovation with previously used technologies;
- the perceived complexity of deploying and using the innovation;
- the availability of the innovation for trial and testing prior to deployment;
- results of peers using the innovation.

Moore and Benbasat [16], based on the ideas of innovation diffusion theory, proposed a questionnaire to assess users' perceptions of IT innovations. In the study [16], they identified the most important factors influencing the user's decision to adopt and use IT innovations:

- voluntary use of an IT innovation;
- advantage of the new IT innovation;
- compatibility with existing practices;
- ease of use of the IT innovation;
- the opportunity to test the IT innovation before implementation;
- visibility of the results of using the IT innovation.

Followers of the theory of diffusion of innovations by Rogers consider the characteristics of the technology being introduced [17]. Proponents of the *technology acceptance model*, developed based on the theories of reasoned action and planned behavior of Ajzen [18], consider the problem of innovation implementation from the point of view of an individual user. The theory of technology acceptance pays special attention to the user's attitude towards technology and their intention to implement the innovation. Davis initially proposed the technology acceptance model [14], which was later revised by Davis, Bagozzi, and Warshaw [15].

The technology acceptance model considers behavioral intention, which is influenced by subjective norms and social attitudes which condition actual behavior (individual's acceptance of technologies).

As an illustration, let us cite the conceptual model of information technology acceptance proposed in the article by Venkatesh et al. [19] and improved in [20] (*Fig. 1*).

The technology acceptance model identifies the following key factors that influence users' perceptions of new technologies:

- perceived usefulness is the individual expected benefit of the innovation: the greater the perceived usefulness, the easier it is for the user to adopt the innovation;
- perceived ease of use – individual expected ease of use of the innovation: if the technology is easy to learn, the user will adopt it more quickly, but if the interface of the technology is awkward, the users' attitude towards the innovation will be determined accordingly;
- external variables, such as social influence, are an important factor in determining user attitudes towards an innovation.

1.2. Research design

The review of studies [12–16] allowed us to identify the determinants of successful implementation of innovative products. In order to build the model, we classified these factors into the following constructs: (1) "Personal characteristics of the user", (2) "Innovative characteristics of Russian software". The construct "Attitude towards acceptance of the necessity to switch to Russian software" will act as an intermediate dependent variable of the model. We will consider the concept of acceptance of Russian software as a symbiosis of individual acceptance and use of the innovation by the user in the educational process, and awareness of the importance of software import substitution to stimulate national economic interests. When forming the research design, we considered the results of [21], devoted to the determinants of successful digital transformation.

1.3. Research variables

1.3.1. Personal characteristics of the user

"The user personality characteristics" construct includes four variables based on the analysis of studies [12–16]: (1) "Knowledge," (2) "Individual innovation sensitivity," (3) "Self-efficacy," and (4) "Engagement." We explain the choice of variables below.

Knowledge refers to the accumulated experience related to the application of a technology or product. Knowledge allows us to assess the relative advantage of the innovation compared to the technologies in use, the perceived difficulty of implementing the innovation and the compatibility of the innovation with the technologies in use - the key factors of innovation adoption stated by Rogers [12]. Rogers argued that the faster a user realizes how to use a new technology, the faster it will be adopted. We can argue that knowledge is one of those key individual characteristics that is important for the initial stage of the innovation adoption process.

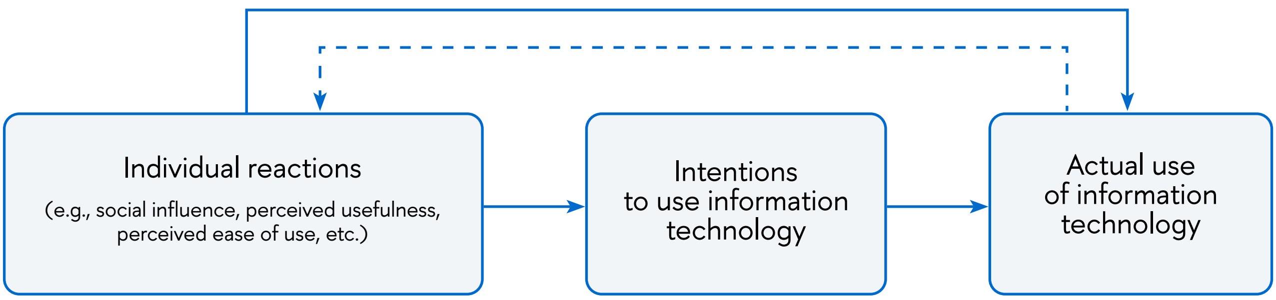

Fig. 1. Basic concept underlying user acceptance models [19, 20].

Individual innovation susceptibility is the user's endorsement of new technologies. Users with a high level of individual innovation sensitivity adopt and use new technologies before others do [12]. Individual innovation susceptibility influences the voluntariness of innovation application at the individual level, which is one factor of successful innovation adoption planned by Moore and Benbasat [16]. According to Rogers' model of innovation diffusion [12], innovators and early adopters are individuals with high innovation susceptibility and are likely to inform others about new technologies. We can say that innovativeness is a characteristic of pioneers in the application of new technologies who are subsequently looked up to by others.

Self-efficacy refers to a person's belief in their ability to solve a particular task successfully. Self-efficacy is a user's subjective confidence that Russian software products are easy for him/her to learn and use. Confidence in self-efficacy when mastering a new software product correlates with the factor "ease of use of information technology" [16].

Involvement means interest in acquiring a new skill and also correlates with the perception of the value, significance and importance of a particular technology. Engagement is the active interest and motivation to use new software products. High involvement has the goal of acquiring knowledge and skills related to IT products and encourages the acceptance and application of new technologies in work. An engaged user will voluntarily decide to use Russian software, which, according to the study [16], characterizes high technology acceptance.

1.3.2. Innovative characteristics of Russian software

The construct "Innovative characteristics of Russian software" comprises two variables: (1) "Relative advantage of Russian software" and (2) "Technological innovativeness of Russian software".

Relative advantage is a criterion for comparing an innovation and a traditional product or technology [12, 16]. The more confidently the user realizes the relative advantages of innovation, the more effective the process of its adoption [12, 16]. The perception that Russian software will be more functional, convenient, reliable and superior to previously used software is a relative advantage. The higher the level of relative advantage of the software, the higher its level of recognition and acceptance.

We can interpret **technological innovativeness** as a necessary condition for the process of innovative decision making [12]. A new technology, by definition, must be original and different from existing technologies [12]. For software to be innovative, its technological innovativeness must be sufficiently high and perceived as such by the consumer [22].

1.3.3. Attitudes towards the need to switch to Russian software

According to the studies of Davis et al. [14], who proposed the technology acceptance model, the actual acceptance (use) of technologies is caused by the intention (desire) to use them to solve problems. A positive attitude to the need to switch to Russian software means the intention to use it.

Davis et al. [14] pointed out that behavioral intentions and beliefs influence technology adoption both by individual users and by all members of an organization. In our study, attitudes towards the need to switch to Russian software will be as a variable mediating individual acceptance and recognition of import substitution as an economic policy of the country.

1.3.4. Individual acceptance of the transition to Russian software and recognition of import substitution as the country's economic policy

In order for a particular technology to be adopted, it is necessary that it have a certain value and advantage over traditional technologies. Adoption of an innovation, depending on the scale of impact, has two components: acceptance of the value of the innovation for the individual and acceptance of the value of the innovation for society.

Accepting the value of Russian software at the individual level, the user is ready to use it for personal purposes and integrate it into the educational process.

Recognizing the value of the innovation for society, the user believes that the recipient of the benefit when using the innovative technology will be the public. Recognizing the value of innovation for society in our study means realizing the importance of software import substitution to stimulate Russia's national economic interests.

As dependent variables of the model, we will consider the constructs "Individual acceptance of switching to Russian software" and "Recognition of import substitution as an economic policy of the country".

1.4. Hypotheses of the study

Within the framework of the research, we assume that individual user characteristics and innovative characteristics of Russian software form the attitude towards the need to switch to Russian software, all of which affects acceptance (both individual and recognition of the value of innovation for society) of Russian software.

Based on the research analysis, we put forward the following hypotheses.

H1: Knowledge, individual innovative sensitivity, self-efficacy and involvement have a positive impact on the user's personal characteristics that determine his/her attitude towards the need to switch to Russian software.

H2: The relative advantage of Russian software and the technological advantage of Russian software have a positive impact on innovative characteristics of software that determine the user's attitude towards software import substitution.

H3: Personal characteristics of the user and innovative characteristics of the software have a positive impact on the user's attitude towards the need to switch to Russian software.

Agreeing to switch to Russian software shows a positive attitude and a willingness to use it. We will use this statement when planning hypotheses H4 and H5.

H4: The user's desire to use Russian software has a positive effect on individual acceptance.

H5: The user's desire to use Russian software positively influences the recognition of the value of this innovation for society.

Figure 2 presents the conceptual model of the study.

1.5. Research methodology

In order to study the factors influencing the acceptance of the need to switch to Russian software, we compiled a questionnaire. The constructs' questions are derived from a study focused on digital transformation adoption issues [21]. The questionnaire comprises 28 questions related to the main scales: personal characteristics of the user – 10 questions; 8 questions – to present innovative characteristics of Russian software; 4 questions concerning the attitude to the necessity of switching to Russian software; 3 questions – to characterize individual acceptance of switching to Russian software and 3 questions – to assess the recognition of the value of switching to Russian software for society. Employees of the Sociological Laboratory and the Department of Economic Theory and Applied Sociology of the Ural State University of Economics took part in the survey. All questions were rated on a 5-point Likert scale, with 1 being the minimum value and 5 being the maximum value. *Appendix 1* presents an operationalization of the research variables.

To validate the theoretical model of the study and build a structural model, we apply the structural equation modeling (SEM) method based on partial fewest squares (PLS) analysis using SmartPLS software [23].

The model using the PLS-SEM approach comprises two sub-models: (1) the *hierarchical measurement model* determines the relationships between latent variables (hypothetical constructs) and observed variables, (2) the *structural equation model* determines causal relationships between constructs.

The reliability and consistency of scales are examined to test the hierarchical measurement model. Assessing the fit of the structural equation model involves estimating the path coefficients and their significance.

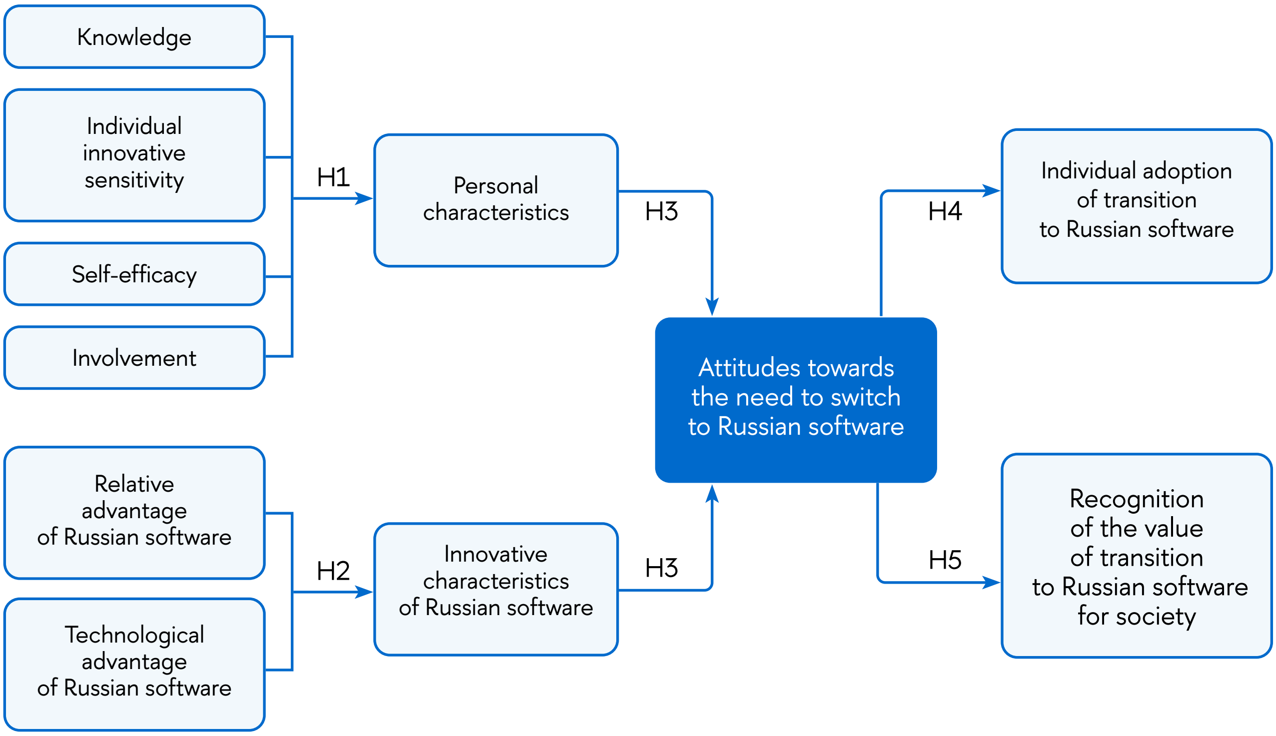

Fig. 2. Conceptual model of the study.

The aim of the PLS-SEM approach is to isolate the maximum proportion of explained total variance in the dependent latent variables in the PLS model. The PLS method allows us to investigate causal relationships under conditions of a small or medium-sized sample and does not require the assumption of a normal distribution of sample data [24–26].

2. Results of the empirical study

112 faculty and staff members of the administrative and educational support sectors of the Ural State University of Economics took part in the survey. We conducted the survey from January to February 2024.

Table 1 contains the results of frequency analysis to study the socio-demographic characteristics of the survey respondents.

Table 1.

Socio-demographic characteristics of respondents

Socio-demographic groups		Share
Age	18–35	23.2%
	36–49	35.7%
	50–64т	30.4%
	65 years and older	10.7%
Gender	Female	72.3%
	Male	27.7%
Position	Head of department, head of institute	5.36%
	Teacher	66.96%
	Administrative and training support staff member	27.68%
Total answers		112

The largest share of respondents (35.7%) belongs to the age group of 36–49 years old. 72.3% of respondents are women. The segment of respondents from faculty members made up about 66% of the respondents.

The two-stage analytical approach of PLS-SEM comprises the following stages: evaluation of the hierarchical measurement model (at this stage the validity and reliability of the selected indicators are assessed) and evaluation of the structural model.

2.1. Evaluation of a hierarchical measurement model

At the first stage of modeling using the structural equation method, it is necessary to verify the structure of the diagnostic toolkit.

To verify the reliability, we analyzed the factor loadings of each of the variables included in the analysis (*Table 2*).

Factor loadings show how significantly each variable affects the factor. Factor loadings greater than 0.7 are preferred in the model [27]; factor loadings having a value greater than 0.4 are an acceptable result. Note that the factor loadings of all variables in the model exceed the recommended value of 0.7.

Another important indicator that assesses representativeness of variables within individual constructs is convergent validity, as measured by the average variance extracted (AVE) indicator. The convergent validity criterion has a value of AVE > 0.5, showing that the variance explained by the factors included in the model is higher than the measurement error.

Table 2.

Verification of reliability of model variables

Variable	Factor load	Convergent validity (average variance extracted, AVE)
Knowledge		
I know what Russian software products can replace foreign software used in implementing disciplines nowadays	0.853	0.740
I am well aware of the pros and cons of Russian software – analog of foreign software used in implementing disciplines at the moment	0.863	
I can tell others about the possibilities of Russian software – analog of foreign software used in implementing disciplines now	0.864	
Individual innovation sensitivity		
I usually start using innovative technologies before anyone else	0.844	0.729
I tend to update my devices as new technologies emerge and new models are released	0.853	
I tend to inform others about devices created using innovative technologies	0.864	
Self-efficacy		
I think it will be easier for me to learn and start using previously unfamiliar software than for others	0.770	0.773
I think I will master the skills of working with Russian software in a relatively short time	0.937	
I am confident in my software skills, and I think I will have no difficulties working with Russian software	0.921	

Variable	Factor load	Convergent validity (average variance extracted, AVE)
Involvement		
I am open to the use of Russian software – the analog of foreign software used in implementing disciplines now	1.000	0.749
Relative advantage of Russian software		
Russian software is likely to have more capabilities than the foreign software used in implementing disciplines now	0.910	0.781
Using Russian software will be more convenient than the use of foreign software used in implementing disciplines now	0.934	
Russian software is more reliable compared to foreign software used in implementing disciplines now	0.900	
Training to work with Russian software is more relevant than training to work with foreign software used in implementing disciplines now	0.782	
Technological innovativeness of Russian software		
I think that Russian software is created using innovative technologies	0.874	0.836
Russian software is innovative	0.941	
Russian software products are original and new	0.915	
Russian software products noticeably differ for the better from foreign software used in implementing disciplines now	0.927	
Attitude towards the need to switch to Russian software		
I have a positive attitude to the use of Russian software	0.872	0.725
I have no difficulties in working with Russian software	0.822	
I am actively in favor of using Russian software in the educational process	0.860	
Individual acceptance of switching to Russian software		
I am ready to use Russian software in the educational process	0.909	0.729
If the need arises, I will use Russian software in the learning process	0.849	
I will continue to use Russian software in the future	0.837	
Recognizing the value of switching to Russian software for the society		
Our society should actively use Russian software	0.958	0.836
Organizations of various spheres of activity should use Russian software in their work	0.968	
We need to gradually increase the use of Russian software	0.931	

This value was achieved in all constructs of the model (*Table 2*).

In the next step, we conducted an internal consistency check of the variables given by the survey questions in order to determine how well each individual question describes the trait-construct. *Table 3* contains the results of the internal consistency test of the variables. Cronbach's Alpha measures consistency of the variables forming each construct. The composite reliability value (*rho_c*) demonstrates the extent to which the construct variables represent their construct. We can conclude that internal consistency is confirmed because all values of Cronbach's Alpha and composite reliability (*rho_c*) are above 0.8.

That for all model constructs, the value of the reliability coefficient (*rho_a*) lies within the bounds given by Cronbach's Alpha and composite reliability (*rho_c*) demonstrates a high level of consistency.

To test the statistical independence of the model constructs, it was necessary to assess their discriminant validity. We evaluated them using the HTMT (heterotrait-monotrait ratio) criterion, according to which one construct differs from another construct and can be included in the model if the HTMT value between constructs exceeds the threshold value of 0.9 [27]. The test showed sufficient discriminant validity of the model constructs: the maximum HTMT value was 0.856.

We can argue that the hierarchical measurement model has an adequate level of convergent reliability, internal consistency, and discriminant validity.

2.2. Evaluating the structural model

We start the structural model estimation by analyzing the values of variance inflation factor (VIF) – a metric for assessing the collinearity of the model variables. The value of VIF > 5 shows a high correlation between the variables [27]. The maximum value of VIF of the model was 3.865.

The scheme (*Fig. 3*) represents the configuration of the structural model.

Table 3.

Checking the internal consistency of the model

Construct	Cronbach's Alpha	Reliability factor (*rho_a*)	Composite reliability (*rho_c*)
Knowledge	0.824	0.825	0.895
Individual innovation sensitivity	0.816	0.825	0.890
Self-efficacy	0.853	0.900	0.910
Relative advantage of Russian software	0.905	0.911	0.934
Technological innovativeness of Russian software	0.934	0.935	0.953
Attitude towards the necessity of switching to Russian software	0.810	0.814	0.888
Individual acceptance of switching to Russian software	0.833	0.848	0.899
Recognition of the value of transition to Russian software for society	0.934	0.935	0.953

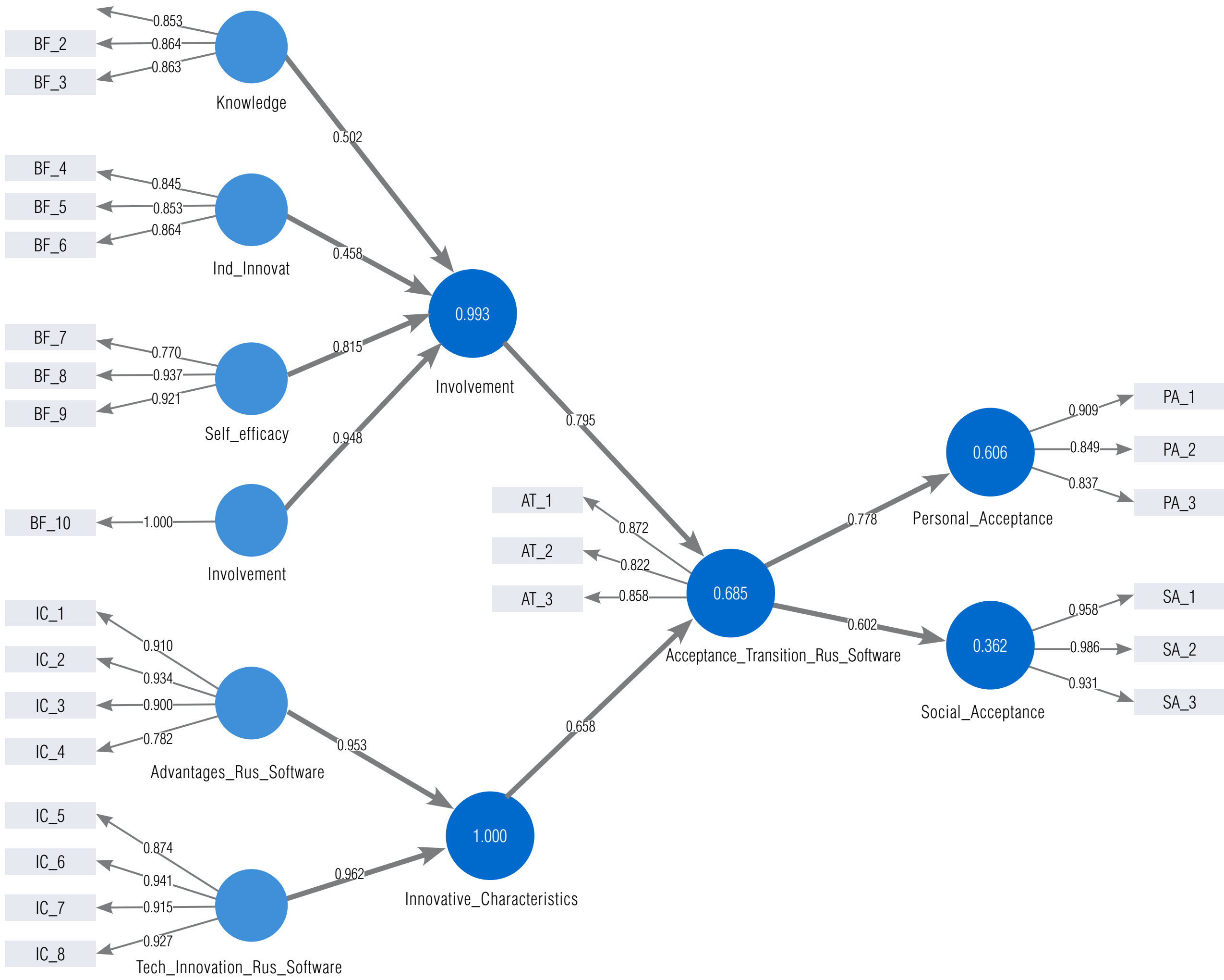

Fig. 3. Configuration of the structural model.

The structural model's adequacy is evaluated by utilizing the R^2 coefficients of determination. *Figure 3* shows the values of R^2 coefficients in circles denoting the model constructs.

The analysis of relationships between the model constructs includes the interpretation of β-coefficients and their corresponding values of *t*-statistics.

Table 4 presents the results of the structural model analysis.

Let us note the criteria necessary to analyze the data in *Table 4*:

- high *p*-values (>0.05) show that the research hypothesis is rejected;
- values of β-coefficients show the closeness of the relationship between the constructs.

3. Discussion

Thus, analyzing the results of hypothesis testing we can draw the following conclusions:

- the research hypothesis about the influence of knowledge level on personal characteristics

Table 4.

β-coefficient values and results of hypothesis testing

Hypothesis	Influence	β-coefficient	t-statistic	p-value	Solution
H1	Individual innovation responsiveness → Personal characteristics of the user	–0.004	0.040	0.968	rejected
H1	Involvement → Personal characteristics of the user	0.720	8.177	0.000	accepts
H1	Knowledge → Personal characteristics of the user	0.113	1.327	0.184	rejected
H1	Self-efficacy → Personal characteristics of the user	0.313	2.531	0.011	accepted
H2	Relative advantage of Russian software → Innovative characteristics of Russian software	0.495	30.845	0.000	accepted
H2	Technological innovativeness of Russian software → Innovative characteristics of Russian software	0.549	29.820	0.000	accepted
H3	Personal characteristics of the user → Attitude to the necessity of switching to Russian software	0.625	8.193	0.000	accepted
H3	Innovative characteristics of Russian software → Attitude towards the need to switch to Russian software	0.286	3.617	0.000	accepted
H4	Attitude towards the need to switch to Russian software → Individual acceptance of switching to Russian software	0.778	17.995	0.000	accepted
H5	Attitude towards the need to switch to Russian software → Recognition of the value of switching to Russian software for society	0.602	8.180	0.000	accepted

influencing the user's attitude towards the necessity of switching to Russian software was rejected;

- the research hypothesis about the influence of individual innovation susceptibility on personal characteristics influencing the user's attitude towards the need to switch to Russian software was rejected;
- all other research hypotheses were confirmed.

Complementing the findings on the results of hypothesis testing by analyzing the values of β-coefficients we can state that:

- there is a significant influence of involvement and self-efficacy on personal characteristics influencing the user's attitude towards the necessity of switching to Russian software, the value of the influence of involvement (0.720) exceeds the value of the influence of self-efficacy (0.313);
- there is **a** statistically significant average in strength influence of relative advantage of Russian software (0.495) and technological innovativeness of Russian software (0.549) on innovative characteristics of Russian software influencing user's attitude to the necessity of switching to Russian software;
- comparing the power of influence of personal characteristics of the user (0.625) and innovative

characteristics of Russian software (0.286) on the intention to use Russian software we note the excess of the power of influence of personal characteristics;

- intention to use Russian software significantly influences both individual acceptance of switching to Russian software (0.778) and recognition of the value of switching to Russian software for society (0.602).

Personal characteristics, specifically involvement and self-efficacy, have the greatest influence on the intention to use and adoption of Russian software, as revealed by the results of the structural model analysis. Considering the mechanisms of working with resistance when implementing a project on software import substitution in a higher educational institution, it is necessary to influence these two factors of motivation for successful implementation of Russian software.

The indicators of personal characteristics of the user and innovative characteristics of Russian software together explain 68.5% of the variance of the indicators of attitude to the need to switch to Russian software ($R^2 = 0.685$). Whereas the indicators of attitude to the necessity of switching to Russian software explain 60.6% of the variance of indicators of individual acceptance of switching to Russian software ($R^2 = 0.606$) and 36.2% of the variance of indicators of recognizing the value of switching to Russian software for society ($R^2 = 0.362$).

Conclusion

The purpose of the study was to identify significant factors influencing the adoption of Russian software solutions in educational organizations. The concepts of innovation diffusion theory and technology acceptance model served as the basis for the research model. To test the hypotheses of the research, we used the method of modeling by structural equations with the use of the results of a questionnaire survey of teachers and staff of the university.

The results of the study have theoretical significance and prospects for further practical application.

The study confirmed one conclusion of the technology acceptance model about the influence of behavioral intentions to use information technology on its direct use. We got a statistically significant result that the attitude towards the need to switch to domestic software acts as a mediating factor between the independent and dependent variables of the study. The analysis is consistent with the ideas of the followers of the theory of diffusion of innovations: individual acceptance of import substitution and recognition of the value of switching to domestic software for society are influenced by both personal characteristics of the user and innovative characteristics of Russian software. A positive attitude to the necessity of import substitution mediates the influence on individual acceptance to a greater extent than on the recognition of the value of transition to Russian software for society. This shows that there are additional factors not considered in the model that influence the recognition of import substitution as a basic priority of Russian economic development.

The results of the study have practical significance. By systematically analyzing the factors that impact successful transition to domestic software in higher educational institutions, university management can strategically plan and improve import substitution activities.

The study has several limitations. First, the study was carried out on a relatively limited sample size, which may cause a representativeness error. Second, the social desirability effect may have influenced the respondents' answers. Interviewees may have consciously or unconsciously chosen socially approved answers and overestimated or underestimated their agreement with the need for import substitution.

Subsequent research should consider other important factors influencing the acceptance of software import substitution. It would be interesting to conduct similar studies in organizations of various sectors of the Russian economy, both in commercial companies and in government agencies. ■

References

1. Ministry of Education and Science of Russia (2022) *Methodological recommendations for the transition of educational organizations of higher education to the primary use of domestic software (messengers and video conferencing), including domestic office software*. Available at: https://minobrnauki.gov.ru/importozameshcheniye/Метод_рекомендации_отеч_ПО_ВУЗы.pdf (accessed 24 May 2024) (in Russian).

2. Ministry of Education and Science of Russia (2022) *Import substitution of IT in the field of science and higher education*. Available at: https://minobrnauki.gov.ru/importozameshcheniye/ (accessed 24 May 2024) (in Russian).

3. Burnyashov B.A. (2022) Import substitution of the software used in the educational process in Russian universities. *Informatics and Education*, vol. 37, no. 1, pp. 27–36 (in Russian).

4. Burnyashov B.A. (2023) Russian cloud office application packages in the educational process of universities. *Informatics and Education*, vol. 38, no. 2, pp. 5–15 (in Russian).

5. Innopolis University (2022) *Analytical report on determining the potential for import substitution of software (in the system of higher and secondary vocational education)*. Innopolis: Innopolis University (in Russian).

6. Pasurin D.A. (2023) Problems of software import substitution in universities. *Digital models and solutions,* vol. 2, no. 4, pp. 63–75 (in Russian).

7. Muzalevskaya A.A., Gajdamakina I.V. (2023) Prerequisites and opportunities for import substitution of software in the training of specialists in Russian universities. *Scientific Notes of Orel State University*, no. 2(99), pp. 275–279 (in Russian).

8. Polonskij A.M. (2022) Import substitution of software and organization of student education using domestic or free software. *Aktual'nye problemy ekonomiki i upravleniya (Actual problems of economics and management)*, no. 2(34), pp. 65–82 (in Russian).

9. Savvateeva T.P. (2019) Problems of software import substitution for teaching of bachelor students to information systems design. *Modern problems of science and education,* no. 5 (in Russian).

10. Zagvyazinskij V.I. (2001) *Learning Theory: A modern interpretation*. Moscow: Academia (in Russian).

11. Kalinichenko N.S., Velichkovskij B.B. (2022) The technology acceptance phenomenon: current state and future research. *Organizational Psychology*, vol. 12, no. 1, pp. 128–152 (in Russian).

12. Rogers E.M. (2010) *Diffusion of innovations*. New York: Simon and Schuster.

13. Rogers E.M., Agarwala-Rogers R. (1980) *Communication in organizations*. Moscow: Economy (in Russian).

14. Davis F.D. (1985) *A Technology Acceptance Model for empirically testing new end-user information systems: theory and results*. PhD dissertation. Cambridge: Massachusetts Institute of Technology.

15. Davis F.D., Bagozzi R.P., Warshaw P.R. (1989) User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, vol. 35, no. 8, pp. 982–1003.

16. Moore G.C., Benbasat I. (1991) Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*, vol. 2, pp. 192–222. https://doi.org/10.1287/isre.2.3.192

17. Parasuraman A. (2000) Technology Readiness Index (TRI): A multiple-item scale to measure readiness to embrace new technologies. *Journal of Service Research*, vol. 2, no. 4, pp. 307–320. https://doi.org/10.1177/109467050024001

18. Ajzen I. (1991) The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, vol. 50, pp. 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

19. Venkatesh V., Morris M., Davis G., Davis F. (2003) User acceptance of information technology: Toward a unified view. *MIS Quarterly*, vol. 27, no. 3, pp. 425–478. https://doi.org/10.2307/30036540

20. Abraham C., Boudreau M., Junglas I., Watson R. (2013) Enriching our theoretical repertoire: The role of evolutionary psychology in technology acceptance. *European Journal of Information Systems*, vol. 22, no. 1, pp. 56–75. https://doi.org/10.1057/ejis.2011.25

21. Oh K., Kho H., Choi Y., Lee S. (2022) Determinants for successful digital transformation. *Sustainability,* vol. 14, no. 3, 1215. https://doi.org/10.3390/su14031215

22. Ram S. (1987) A model of innovation resistance. *Advances in Consumer Research*, vol. 14, pp. 208–212.

23. SmartPLS (2024) *SmartPLS 4. The world's most user-friendly statistical software*. Available at: https://www.smartpls.com/ (accessed 25 May 2024).

24. Chu Y., Chi M., Wang W. (2019) The impact of information technology capabilities of manufacturing enterprises on innovation performance: Evidences from SEM and fsQCA. *Sustainability*, vol. 11, no. 21, 5946. https://doi.org/10.3390/su11215946

25. Miceli A., Hagen B., Riccardi M.P. (2021) Thriving, not just surviving in changing times: How sustainability, agility and digitalization intertwine with organizational resilience. *Sustainability*, vol. 13, no. 4, 2052. https://doi.org/10.3390/su13042052

26. Zhang J., Long J., von Schaewen A.M.E. (2021) How does digital transformation improve organizational resilience? – Findings from PLS-SEM and fsQCA. *Sustainability*, vol. 13, no. 4, 11487. https://doi.org/10.3390/su132011487

27. Hair J.F., Hult G.T.M., Ringle C., Sarstedt M., Danks N., Ray S. (2021) *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook*. Springer.

Appendix 1.

Operationalization of the study variables

Variable	Name	Questions
Construct 1: Personal characteristics of the user		
Knowledge	BF_1	I know what Russian software products can replace foreign software used in implementing disciplines nowadays
	BF_2	I am well aware of the pros and cons of Russian software – analog of foreign software used in implementing disciplines at the moment
	BF_3	I can tell others about the possibilities of Russian software – analog of foreign software used in implementing disciplines now
Individual innovation sensitivity	BF_5	I usually start using innovative technologies before anyone else
	BF_6	I tend to update my devices as new technologies emerge and new models are released
	BF_7	I tend to inform others about devices created using innovative technologies
Self-efficacy	BF_8	I think it will be easier for me to learn and start using previously unfamiliar software than for others
	BF_9	I think I will master the skills of working with Russian software in a relatively short time
	BF_10	I am confident in my software skills and I think I will have no difficulties working with Russian software
Involvement	BF_11	I am open to the use of Russian software – the analog of foreign software used in implementing disciplines now
Construct 2. Innovative characteristics of Russian software		
Relative advantage of Russian software	IC__1	Russian software is likely to have more capabilities than the foreign software used in implementing disciplines now
	IC__2	Using Russian software will be more convenient than the use of foreign software used in implementing disciplines now
	IC__3	Russian software is more reliable compared to foreign software used in implementing disciplines now
	IC__4	Training to work with Russian software is more relevant than training to work with foreign software used in implementing disciplines now
Technological innovativeness of Russian software	IC__5	I think that Russian software is created using innovative technologies
	IC__6	Russian software is innovative
	IC__7	Russian software products are original and new
	IC__8	Russian software products noticeably differ for the better from foreign software used in implementing disciplines now
Construct 3. Attitudes towards the need to switch to Russian software		
Attitudes towards the need to switch to Russian software	AT_1	I have a positive attitude to the use of Russian software
	AT_2	I have no difficulties in working with Russian software
	AT_3	I am actively in favor of using Russian software in the educational process
Construct 4: Individual acceptance of the transition to Russian software		
Individualized acceptance	PA_1	I am ready to use Russian software in the educational process
	PA_2	If the need arises, I will use Russian software in the learning process
	PA_3	I will continue to use Russian software in the future
Construct 5. Recognizing the value of the transition to Russian software for society		
Recognizing the value of transition to Russian software for the society	SA_1	Our society should actively use Russian software
	SA_2	Organizations of various spheres of activity should use Russian software in their work
	SA_3	We need to gradually increase the use of Russian software

About the authors

Svetlana V. Begicheva

Cand. Sci. (Econ.);

Associate Professor, Department of Business Informatics, Ural State University of Economics, 62, 8 Marta Str., Yekaterinburg 620144, Russia;

E-mail: begichevas@mail.ru

ORCID: 0000-0002-0551-1622

Antonina K. Begicheva

Lecturer, Postgraduate Student,, School of Software Engineering, Faculty of Computer Science, HSE University, 11, Pokrovsky Blvd., Moscow 109028, Russia;

Research Assistant, Laboratory of Process-Aware Information Systems (PAIS Lab), Faculty of Computer Science, HSE University, 11, Pokrovsky Blvd., Moscow 109028, Russia;

E-mail: abegicheva@hse.ru

ORCID: 0000-0001-6657-1760

Dmitry M. Nazarov

Dr. Sci. (Econ.);

Head of Department of Business Informatics, Ural State University of Economics, 62, 8 Marta Str., Yekaterinburg 620144, Russia;

E-mail: slup2005@mail.ru

ORCID: 0000-0002-5847-9718

DOI: 10.17323/2587-814X.2024.3.24.40

Counterfactual explanations based on synthetic data generation

Yuri A. Zelenkov
E-mail: yzelenkov@hse.ru

Elizaveta V. Lashkevich
E-mail: evlashkevich@hse.ru

Graduate School of Business, HSE University, Moscow, Russia

Abstract

A counterfactual explanation is the generation for a particular sample of a set of instances that belong to the opposite class but are as close as possible in the feature space to the factual being explained. Existing algorithms that solve this problem are usually based on complicated models that require a large amount of training data and significant computational cost. We suggest here a method that involves two stages. First, a synthetic set of potential counterfactuals is generated based on simple statistical models (Gaussian copula, sequential model based on conditional distributions, Bayesian network, etc.), and second, instances satisfying constraints on probability, proximity, diversity, etc. are selected. Such an approach enables us to make the process transparent, manageable and to reuse the generative models. Experiments on three public datasets have demonstrated that the proposed method provides results at least comparable to known algorithms of counterfactual explanations, and superior to them in some cases, especially on low-sized datasets. The most effective generation model is a Bayesian network in this case.

Keywords: counterfactual explanations, synthetic data generation, multimodal distribution modelling, Bayesian network, credit scoring

Citation: Zelenkov Yu.A., Lashkevich E.V. (2024) Counterfactual explanations based on synthetic data generation. *Business Informatics*, vol. 18, no. 3, pp. 24–40. DOI: 10.17323/2587-814X.2024.3.24.40

Introduction

Recently, concern in interpretable AI (XAI) has grown rapidly, driven by the expanding use of machine learning algorithms in various fields of human endeavor [1, 2]. Moreover, many national and international regulators require transparency of algorithm-based decisions. In particular, the EU's General Data Protection Regulation (GDPR) provides citizens with the right to request "meaningful information about the logic involved and the meaning and intended consequences" of automated decisions, and US credit laws require that consumers be provided with reasons for unfavorable decisions [3]. Bank of Russia also follows the OECD recommendations on AI usage, whereby models should be transparent and interpretable to limit modeling risks and allow for independent external, internal and regulatory validation.

XAI methods can be categorized into two groups [4]. The first includes models where interpretability is a core property (e.g., decision trees or linear regression). The second group comprises methods that treat the model as a black box. In contrast to the models of the first group, they lack properties that provide a meaningful interpretation, so additional efforts must be made to explain the decision logic post facto (explainability). In the second group, in turn, we can distinguish the methods of model explanation, local result explanation and black-box examination [5].

This paper examines methods of counterfactual explanations [5–8]. A counterfactual explanation (CE) allows us, for a specified sample, to find a set of objects that belong to the opposite class but are as close as possible to the instance explained in the feature space. An example commonly cited in the literature is a borrower who was denied a loan based on the decision of an algorithm used at a bank. The objective of CE is to generate a profile for this borrower such that his application is approved (e.g., reducing the amount of the requested loan). An obvious constraint is the feasibility of the proposed changes, so the mandatory parameter minimized in this type of problem is the distance between the sample and the counterfactual. It follows from this example that, according to the above classification, CE belongs to the group of local ex post explanation methods, since it explains the solution of the trained model, treated as a black box, for a particular sample. In Russian, the concept of CE was first presented in a translated book [9].

A counterfactual is defined as a conditional statement in philosophy, the antecedent of which (a previous event that helps to understand the present) is false, while the consequent describes what the world would be like if the antecedent had occurred (an answer to the "what-if" question). According to the Great Russian Encyclopedia, counterfactual thinking is a type of thinking characterized by a person's tendency to imagine possible other variants of events that have already occurred, i.e. reflection contrary to facts.

While most XAI methods aim to answer the "why" question [4], counterfactual statements provide a means of interpretation by indicating what changes would be required to achieve a desired goal (prediction) rather than helping to understand why the current situation has a particular predicted outcome [8]. Therefore, many authors [5] state that CE corresponds to the third level of Pearl's causality models [10], which need to answer questions involving retrospective reasoning, e.g., "what is the probability of event y at x if there are x' and y' observed". At the same time, CE also does not impose restrictions on model complexity and does not require disclosure of model information [3].

Obviously, CE methods are a powerful decision support tool in various fields such as finance [11, 12] and healthcare [13]. Several dozen CE algorithms are already known by now (see reviews [5, 6, 8] and other papers). Most of them are premised on optimizing some target function, and when this problem is solved each time a set of counterfactuals needs to be computed for a given sample. This imposes limitations on the performance and scalability of CE [6]. A possible alternative is to use methods that allow modeling the joint distribution of the features of the objects under study. In this case, a once-trained model can generate counterfactuals for different samples without significant computational costs.

Note that in this formulation the task can be viewed as generation of synthetic tabular data [14, 15]. Both statistical methods (copulas, Bayesian networks), and machine learning methods (variational autoencoders, generative adversarial networks, etc.) are used to create such models. [16]. Some scholars also adapt for this purpose oversampling methods that are designed to generate minor class objects in the case of imbalanced data [17].

Considering these circumstances, an approach to CE based on synthetic data generation principles is proposed here, involving two steps. In the first, a set of potential counterfactuals is generated, and in the second, a selection is made of those that satisfy the constraints of actionability, proximity, cost, etc. This organization allows to make the CE process transparent, manageable, reuse generation models and thus significantly reduce computational costs.

The rest part of the paper is organized as follows. After the review of literature, the proposed method is presented in section 2. Sections 3 and 4 set out the experimental results, comparing the proposed approach with other existing known CE methods. Finally, the limitations of the proposed method as well as directions for future research are discussed.

1. Literature review

1.1. Counterfactual generation

CE is based on several implicit assumptions [3]:

- the recommended variation of attribute values is unambiguously realized in the real world;
- the distribution of feature values can be reconstructed from available training data;
- the proposed modifications are relevant only to the decision being taken and do not affect other aspects;
- the model is stable in time, monotonic and restricted to binary outcomes.

As discussed above, CE is an actively growing area of research. The very term "counterfactual explanation" as applied to AI systems was first used in [18], However, papers using a similar approach have begun to appear since the mid-2010s [5].

Let's give formal definitions. Consider a classifier $h: \mathcal{X} \to \mathcal{Y}$ trained on the data set

$$\mathcal{D} = \{(x_1, y_1), \ldots, (x_n, y_n)\}, x_i \in \mathcal{X}, y_i \in \mathcal{Y},$$

where $\mathcal{X} \subset \mathbb{R}^m$ is a feature space, and $\mathcal{Y}$ is class label range.

It is usually assumed $\mathcal{Y} = \{0, 1\}$ but all the proposed definitions can be simply generalized to the case of multiclass classification. Each instance x_i is a vector m feature pairs, $x_i = \{(a_j, v_{ij})\}_{j=1}^{m}$, where a_j is an attribute, v_{ij} is its value from the domain a_j. Attributes can be either categorical, ordinal or continuous.

Definition 1. If a classifier h assigns the label $y = h(x)$ to an instance x then the counterfactual explanation of x is an instance x^* whose label is different from y, i.e., $h(x^*) \neq y$, with the difference betweeny x and x^* being minimal.

The concept of minimal difference is not specified here as it depends on the context of the problem and will be discussed later.

Definition 2. The counterfactual explainer is a function f_k that, for a dataset $\mathcal{D}$, a classifier h and an instance x, returns a set $C = f_k(h, \mathcal{D}, x)$ of $l \leq k$ valid counterfactual examples $C = \{x_1^*, \ldots, x_l^*\}$, where k is the number of counterfactuals required.

Characteristics that enable us to evaluate the quality of the counterfactual generation algorithm:

1. Validity is measured by the ratio of the number of counterfactuals that have the required class label to the total number of generated objects [11]:

$$V = |C_v| / |C|,$$

where C_v is a set of valid counterfactuals generated by the model f_k; C is a set of samples generated f_k, $C_v \subset C$.

The validity of the generated sample is determined using a predictive model h, for a valid example. The following condition must be fulfilled $h(x^*) \neq h(x)$. As follows from the definition, the maximum validity value is $V = 1$; values less than 1 indicate insufficient efficiency of the model.

2. Proximity is the distance of a counterfactual from the sample for which the explanation is generated. The proximity of a set of counterfactuals is estimated by the average distance on this set [19]:

$$P = \frac{1}{|C_v|} \sum_{x^* \in C_v} dist\,(x^*, x).$$

To measure distance $dist(x^*, x)$ most used L_0, L_1, L_2 and L_∞ norms, $L_k = \left(\sum_i |x_i|^k\right)^{1/k}$, and its weighted combinations. The lower the value P, the closer the objects found are to the explained factual.

3. Sparsity estimates the number of features that need to be changed to move into the counterfactual class. It is preferable for counterfactuals to have the smallest possible changes in their features. This property allows for more efficient, human-understandable and interpretable counterfactuals [18].

$$S = \frac{1}{|C_v|} \sum_{x^* \in C_v} K(x^*).$$

$K(x^*)$ is the number of counterfactual attributes which value changes in comparison with the factual x. Thus, models with a lower value of S are preferred.

4. Diversity. Searching for the closest instances according to a distance function can lead to very alike counterfactual candidates with few differences between them. Diversity implies that the counterfactual generation process produces different explanations for the same sample. This leads to explanations that are more interpretable and more understandable to the user. The authors [19] propose to use the average distance between all pairs of valid counterfactuals as a measure of diversity:

$$D = \frac{1}{|C_v|^2} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} dist(x_i^*, x_j^*),$$

here $dist(x_i^*, x_j^*)$ is a measure of the distance between two counterfactuals x_i^* and x_j^*. The higher the diversity, the more efficient the CE algorithm.

5. Plausibility. This property emphasizes that the generated counterfactuals must be legitimate, and the search process must provide logically valid results. This means that the counterfactual found should never change immutable characteristics such as gender or race. Three categories of plausibility are distinguished in the literature [20]:

- domain consistency, which limits the range of acceptable values of counterfactual attributes;
- distributional consistency requires that the probabilities of counterfactual feature values are matched to the (empirical) distribution of the data. This property can be measured [6] as the average distance to k nearest neighbors e.g., local outlier factor (LOF) [21], as well as by kernel density estimation (KDE). In the last case, the density of distribution of each attribute is estimated based on KDE, and then the probability of the corresponding counterfactual attribute belonging to this distribution is calculated. This approach has obvious limitations – each attribute is considered separately, and it is applicable only to continuous attributes. The nearest neighbors method has no such limitations;
- prototype consistency selects counterfactual instances that are either directly present in the dataset or are close to the data object being explained.

Note that this property is close to the definition of proximity presented above.

In this paper, we will use a plausibility measure based on the LOF value, i.e.

$$U = \frac{1}{|C_v|}\sum_{x^* \in C_v} LOF(x^*).$$

Note that LOF values are difficult to interpret due to the local nature of the method. Values about 1, says that the point is interior; the higher the value, the more likely it is an outlier. Thus, from the point of view of evaluating the CE algorithm, values close to 1 are preferred.

6. Actionability / feasibility. Finding the closest counterfactual for a data instance does not necessarily result in a feasible change in characteristics. The feasibility of changing a particular variable is described by one of three categories:

- the attribute can be actionable and therefore the attribute is mutable, e.g., balance sheet data;
- the attribute is changeable, but the change is not feasible (e.g., credit rating);
- the attribute is unchangeable (e.g., place of birth).

Remark that the user cannot change the values of the variables of the last two categories, but these values can change because of their ancestors in the causal model [20]. Some authors assume that fulfilling the feasibility requirement automatically guarantees the plausibility of a counterfactual recommendation [22], but despite some overlap, these are different concepts [20]. Feasibility restricts the set of actions to those that can be performed, while plausibility requires that the resulting counterfactual be realistic.

Authors of review articles [5–8] use different taxonomies of CE methods. Here we propose a classification based on the architecture of the models used (*Fig. 1*).

The first group of techniques is based on the solution of an optimization problem in which some of the above properties are treated as a target function and the remaining properties are treated as constraints. For example, in [18], the distance is used as a target $dist(x^*, x)$ with a counterfactual label restriction $h(x^*) = y^*$. This task can be transformed into a problem described by a differentiable function without constraints:

$$x^* \in \arg\min_{x^*}\max_{\lambda} \lambda(h(x^*) - y^*)^2 + dist(x^*, x).$$

The $\lambda(h(x^*) - y^*)^2$ term ensures that the counterfactual label matches the desired class.

Such an approach can be extended to include constraints on actionability, sparsity, data manifold closeness and others, for example [6]:

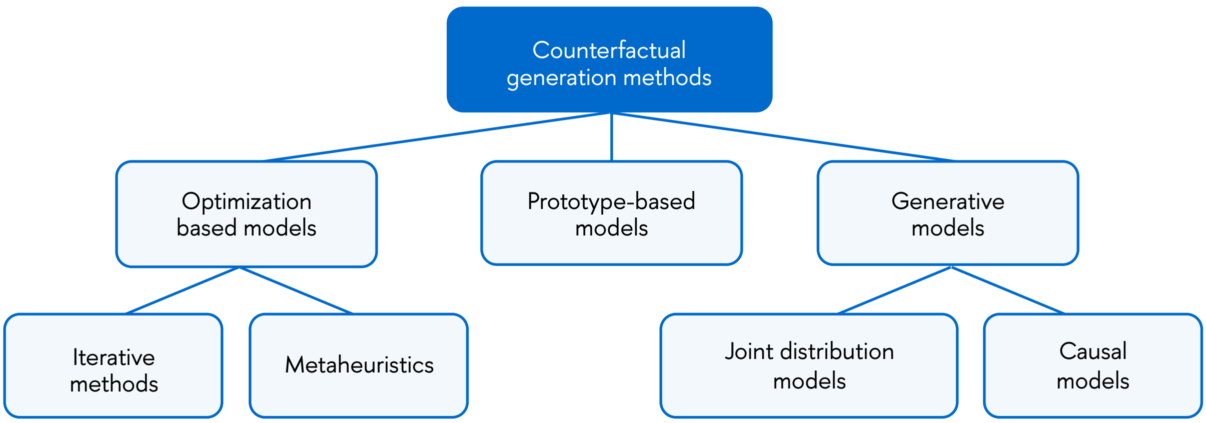

Fig. 1. Classification of counterfactual generation algorithms.

$$x^* \in \arg\min_{x^* \in \mathcal{A}} \max_{\lambda} \lambda \left(h(x^*) - y^*\right)^2 + dist(x^*, x) + \\ + g(x^* - x) + l(x^*; \mathcal{X}).$$

The condition $x^* \in \mathcal{A}$ restricts the list of attributes x^* to ones whose modification is feasible, $g(x^* - x)$ is a penalty function for the difference between the original instance and the counterfactual (e.g., L_0, L_1 norm), and $l(x^*; \mathcal{X})$ is a penalty function for deviation from the data manifold.

Authors of optimization-based CE methods focus first on defining a target function including various metrics for the above properties, and then on choosing an algorithm to find the optimum. It is usually not possible to guarantee the convexity of the target function in this process. Iterative methods of various orders are often used, and metaheuristics (e.g., genetic algorithms) are also widely used. However, this approach requires solving an optimization problem when generating counterfactuals for each new sample. Therefore, [6] recommends authors of such works to cite computation time as one of the algorithm characteristics.

The second group of methods is based on searching in $\mathcal{D}$ for prototypes that will be used to generate counterfactuals [23]. Conceptually, this approach is close to Case Based Reasoning (CBR) [24], which includes four steps: (1) retrieve – extract a case relevant to the problem to be solved, (2) reuse – map the solution found to the problem, (3) revise – test the solution and revise it, if necessary, (4) retain – save the successfully adapted solution.

In particular, an algorithm is proposed in [25], according to which a dataset $\mathcal{D}$ is considered as a set of pairs (x, x^*), where (x, x^*) are the closest objects for which $h(x^*) \neq h(x)$. For a given factual z the closest instances of x, belonging to the same class are found, $h(z) = h(x)$. The attribute values of the counterfactual z^* are initialized with the values from z, then those attributes that differ in x and x^* are changed until z^* is found such that $h(z^*) = h(x^*)$. If this condition is not achieved, the following pair (x, x^*) is used. The idea is that z^* should differ from z in the same way that x^* differs from x.

The third group of CE methods (generative models) is based on modeling the process of data generation. Two types of models can be distinguished in this group: joint distribution modeling and causal models.

A model of the joint distribution $P(\mathcal{X})$ is trained from observations $\mathcal{D}$ and then used to find counterfactuals. As such a model, CE most commonly uses variational autoencoders (VAE), which consist of two parts – an encoder that maps the feature distribution $P(\mathcal{X})$ in $\mathbb{R}^m$ space into the distribution of latent variables $P(\mathcal{Z})$ in a space of lower dimensionality $\mathcal{Z} \subset \mathbb{R}^k$ $(k < m)$, and a decoder that generates the value x' corresponding to the point z' in $P(\mathcal{Z})$. The VAE-based approach offers the interesting prospect of searching for counterfactuals in latent space; in particular, some authors use gradient descent for this purpose [26, 27], but as shown in [28], this is associated with potential problems.

The authors of VAE-based CE methods must consider the above requirements for generating counterfactual explanations, so they introduce additional constraints to the latent representation model. Thus, [29] adapts the traditional scheme, where the encoder is only used to find $P(\mathcal{Z})$ and is not involved in data generation and includes it in the generation process. The encoder is used to find a point z in the latent space corresponding to a given factual x; the counterfactual is generated from the point $z^* = z + \delta$, where δ is a small perturbation. This should enforce the proximity requirement. In addition, the authors of this paper cluster the latent space based on a Gaussian mixture to obtain a conditional distribution $P(\mathcal{Z} \mid \mathcal{I})$, where $\mathcal{I}$ is the set of immutable features.

The authors [28] use a VAE model adapted to find latent variables correlated with class labels [30]. This divides the latent domain into two parts: one for training the representations that predict the labels, and the other for training the rest of the latent representations needed to generate the data. This allows counterfactuals to be generated by modifying only the relevant latent features. The generated examples are then filtered according to causal constraints (e.g., an increase in a borrower's education level must be matched by a corresponding increase in his age).

Note that besides VAE, other models of joint distribution of $P(\mathcal{X})$, can be used, for example statistical models such as copulas and Bayesian networks. However, these techniques are much less frequently used in CE tasks (see reviews of algorithms in [5, 8]). In addition, generative adversarial networks can be applied in some specific cases, such as image analysis tasks [13].

The causal model can be represented as a directed acyclic graph (DAG), which allows for a compact and visual representation of the structure of the system under study [10]. The ability of DAG to encode causal relationships is based on the criterion of d-separation, which corresponds to the conditional independence of variables in the data set. In other words, for any three non-overlapping subsets of variables (X, Y, Z), if nodes X and Y are conditionally independent given Z in the joint distribution $\mathcal{P}$, then they will be d-separated in graph $\mathcal{G}$ (Markov condition): $(X \perp\!\!\!\perp_{\mathcal{P}} Y) | Z \Rightarrow (X \perp\!\!\!\perp_{\mathcal{G}} Y) | Z$. DAG nodes correspond to variables, edges correspond to relationships between them, and the direction of edges corresponds to causal relationships.

DAG corresponds to the structural model $\mathcal{M}$:

$$\mathcal{M} = (\mathbf{S}, P_U), \;\; \mathbf{S} = \left\{X_j := f_j\left(X_{pa(j)}, U_j\right)\right\}_{j=1}^{m},$$
$$P_U = P_{U1} \times \ldots \times P_{Um}.$$

Here $\mathbf{S}$ are structural equations specifying the rules of generation of observed variables X_j as a deterministic function of their ancestors in the causal model $X_{pa(j)} \subseteq \mathrm{X} \setminus X_j$. The assumption of mutual independence of the noises U_j (full factorization of P_U) implies the absence of unobserved confounders, e.g. variables affecting cause and effect simultaneously. Note that many studies assume that noise is additive, i.e., $\mathbf{S} = \left\{X_j := f_j\left(X_{pa(j)} + U_j\right)\right\}_{j=1}^{m}$. This allows one to build efficient algorithms for model identification from data [31].

An important component of causal modeling is the apparatus of do-calculus [10]. For example, an intervention, i.e., the assignment of values $\boldsymbol{\theta}$ to a subset of variables $\mathbf{X}_K$ ($K \subseteq |m|$), is described using the $do(\mathbf{X}_K = \boldsymbol{\theta})$ operator. The distribution of the remaining variables $\mathbf{X}_{-k}$ can be obtained from the system $\mathbf{S}^{do(\mathbf{X}_K = \boldsymbol{\theta})}$, in which the equations for $\mathbf{X}_k$ are replaced by the corresponding values. Thus, the causal model can be used to find counterfactuals [20], for an instance x a counterfactual is defined as $x^* = \mathbf{X}(a)|x$ where $a = do(\mathbf{X}_K = \boldsymbol{\theta})$, $a \in A$, a is an action, and A is the set of admissible actions.

Causal models can be recovered from observed data or constructed from expert knowledge. However, a model $\mathcal{M}$ trained on data may be imperfect, for example, because of sample limitations or, more importantly, because of incorrect specification of the model (i.e., assuming the incorrect parametric form of the structural equations). On the other hand, although in many cases expert knowledge allows the construction of a causal model, assumptions about the form of the structural equations are usually not verifiable [32]. As a result, counterfactual explanations computed based on an ill-defined causal model may be inaccurate and recommend suboptimal or even worse, ineffective actions.

To circumvent these limitations, the authors of [20] propose two probabilistic approaches to selecting optimal actions when there is limited knowledge of causality (e.g., when only the DAG is known). The first one applies to models with additive Gaussian noise and uses Bayesian averaging to estimate the counterfactual distribution. The second excludes any assumptions about the structural equations and instead calculates the average effect of actions on objects that are similar to the factual under consideration.

1.2. Synthetic tabular data generation

Synthetic data generation (SDG) is a core element in solving several machine learning problems: data anonymization, augmentation of small datasets, class balancing in case of severe imbalance, etc. [14].

Definition 3. A synthetic generation model is a function $g \in \mathcal{G}$ that, for an observed data set $\mathcal{D} \sim \mathbb{P}$, returns a data set $\mathcal{D}^S = g(\mathcal{D}, \theta)$ of a given size, $\mathcal{D}^S \sim \mathbb{P}^S$, such that the condition $\mathbb{P}^S \approx \mathbb{P}$, $x_i \neq x_j$, $\forall x_i \in \mathcal{D} \wedge \forall x_j \in \mathcal{D}^S$ is fulfilled. Here θ is a vector of hyperparameters defining the generation policy and $\mathcal{G}$ is a generative function class family.

Mathematically, this can be represented as a Kullback–Leibler distance minimization problem:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \sum_i \mathbb{P}(x_i) \log g(x_i, \theta).$$

Based on this definition, the key performance metric of a generative model is the fidelity of the synthetic data distribution $\mathbb{P}^S$ to the empirical distribution $\mathbb{P}$. Moreover, additional metrics can be introduced [33], such as diversity and generalization. The diversity requirement requires that synthetic instances should cover the entire range of variation $\mathcal{D}$. The generalization property requires that synthetic data should not be copies of real observations.

In this review, we confine our attention to synthetic table (cross-sectional) data generation (tSDG). The following classes of tSDG methods can be distinguished:

- Randomization models based on mixing, interpolation and geometric transformation of the original data and the addition of random noise.
- The probabilistic algorithms that generate data based on a multivariate distribution $\mathbb{P}^S$, modeling the real distribution $\mathbb{P}$. Several approaches can be distinguished here, as follows:
  - modeling of the joint distribution of $\mathbb{P}$, e.g., based on a Gaussian mixture or copulas [15];
  - sequential generation of $\mathcal{D}$ attributes based on conditional distributions $\mathbb{P}(x_i \,|\, \mathcal{D} \backslash \{x_1, ..., x_{i-1}\})$;
  - modeling $\mathbb{P}$ using factorization based on a graphical probability model (Bayesian network) [34].
- Models generating data from lower dimensional latent space.
- Sampling modeling based on generative adversarial networks (GAN).
- Models based on a priori known causal structure.

We remark that the conditional distributions model approach synthesizes the variables x_i sequentially using regression models $x_i = f(x_1, ..., x_{i-1})$, which can be constructed by both parametric (linear regression) and non-parametric (decision tree) methods [35, 36]. Thus, the conditional distributions $\mathbb{P}(x_i \,|\, \mathcal{D} \backslash \{x_1, ..., x_{i-1}\})$, from which the synthetic values of x_i, are derived, are defined for each variable separately and depend on the attributes $x_1, ..., x_{i-1}$, that are earlier in the synthesis sequence. The value of the very first variable in the sequence is generated based on its marginal distribution.

A comprehensive analysis of tSDG methods is presented in [14]. Several publications [16, 17] compare some of the approaches considered on real datasets. From the results presented we can conclude that there is no dominant method, and the quality of generation depends on the specific problem.

It can also be observed that conceptually synthetic data generation methods are close to CE algorithms: both are based on modeling the distribution of observed data but differ in the result. While the objective of CE is to find an instance as close as possible to the sample under study but with the opposite label (see Definition 1), the objective of tSDG is to generate a set of instances that belong to the distribution of the observed data (Definition 3). Accordingly, they are based on different performance metrics.

2. Proposed method

As can be deduced from the review presented above, the known CE algorithms have several limitations. Optimization-based methods require repeated model building for each factual; prototype-based approaches require "factual – counterfactual" pairs in the training set $\mathcal{D}$; generative model-based approaches introduce additional constraints into the algorithm, which also complicates the computation. At the same time, as noted above, synthetic data generation methods are conceptually close to CE, differing only in the result and its evaluation metrics.

Based on these considerations, we propose a two-stage method for generating counterfactual explanations (*Fig. 2*). In the first stage, a model $g(\mathcal{D}, \theta)$ of synthetic data generation is trained. According to Definition 3, this model emulates the empirical distribution $\mathbb{P}$ of real data. Using this model, a set of potential counterfactuals $\{x^*\}_g$ is generated for a given factual x.

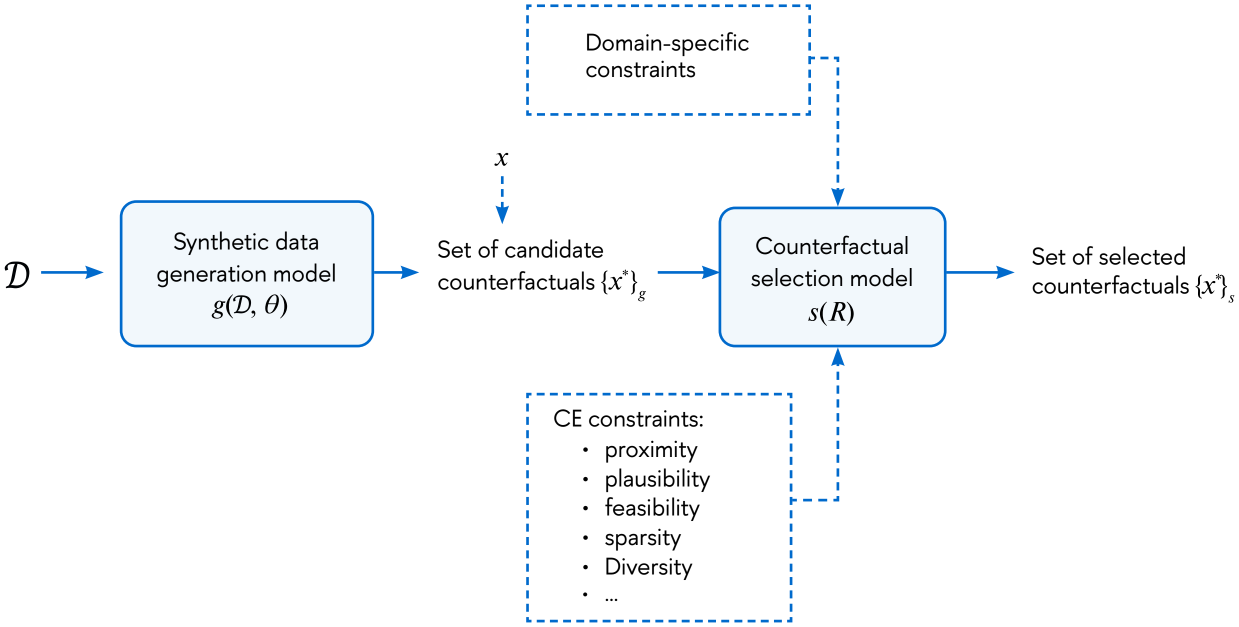

Fig. 2. A two-stage method for generating counterfactual explanations.

In the second step, using the selection model $s(R)$ a set $\{x^*\}_g$ is selected from $\{x^*\}_s$, the elements of which meet the constraints of R. The set $\{x^*\}_s$ is the solution to the CE problem. The selection model can include any constraints formulated as equations of the form $r(c) = m(c) \le v(c),\ \ r \in R$. Here c is a requirement for the result (e.g., validity, proximity, sparsity for CE, or implementation cost), $m(c)$ is the corresponding metric, and $v(c)$ is the boundary of acceptable values. Note that requirements may also include constraints of a particular subject area.

The proposed approach has the following advantages:

- the generative model is built once and allows computing counterfactuals for any new observations without re-training;
- splitting the process into two steps allows the use of simple, easily modifiable selection rules;
- the selection model can include not only the requirements of CE tasks, but also any constraints specific to the subject area under consideration.

3. Experiment

To validate the proposed method, it is first necessary to verify that tSGD methods can generate counterfactuals that satisfy the requirements listed in Section 1.1, and to compare the results with existing known CE methods.

The $g(\mathcal{D}, \theta)$ generation models to be used in the experiment are summarized in *Table* 1. We have selected the simplest statistical models since our goal is to propose an efficient method for generating counterfactuals with low computational cost. These models include the Gaussian copula (GC), a sequential nonparametric model based on conditional distributions (CD), and the Bayesian network (BN), which models the distribution $\mathbb{P}$ as a multiplication of conditional distributions of factors (features). For comparison, we also include a model that generates data based on marginal distributions of features (MD). It can be regarded as a degenerate case of BN in which the relationships between features are not considered. As

Table 1.

Methods for generating synthetic tabular data

ID	Model type	Description	Source
GC	Joint distribution $\mathbb{P}$	Gaussian copula	[15][1]
CD	Conditional distributions $\mathbb{P}(x_i \mid \mathcal{D} \backslash x_i)$	Non-parametric method / decision tree	[36][2]
BN	Factorisation $\mathbb{P} = \prod p(x_i)$	Bayesian network	[34][3]
MD	Marginal distributions x_i	Sampling based on marginal distributions	[38][4]
GAN	Deep learning	Generative adversarial network	[37][5]

mentioned above, such simple models are hardly used in CE tasks, however, we suggest that their potential can be utilized much more effectively using the two-stage approach proposed here.

In addition, as the literature review suggests, most researchers using generative models to solve the CE problem focus on complex algorithms based on deep neural networks, so we also included GAN. We also investigated the possibility of applying VAE, but in our experiments these models did not achieve robust generation of $\{x^*\}_g$. This is most likely due to the insufficient amount of data for training (cf. *Table 2*).

The selection model $s(R)$ is given in the form of a rule:

$$R: h(x^*) \neq h(x) \wedge x_i^* \in \left\{ \overline{x}_i \mp 1.5 \cdot IQR(\mathcal{D}) \right\} \wedge dist(x^*, x) \to$$
$$\to \min \wedge \left| \{x^*\}_s \right| = k.$$

It means that for a particular x, k instances will be selected from the generated set $\{x^*\}_g$ whose label $h(x^*)$ is not equal to the label $h(x)$, the attribute values of x^* are within the range of three $IQR(\mathcal{D})$ interquartile intervals with respect to the mean x_i (Tukey Outlier Definitions), and the distance between x and x^* is minimal.

Table 2 presents the three datasets used in the experiments, their general characteristics and the classification of features in terms of change feasibility (feature changeable, change not feasible, feature not changeable). These public datasets are widely used in machine learning work and, in particular, CE research. The use of public datasets ensures the repeatability of the results.

Table 2 also presents the training results of the classifier h used in the CE finding process: the ROC AUC metric obtained using 10-fold cross-validation and

[1] https://sdv.dev

[2] https://www.synthpop.org.uk/

[3] https://github.com/DataResponsibly/DataSynthesizer

[4] https://github.com/vanderschaarlab/synthcity

[5] https://github.com/NextBrain-ai/nbsynthetic

Table 2.

Data sets

$\mathcal{D}$	$n \times m$	Features			Classifier		Description
		Immutable	The change is not implementable	Mutable	Model $h(x)$	AUC	
German Credit[6]	1000 × 20	3	14	3	RF	0.79 (0.03)	700 approved and 300 rejected loan applications
Adult[7]	48842 × 14	8	3	3	CB	0.93 (0.002)	Income levels based on census data
Loan Default[8]	255347 × 16	8	3	5	CB	0.76 (0.002)	29 653 rejected and 225 694 approved loan applications

the best performing model. In one case it is Random Forest (RF), in all other cases – CatBoost (CB).

One of the most popular libraries implementing CE methods is DiCE [19], which supports three counterfactual search methods. In addition to random search, they are genetic algorithm-based optimization and a method for searching and then adapting prototypes in a training sample [23]. We used these models to comparatively evaluate the results obtained. For each dataset, all three types of models were trained, and the best one was selected. It should be noted that the prototype-based approach failed to find counterfactuals for any dataset. This is obviously due to the limitation noted above: there must be a set of pairs (x, x^*) in $\mathcal{D}$ for a wide range of factuals.

To assess the results of calculating counterfactuals we will use the metrics of validity (V), proximity (P), sparsity (S), diversity (D), and plausibility (U), described above. Specifying the features whose variation is possible is carried out at the level of the generation model $g(\mathcal{D}, \theta)$.

4. Analysis of experimental results

Consider the process of applying the proposed method on the example of the German Credit dataset. This dataset contains records of 1 000 credit applications, 700 of which were approved. The attributes include the amount and term of the loan, as well as indicators of the borrower's social and financial status (credit rating, duration of employment, proportion of loan payments in the borrower's total income, etc.). Most attributes are either categorical or ordinal.

The task of CE in this case is to generate counterfactuals for borrowers who have been denied a loan. The data analysis shows that the attributes that are modifiable are *laufzeit* – loan term in months, *hoehe* – loan amount and *buerge* – presence of a co-borrower or guarantor. All other attributes are either not changeable (gender, citizenship) or cannot be changed by direct influence (credit rating).

[6] South German Credit (2019) UCI Machine Learning Repository. https://doi.org/10.24432/C5X89F
[7] Becker B., Kohavi R. (1996) Adult. UCI Machine Learning Repository. https://doi.org/10.24432/C5XW20
[8] Loan Default Dataset. https://www.kaggle.com/datasets/nikhil1e9/loan-default/data

The computation is performed according to the method presented in *Fig. 2*. At the first stage, the generation model $g(\mathcal{D}, \theta)$ is trained, which is used to generate 200 synthetic instances for the investigated factual. From this set, instances are selected according to the rule given by equation (1).

Table 3 shows an example of the data generated for a rejected application of DM 2348 for a term of 36 months. As can be seen from the data presented, the loan for this borrower can be approved if the term is reduced to 8 months and the amount to DM 1 956. If the borrower presents a co-borrower (*buerge* = 2), the amount can be increased to DM 2234 for a period of 14 months. If there is a guarantor (*buerge* = 3), the loan can be DM 4276 for a period of 26 months. As we can see, even the 3 presented counterfactuals allow us to describe the situation for a particular borrower and suggest a possible way for the borrower to achieve his goal.

Table 3.

An example of the generated data (attributes are explained in the body of the text)

	laufzeit	*hoehe*	*buerge*	Class label
Factual	36	2384	1	0
Counterfactuals	8	1956	1	1
	14	2234	2	1
	26	4276	3	1

Table 4.

Average values and standard deviations of model quality metrics for the three datasets

Model	*V*	*P*	*S*	*D*	*U*
BN	**1.000 (0.000)**	**1.230 (0.805)**	3.790 (1.251)	1.310 (0.340)	**2.053 (0.855)**
CD	**1.000 (0.000)**	1.400 (1.155)	3.303 (1.036)	1.135 (0.228)	2.128 (0.836)
DiCE	0.869 (0.273)	3.138 (1.519)	**1.962 (0.556)**	1.159 (0.206)	2.797 (0.734)
GAN	0.966 (0.186)	1.984 (1.045)	4.190 (1.640)	1.285 (0.372)	2.298 (0.742)
GC	0.967 (0.183)	1.333 (0.851)	3.887 (1.521)	1.284 (0.402)	2.089 (0.828)
MD	**1.000 (0.000)**	3.198 (2.094)	4.177 (1.506)	**1.314 (0.299)**	2.851 (0.760)

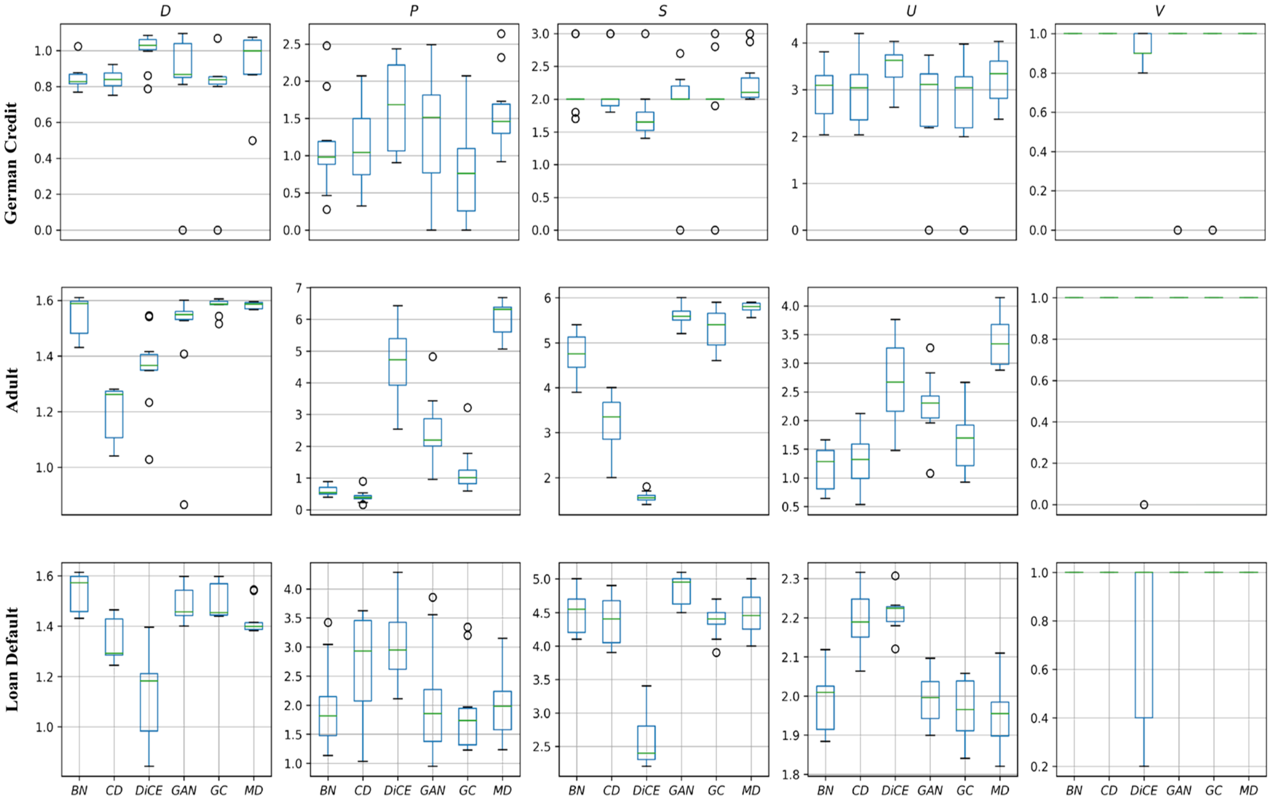

Fig. 3. Performance metrics of the models considered by datasets.

Table 4 lists the averages and standard deviations of the quality metrics for the considered methods computed over all three datasets, and *Fig. 3* presents distributions of metrics across datasets. The best values of the metrics in *Table 4* are in bold type.

It should be mentioned that for most metrics (validity, closeness, and plausibility), the best results are demonstrated by the Bayesian network (BN) model, which generates samples based on conditional distributions of features, i.e., considering the dependencies between them. Considering that this model is only slightly inferior to MD in terms of diversity, the choice of BN for CE seems quite justifiable. The high diversity of counterfactuals generated by MD is because this model considers only marginal distributions of features and does not consider the relationships between them. This model should work well in the case of uncorrelated features but may generate challenges when such correlations are present (see distribution *D* for the Loan Default dataset in *Fig. 3*). On the contrary, BN performs the best in terms of diversity among all models.

In our experimentation, the most sophisticated GAN model lost out to other models, possibly because there was not enough data to train, although the authors of the implementation we used [37] emphasize that it focuses specifically on small training samples. *Figure 3* shows that as the sample size increases, the GAN results improve but do not outperform the other models.

The methods developed directly for the CE problem (DiCE) showed the best result on the sparsity metric (*Table 4*), but *Fig. 3* shows that this was achieved by performing well on the largest dataset

(Loan Default). On smaller data, this method is inferior to simpler models, in particular GC and BN. Furthermore, it should be noted that DiCE on all datasets fails to find the required number of counterfactuals ($V = 1$) and in this sense this is the worst of the methods considered.

Conclusion

Therefore, we can conclude that the proposed method of counterfactual search based on synthetic data generation can achieve results at least comparable to the "standard" CE methods, and in some cases, it outperforms them, especially on small datasets. According to our results, the most obvious choice in this case is a generation model based on a Bayesian network that considers the interconnections between attributes.

This result reveals new possible research directions. The Bayesian network is a statistical model because it is built on associations measured by correlations. Therefore, it is of interest to study causal models that capture causal relationships in a dataset.

It should be noted, however, that to the best of our knowledge, the direction related to the use of causal models for CE is only beginning to be explored [20], and there are no works devoted to their application to the generation of synthetic data. ■

References

1. Samek W., Muller K.-R. (2019) Towards explainable artificial intelligence. *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning. Lecture Notes in Computer Science*, vol. 11700, pp. 5–22. https://doi.org/10.1007/978-3-030-28954-6_1

2. Giuste F., Shi W., Zhu Y., Naren T., Isgut M., Sha Y., Tong L., Gupte M., Wang M.D. (2023) Explainable artificial intelligence methods in combating pandemics: A systematic review. *IEEE Reviews in Biomedical Engineering*, vol. 16, pp. 5–21. https://doi.org/10.1109/RBME.2022.3185953

3. Barocas S., Selbst A. D., Raghavan M. (2020) The hidden assumptions behind counterfactual explanations and principal reasons. Proceedings of the *2020 Conference on Fairness, Accountability, and Transparency (FAT*'20)*, pp. 80–89. https://doi.org/10.1145/3351095.3372830

4. Murdoch W.J., Singh C., Kumbier K., Abbasi-Asl R., Yu B. (2019) Definitions, methods, and applications in interpretable machine learning. *National Academy of Sciences*, vol. 116(44), pp. 22071–22080. https://doi.org/10.1073/pnas.1900654116

5. Guidotti R. (2022) Counterfactual explanations and how to find them: Literature review and benchmarking. *Data Mining and Knowledge Discovery*. https://doi.org/10.1007/s10618-022-00831-6

6. Verma S., Boonsanong V., Hoang M., Hines K. E., Dickerson J. P., Shah C. (2020) Counterfactual explanations and algorithmic recourses for machine learning: A review. *arXiv:2010.10596*. https://doi.org/10.4550/arxiv.2010.10596

7. Stepin I., Alonso J.M., Catala A., Pereira-Fariña M. (2021) A survey of contrastive and counterfactual explanation generation methods for explainable artificial intelligence. *IEEE Access*, vol. 9, pp. 11974–12001. https://doi.org/10.1109/ACCESS.2021.3051315

8. Chou Y.L., Moreira C., Bruza P., Ouyang C., Jorge J. (2022) Counterfactuals and causability in explainable artificial intelligence: Theory, algorithms, and applications. *Information Fusion*, vol. 81, pp. 59–83. https://doi.org/10.1016/j.inffus.2021.11.003

9. Mishra P. (2022) *Practical explainable AI using Python: Artificial Intelligence model explanations using python-based libraries, extensions, and frameworks*. Apress.

10. Pearl J. (2009) *Causality: models, reasoning, and inference*. *2nd ed.* New York: Cambridge University Press.

11. Cho S.H., Shin K.S. (2023) Feature-weighted counterfactual-based explanation for bankruptcy prediction. *Expert Systems with Applications*, vol. 216, article 119390. https://doi.org/10.1016/j.eswa.2022.119390

12. Wang D., Chen Z., Florescu I., Wen B. (2023) A sparsity algorithm for finding optimal counterfactual explanations: Application to corporate credit rating. *Research in International Business and Finance*, vol. 64, article 101869. https://doi.org/10.1016/j.ribaf.2022.101869

13. Mertes S., Huber T., Weitz K., Heimerl A., André E. (2022) Ganterfactual – counterfactual explanations for medical non-experts using generative adversarial learning. *Frontiers in Artificial Intelligence*, vol. 5, article 825565. https://doi.org/10.3389/frai.2022.825565

14. Fonseca J., Bacao F. (2023) Tabular and latent space synthetic data generation: A literature review. *Journal of Big Data*, vol. 10(1), article 115. https://doi.org/10.1186/s40537-023-00792-7

15. Patki N., Wedge R., Veeramachaneni K. (2016) The synthetic data vault. Proceedings of the *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 399–410. https://doi.org/10.1109/DSAA.2016.49

16. Dankar F., Ibrahim M., Ismail L. (2022) A multi-dimensional evaluation of synthetic data generators. *IEEE Access*, vol. 10, pp. 11147–11158. https://doi.org/10.1109/ACCESS.2022.3144765

17. Endres M., Mannarapotta Venugopal A., Tran T.S. (2022) Synthetic data generation: A comparative study. Proceedings of the *26th International Database Engineered Applications Symposium*, pp. 94–102. https://doi.org/10.1145/3548785.3548793

18. Wachter S., Mittelstadt B., Russell C. (2017) Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harvard Journal of Law & Technology (Harvard JOLT)*, vol. 31, article 841.

19. Mothilal R.K., Sharma A., Tan C. (2020) Explaining machine learning classifiers through diverse counterfactual explanations. Proceedings of the *2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*, pp. 607–617. https://doi.org/10.1145/3351095.3372850

20. Karimi A.H., Barthe G., Schölkopf B., Valera I. (2023) A survey of algorithmic recourse: Contrastive explanations and consequential recommendations. *ACM Computing Surveys*, vol. 55(5), article 95. https://doi.org/10.1145/3527848

21. Breunig M.M., Kriegel H.-P., Ng R.T., Sander J. (2000) LOF: identifying density-based local outliers. Proceedings of the *2000 ACM SIGMOD International Conference on Management of Data (ICDM)*, pp. 93–104. https://doi.org/10.1145/335191.335388

22. Poyiadzi K., Sokol K., Santos-Rodriguez R., De Bie T., Flach P. (2020) FACE: feasible and actionable counterfactual explanations. Proceedings of the *2020 AAAI/ACM Conference on AI, Ethics, and Society (AIES 2020)*, pp. 344–350. https://doi.org/10.1145/3351095.3372850

23. van Looveren A., Klaise J. (2021) Interpretable counterfactual explanations guided by prototypes. Proceedings of the *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2021)*, pp. 650–665. https://doi.org/10.1007/978-3-030-86520-7_40

24. Aamodt A., Plaza E. (1994) Case-based reasoning: Foundational issues, methodological variations, and system approaches. *Artificial Intelligence Communications*, vol. 7(1), pp. 39–59.

25. Keane M.T., Smyth B. (2020) Good counterfactuals and where to find them: A case-based technique for generating counterfactuals for explainable AI (XAI). Proceedings of the *28th International Conference on Case-Based Reasoning Research and Development (ICCBR)*, pp. 163–178. https://doi.org/10.1007/978-3-030-58342-2_11

26. Joshi S., Koyejo O., Vijitbenjaronk W., Kim B., Ghosh J. (2019) Towards realistic individual recourse and actionable explanations in black-box decision making systems. *arXiv:1907.09615*. https://doi.org/10.48550/arXiv.1907.09615

27. Guyomard V., Fessant F., Bouadi T., Guyet T. (2021) Post-hoc counterfactual generation with supervised autoencoder. Proceedings of the *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2021)*, pp. 105–114. https://doi.org/10.1007/978-3-030-93736-2_10

28. Downs M., Chu J.L., Yacoby Y., Doshi-Velez F., Pan W. (2020) CRUDS: Counterfactual recourse using disentangled subspaces. Proceedings of the *2020 ICML Workshop on Human Interpretability in Machine Learning (WHI 2020)*, pp. 1–23.

29. Pawelczyk M., Broelemann K., Kasneci G. (2020) Learning model-agnostic counterfactual explanations for tabular data. Proceedings of the *Web Conference 2020 (WWW'20)*, pp. 3126–3132. https://doi.org/10.1145/3366423.3380087

30. Klys J., Snell J., Zemel R. (2018) Learning latent subspaces in variational autoencoders. *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*.

31. Hoyer P., Janzing D., Mooij J.M., Peters J., Schölkopf B. (2008) Nonlinear causal discovery with additive noise models. *Advances in Neural Information Processing Systems 21 (NIPS 2008)*.

32. Peters J., Janzing D., Schölkopf B. (2017) *Elements of causal inference: foundations and learning algorithms*. MIT press.

33. Alaa A., van Breugel B., Saveliev E.S., van der Schaar M. (2022) How faithful is your synthetic data? Sample-level metrics for evaluating and auditing generative models. Proceedings of the *39th International Conference on Machine Learning*, pp. 290–306.

34. Ping P., Stoyanovich J., Howe D. (2017) DataSynthesizer: Privacy-preserving synthetic datasets. Proceedings of the *29th International Conference on Scientific and Statistical Database Management (SSDBM'17)*. https://doi.org/10.1145/3085504.3091117

35. Drechsler J., Reiter J.P. (2011) An empirical evaluation of easily implemented nonparametric methods for generating synthetic datasets. *Computational Statistics & Data Analysis*, vol. 55(12), pp. 3232–3243. https://doi.org/10.1016/j.csda.2011.06.006

36. Nowok B., Raab G.M., Dibben C. (2016) Synthpop: Bespoke creation of synthetic data in R. *Journal f Statistical Software*, vol. 74, pp. 1–26. https://doi.org/10.18637/jss.v074.i11

37. Marin J. (2022) Evaluating synthetically generated data from small sample sizes: An experimental study. *arXiv:2211.10760*. https://doi.org/10.48550/arXiv.2211.10760

38. Qian Z., Cebere B.C., van der Schaar M. (2023) Synthcity: facilitating innovative use cases of synthetic data in different data modalities. *arXiv:2301.07573.* https://doi.org/10.48550/arxiv.2301.07573

About the authors

Yuri A. Zelenkov

Dr. Sci. (Tech.);

Professor, Department of Business Informatics, Graduate School of Business, HSE University, 28/11, Shabolovka Str., Moscow 119049, Russia;

E-mail: yzelenkov@hse.ru

ORCID: 0000-0002-2248-1023

Elizaveta V. Lashkevich

Doctoral Student, Department of Business Informatics, Graduate School of Business, HSE University, 28/11, Shabolovka Str., Moscow 119049, Russia;

E-mail: evlashkevich@hse.ru

ORCID: 0000-0002-3241-2291

DOI: 10.17323/2587-814X.2024.3.41.55

Hidden Markov model: Method for building a business process model*

Artem Yu. Varnukhov

E-mail: varnuhov_ayu@usue.ru

Ural State University of Economics, Ekaterinburg, Russia

Abstract

More and more companies are influenced by the rapid development of technology (Industry 4.0/5.0 concept), are embracing digital transformation processes. The introduction of information systems makes it possible to accumulate a large amount of data about the company's activities. Study of such information expands the opportunities for applying a data-driven approach to business process management (BPM). Processing and studying data from event logs using process mining methods make it possible to build digital models of business processes which turn out to be a useful source of information when carrying out analysis, modeling and reengineering within the framework of the process approach. In this paper, we develop a method for building a business process model based on a hidden Markov model, taking into account the restrictions imposed by the subject area. The use of a hidden Markov model allows us to use the apparatus of probability theory and mathematical statistics to analyze business processes, as well as to solve classification and clustering problems. This article describes the capabilities of a data-driven approach to business process management and demonstrates examples of the practical application of the method to solve business challenges: drawing a dependency graph that can be used to identify discrepancies between actual and expected execution, as well as a method for predicting the outcome of a business process based on the sequence of observed events.

* The article is published with the support of the HSE University Partnership Programme

Keywords: business processes, hidden Markov models, process mining, business analysis, prediction, classification, data-driven approach, information systems, event logs

Citation: Varnukhov A.Yu. (2024) Hidden Markov model: Method for building a business process model. *Business Informatics,* vol. 18, no. 3, pp. 41–55. DOI: 10.17323/2587-814X.2024.3.41.55

Introduction

The development of the capabilities of modern information technologies stimulates enterprises in various fields to transfer their business processes from "analog" to digital form. There are many methodologies and techniques that enable us to model, reengineer, manage and monitor business processes [1], and they are constantly being improved. Quite often, modeling is performed "manually" with the involvement of appropriate business analysts and "in-house" experts who have specialized knowledge and expertise with respect to the phenomena being modeled. At the same time, in practice, the process of modeling and reengineering business processes turns out to be a non-trivial task even for experienced specialists [2]. For example, distortions occur due to the human factor, one's own position in the organization's structure and other typical issues characteristic of this modeling approach: idealization, choosing the wrong level of abstraction, or the inability to adequately reproduce the observed interaction [3]. As a result, the generated model may reflect only part of the existing "reality" and it is not functional enough so that ultimately it will have very limited value.

The implementation of automated information systems of various classes and functionality (ERP, CRM, ECM, etc.) leads to the concomitant accumulation of a large amount of useful information about the activities of the enterprise in a data warehouse [4]. Processing and subsequent analysis of data accumulated in the enterprise information systems make it possible to use a data-driven approach. Currently, research is being conducted in the field of process data mining [5, 6], creating digital twins [7], predictive and prescriptive analytics [8, 9], robotic process automation [10], and work is also being carried out on practical implementation in various industries [11, 12].

The purpose of this study is to develop a method for building a business process model based on a hidden Markov model, taking into account the characteristics of the subject area.

1. Application of a data-driven approach to business process management

The process approach allows us to present a company as a set of interconnected business processes, each process is considered as a valuable asset that ensures the delivery of the company's products and services to end consumers. The business process management (BPM) methodology defines the management life cycle, which typically consists of the following main stages: analysis, modeling, execution, monitoring, optimization and reengineering. To formulate the capabilities and context of using a data-driven approach within BPM, we can build a generalized management life cycle as shown in *Fig. 1.*

As can be seen from *Fig. 1*, process models are a source of information for analysis and optimization. They help us at the stage of information systems implementation and contribute to the management and control functions. Data accumulated in information systems can be used to make digital models that will allow for a better understanding of the organization's existing business processes. The use of a data-driven approach provides a capability to establish a

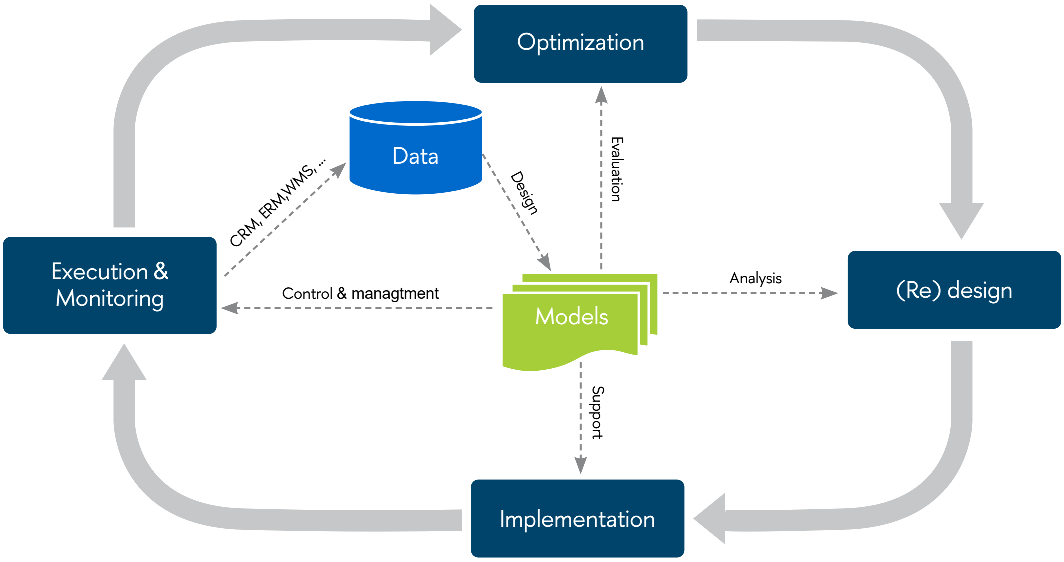

Fig. 1. Generalized BPM life cycle.

close relationship between existing processes and their representation in the form of models. We can single out a few key applications among many options. They include model identification, compliance testing, performance evaluation and process improvement. To identify models, data from event logs of information systems are examined and, using special methods of intelligent analysis, models are built without involving any a priori information. We can use the AS-IS models obtained in this way in further work when carrying out analysis, modeling and reengineering. It is worth noting that the design stage is crucial, since it provides a kind of "entry" point for all other tasks employing the identified digital models. To check compliance, a previously designed reference model of a business process and data from the event log are used: they are compared with each other to determine the degree of compliance. Such verification is useful for monitoring compliance with imposed rules and restrictions, detecting discrepancies between actual and expected performance, finding reasons for deviations, and so on. If we take into account the presence of a time component in the data, then using models it is possible to measure business process performance indicators, detect bottlenecks, evaluate the level of service, etc. For example, variant analysis will reveal differences in control flow and performance indicators between different organizations' departments. In addition to studying the control flow, we can expand the model by including an organizational component. This will allow us to take into account information about the process participants and their relationships. Thus, the use of a data-driven approach allows us to improve the quality and efficiency of business process management.

2. Statement of the problem for building a business process model

If a business process can be represented as a model, then a specific individual case implemented within this model can be described as its instance. An instance of a business process must be characterized by a certain set of sequential or parallel actions (activities) with the capability to determine the order in which they occur. Different instances shall be distinguishable from each other at least by the order of events. *Table 1* shows an example of an event log fragment obtained from the information system.

Table 1.

Event log fragment

Instance ID	Event ID	Timestamp		Event	Employee	...
1001	24837	24.08.2023	13:20	Receipt of request	Ivanov A.	…
1001	25123	25.08.2023	11:05	Availability check	Petrova I.	…
1001	26001	26.08.2023	09:15	Sending an invoice	Ivanov A.	…
1001	26560	27.08.2023	16:07	Shipment of goods	Sidorov V.	…
1002	24842	24.08.2023	14:27	Receipt of request	Ivanov A.	…
1002	24859	24.08.2023	16:20	Availability check	Petrova I.	…
1002	24892	24.08.2023	17:40	Refusal to deliver	Sobolev B.	…
…	…	…		…	…	…

It is assumed that all data in the log relates to one analyzed business process. Each line in the table contains the following mandatory attributes: Instance ID, Event and Timestamp. Multiple lines with the same Instance ID attribute value represent events that are associated with a single business process instance. The Event attribute contains the name of the event, which can be associated with some action (activity). The Timestamp attribute is used to chronologically arrange events within a single instance. The event log may contain other additional attributes (Employee, Cost, Customer, Office, etc.), which can be useful for monitoring a business process using machine learning [13]. For brevity, we will use a multiset, which will consist of chronologically ordered and grouped sequences of events according to the business process log. For example, for the data given in *Table 1*, we can record a multiset:

$$L = \{< a, b, c, d >^{n}, < a, b, e >^{m}, ...\}, \tag{1}$$

where L is a multiset in which each element contains an ordered sequence of events: a — "Receipt of request" event, b — "Availability check" event, c — "Sending an invoice" event, d — "Shipment of goods" event, e — "Refusal to deliver"; n and m are the number of times this ordered sequence occurred in the log.

Thus, we need to develop a method for building a business process model based on incoming input data in the form of multiset L.

3. Analysis of methods for building a business process model

"α-algorithm". It is quite simple and one of the first methods that allows us to recreate a business process model from an existing set of sequential events in the form of a Workflow-net (an individual case of a Petri net) [14]. To do this, the algorithm scans the log searching for a specific set of patterns: sequence, XOR-split and AND-split. Based on this, a matrix of "fingerprints" is recorded, enabling recognition of the existing relationships between events. The final model is built according to this matrix, taking into account the inference rules. We can name the following limitations of the "α-algorithm": difficulties in processing noisy data, inability to recognize 1- or 2-step cycles and problems with local dependencies.

"Heuristic Miner". Unlike the "α-algorithm", it applies the idea of counting the frequencies of occurrence of events and reproduces the process model in the form of a Causal net [15, 16]. First, metrics are calculated that estimate the number of direct connections between each pair of events and measure the degree of their dependence. A dependency graph is drawn using (sequence, XOR, AND, and cycle) patterns based on the calculated metrics. The search for mergers and splits in the dependency graph can be performed by a sliding window over an event log of a given size or based on solving an optimization task in which the target function is the compliance degree of the model to the analyzed event log. The resulting process model in the form of Causal net can be converted to other required notations (BPMN, UML, EPC, WF-net, etc.). This method is less susceptible to noise in the data and eliminates many of the shortcomings of the "α-algorithm", but has problems handling non-local dependencies and detecting duplicate events, and also requires manual adjustment of cutoff threshold levels.

"Region-Based Miner". It is based on the application of the theory of regions and is built according to the assumption that state models can be transformed into Petri nets [17]. There are several approaches to implement this method. The first approach consists in defining a region as a set of states such that the actions in the state and transition model are consistent with this region. In this case, all events can be divided into "incoming", "outgoing" and "internal" in relation to this region. After dividing the regions according to these rules, each region can be associated with a specific position in the Petri net. The second approach uses a specially defined language model instead of a system of states and transitions [18]. The main idea of this approach is that removing a P_i position does not remove any behavioral pattern, but adding a new position may lead to the elimination of some possible behavior options. Advantages include the capability to handle more complex control flow structures. The weakness of this method is the inability to detect some types of process designs, issues with accuracy and generalization ability, and the difficulty of its practical implementation.

"Inductive Miner". It consists of three recursive steps: drawing an oriented graph, searching for a cut, and splitting log entries [19]. The method uses a pre-processed event log as input data. In the first step, the method transforms the data into an oriented graph in which each node corresponds to one event, and the arcs form transitions between events. After this, an attempt is made to detect places of possible cuts. If places of such cuts are detected, then the algorithm generates a cut operator and split segments. Based on the detected segments, the log is decomposed into smaller components. Each fragment thus obtained is then processed recursively until the base case is found: the fragment contains only one event. If, in the process of recursive descent, a fragment is encountered that is not reducible to the base case and at the same time does not have valid places for cutting, then the process of "falling through" is applied. The basic implementation of the method had difficulties detecting fixed-length cycles, handling rare events and limitations associated with the recursive nature of the design. However, further development of the method made it possible to overcome the primary disadvantages and provided the capability of scaling and using distributed computing [20].

The methods presented above allow us to build a business process model in various ways, however, it is of interest to study the capability of building a model that is based on probability assessment. As a basis, we could consider a Markov chain, but given the characteristics and nature of the source data, it would be more acceptable to assume that the events recorded in the log are only the external manifestation of some process hidden from the observer. To model such an assumption, we can consider a first-order hidden Markov model (hereinafter referred to as HMM). It is known that such models with hidden states are successfully applied for text processing tasks in natural languages [21], gesture identification [22], speech recognition [23], bioinformatics [24] and other fields. Based on the information presented in the analyzed sources, it can be expected that the HMM use for business process analysis will make it possible not only to build a business process model, but also to solve the problem of classification and perform data clustering.

4. Suggested method for building a model

The HMM possesses multiple hidden states $S = \{s_1, s_2, s_3, ..., s_N\}$. Each hidden state can be associated with some other hidden states. A schematic representation of the model is shown in *Fig. 2*.

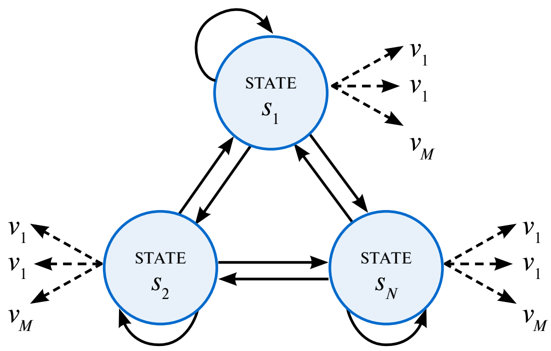

Fig. 2. The HMM schematic representation.

This paper considers a fully connected HMM structure, in which each hidden state s_k is associated with all hidden states different from it, as well as with itself. In addition to the hidden states, a finite alphabet of multiple observed events $V = \{v_1, v_2, v_3, ..., v_M\}$ is defined and each hidden state reproduces events from a given set V. At any individual time t, the model is in one of the hidden states:

$$\forall\, t : q_t \in S,\ 1 \le t \le T. \tag{2}$$

The HMM makes transitions between hidden states. So at time t, being in the hidden state q_t, the model will move to another state with a certain probability and at time $t + 1$ will be in the hidden state $q_{t+1} \in S$. In this paper we consider only discrete instants of time, while the current state and the chain of completed transitions between them are invisible to the observer. Being in some hidden state q_t, the model reproduces the event $o_t \in V$, which is visible to an external observer. The series of transitions between states and the events they reproduce as a result generates a sequence of observations $O = \{o_1, o_2, o_3, ..., o_T\}$. A schematic representation of the HMM operation is shown in *Fig. 3*.

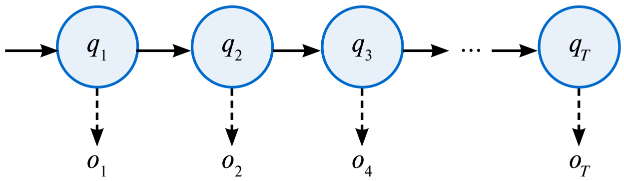

Fig. 3. Schematic representation of the HMM operation.

Since this work considers a first-order HMM, according to the Markov property we will assume that the probability of transition from one state to another is determined only by the previous state of the model:

$$P(q_t = s_i | q_1 = s_n, q_2 = s_m, \ldots, q_{t-1} = s_j) = \\ = P(q_t = s_i | q_{t-1} = s_j). \tag{3}$$

The second assumption will be that the probability of producing the observed event o_i depends only on the state in which the model is at a discrete instant of time t and does not depend on other states and observed events:

$$P\left(o_t = v_k | q_1 = s_n, \ldots, q_t = s_i, o_1 = v_m \ldots, o_{t-1} = v_j\right) = \\ = P\left(o_t = v_k | q_t = s_i\right). \tag{4}$$

Let us determine the initial distribution over the hidden states of the model, which specifies the probability that the model will be in a certain state at the first step:

$$\pi = \{\pi_i\}_{i=1}^{N},\ \pi_i = P(q_1 = s_i),\ \sum_{i=1}^{N} \pi_i = 1. \tag{5}$$

Let us define the distribution of transition probabilities between hidden states as matrix $A = (a_{ij})$, where

$$a_{ij} = P(q_t = s_j | q_{t-1} = s_i),\ 1 \le i, j \le N,\ \sum_{j=1}^{N} a_{ij} = 1. \tag{6}$$

We define the probability distribution of events occurring when the model is in some hidden state as matrix $B = (b_{ik})$, where

$$b_{ik} = P(o_t = v_k | q_t = s_i),\ 1 \le i \le N,\ 1 \le k \le M,$$
$$\sum_{k=1}^{M} b_{ik} = 1. \qquad (7)$$

Based on the above, we define the hidden Markov model θ as

$$\theta = (S, V, A, B, \pi). \qquad (8)$$

Let us assume that the original log is presented and some data pre-processing has been completed. Suppose multiset is

$$L = \{<a,e>^5, <a,b,c,e>^5, <a,c,b,e>^5,$$
$$<a,d,e>^{10}, <a,d,d,e>^5, <a,d,d,d,e>^1,$$
$$<a,b,c,d,e>^3, <a,b,d,c,e>^3, <a,c,b,d,e>^3,$$
$$<a,c,d,b,e>^3, <a,d,c,b,e>^3, <a,d,b,c,e>^3,$$
$$<e,b,c,m>^4, <e,c,b,m>^3\}. \qquad (9)$$

Multiset L contains elements repeated several times, which represent individual instances of a business process performed at different time. It can be observed that some elements of the multiset (for example, <*a, b, c, e*> and <*a, c, b, e*>) contain almost identical sequences of events, except that the order in which events "*b*" and "*c*" follow is rearranged. An apparent rearrangement in the multiset may occur due to the fact that the registered events appear in the source log ordered by time stamp, but in reality, they represent subprocesses of a business process running in parallel. An example of such a situation is shown in *Fig. 4*.

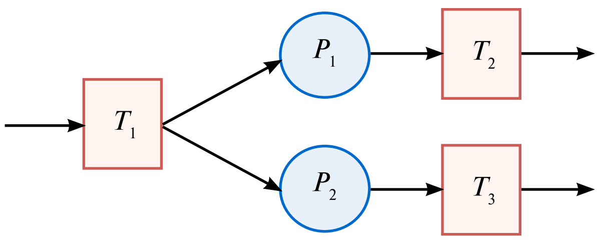

Fig. 4. Subprocesses run in parallel.

Due to the fact that such events are always pairwise and follow each other (albeit in a different order), and also taking into account the nature of the modeled subject area, we will assume that such rearranged events generate a logically single operation of the business process. Thus, we will refer similar pairwise permutations of events, which by their behavior form the logical AND operator of a business process, to the same hidden state of the model. Generally, a business process has one fixed start event at which its execution begins. A business process may have several end events due to the need to present different outcomes. Each such end event is logically final and therefore should not be divisible into several subprocesses. If subprocesses that run in parallel are not in the start or end events, then it should be assumed that there is some event after which the execution branches, as well as an event connecting the parallel execution. Thus, the logical AND operator of a business process shall be located between some start and end event in the observed sequence of events. To determine groups of events that form a set of logical operations AND of a business process, we choose unique elements according to rule (10) from multiset L:

$$FSET = \left\{ \begin{array}{l} \big(\sigma(i) = S_1, \sigma(i+1) = S_2, \sigma(i+2) = \\ = S_3, \sigma(i+3) = S_4\big): \\ \exists\big(\delta(l) = S_1, \delta(l+1) = S_3, \delta(l+2) = \\ = S_2, \delta(l+3) = S_4\big) \end{array} \right\}$$
$$\forall \sigma, \delta \in L : |\sigma| \ge 4,\ |\delta| \ge 4,$$
$$1 \le i \le |\sigma|,\ 1 \le l \le |\delta|. \qquad (10)$$

That is, the members of the *FSET* set are ordered sequences of events, each of which satisfies the following conditions:

- is obtained from elements included in original multiset L;
- the sequence contains at least four consecutive events;
- there is another quadruplet with the same start event S_1 and end event S_4 in which places of events S_2 and S_3 are interchanged.

For the considered multiset from (9), set *FSET* will be generated consisting of the following elements:

$$FSET(L)=\{(a,b,c,e),(a,c,b,e),(a,b,c,d),\\ (b,c,d,e),(a,b,d,c),(b,d,c,e),(a,c,b,d),(c,b,d,e),\\ (a,c,d,b),(c,d,b,e),(a,d,c,b),(d,c,b,e),\\ (a,d,b,c),(d,b,c,e),(e,b,c,m),(e,c,b,m)\}. \quad (11)$$

Each element of the set *FSET* represents the minimum acceptable part of a possible permutation. So, for example, to form a logical AND operator, composed of two subprocesses "*b*" and "*c*" run in parallel, which starts after the event "*a*" and ends with the event "*e*", it is necessary that the set *FSET* contains both parts of such a permutation (*a, b, c, e*) and (*a, c, b, e*). The elements of the set (*a, b, d, c*) and (*a, d, b, c*) also generate the minimal logical AND operator. However, it can be seen from (9) that the elements "*b*" and "*d*" are part of a larger logical AND operator, which also includes the event "*c*". Thus, it is necessary to define a procedure for growing longer logical AND operators, consisting of basic minimal parts. For this purpose, we define sets of start, permutable and end events:

$$FS=\{\,\sigma(1):\sigma\in FSET\,\}. \quad (12)$$

$$PS=\{\,\sigma(2):\sigma\in FSET\,\}. \quad (13)$$

$$ES=\{\,\sigma(4):\sigma\in FSET\,\}\setminus PS. \quad (14)$$

To build up the maximum possible permutation, we take each start event from the set *FS* one by one and, going through the set of *FSET* elements, we will add each permutable symbol encountered in the second and third positions until we reach the element that contains the end event from *ES* in the last position. The procedure for display augmentation is shown in *Listing 1.*

As a result of the augmentation procedure, a multiset will be generated that contains elements with selected start and end states, as well as a set of permutable events between them. For example, a multiset will be formed from (9):

$$FPRM(L)=\{<e,\{b,c\},m>^2,<a,\{b,c\},e>^2,\\ <a,\{b,c,d\},e>^6,<b,\{c,d\},e>^2,<c,\{b,d\},e>^2,\\ <d,\{b,c\},e>^2\}. \quad (15)$$

The multiplicity of elements in multiset *FPERM* reflects the frequency with which a given permutation occurred in the source data. Let us assume that the source data contained a group of events "*b*", "*c*", "*d*", which forms the logical AND operator of the business process. Then the multiplicity of such a group with the same start and end events must be equal to six. Suppose that this group is missing one element, for example (*a, c, b, d, e*). In this case, the group of events should split into two logical AND operators forming parallel sub-

```
Input: FS: Start events
Input: ES: End events
Input: FSET: Set of data
Output: FPERM: Multiset of augmented permutations
    FPERM ← ∅
    for each start event SS in FS do
    ACCPERM ← ∅
    for each element σ in FSET do
        if σ(1) = SS or |ACCPERM| > 0 then
            ACCPERM ← ACCPERM ∪ σ(2) ∪ σ(3)
        if |ACCPERM| > 0 and σ(4) ∈ ES then
            FPERM ← FPERM ⊎ (SS, ACCPERM, σ(4))
            ACCPERM ← ∅
```

Listing 1. Augmentation procedure.

processes: the first group of events starts with event "b" and includes "c" and "d", the second group starts with event "d" and includes "c" and "b". It is also possible that some data may be lost during uploading or preprocessing. To take these situations into account, we introduce a control metric:

$$\mathrm{MinLimit}(\sigma) = |\sigma(2)|! - \varepsilon,\ \sigma \in FPERM,\ \varepsilon \geq 0. \quad (16)$$

Metric (16) allows us to calculate the required number of identical elements, taking into account the capability of adjusting ε for missing or lost data. Let us create set of unique groups of events that form logical AND operators by including only elements whose multiplicity is not less than a given limit:

$$LPERM = \left\{\sigma : \sum_{s \in FPERM} \left[s = \sigma\right] \geq \mathrm{MinLimit}(\sigma)\right\},$$
$$\sigma \in FPERM. \quad (17)$$

In addition to the augmented elements, the resulting set also contains parts of a larger permutation. To exclude such extra elements, we shall perform the check:

$$ANDGROUP = \left\{\sigma : \nexists \delta \mid \delta(2) = \right.$$
$$\left. = \{\sigma(1)\} \cup \sigma(2) \wedge \delta(3) = \sigma(3)\right\},\ \sigma, \delta \in LPERM. \quad (18)$$

Thus, the set $ANDGROUP$ will contain only the necessary groups of events that represent subprocesses of the business process run in parallel. For example, we obtain the following logical AND operators from (9):

$$ANDGROUP(L) =$$
$$= \left\{(e,\{b,c\},m),\ (a,\{b,c\},e),\ (a,\ \{b,c,d\},e)\right\}. \quad (19)$$

From (8) it follows that to determine the HMM, it is necessary to specify sets of hidden states and events, as well as to determine transition and emission matrices, and probability vector characterizing the choice of the start state. For an arbitrary business process, none of these model parameters are known in advance, since there is only an observable sequence of events that is obtained from the input data. Thus, it is necessary to formulate a method that allows us to find parameters if only the sequence of observed events of a business process is known. This problem was described by Rabiner and is one of the three basic problems when working with HMM and at the same time the most challenging among them [25]. The complexity is conditioned by the lack of known analytical methods for solving the task enabling us to determine the model parameters for any finite sequence O. There are several approaches to solving it by reducing the problem to an optimization task for finding such model parameters θ that allow maximizing the probability $P(O \mid \theta)$. One such approach is the Baum–Welch algorithm, which is a type of EM (expectation-maximization) algorithm for computing maximum likelihood estimates. In general, this algorithm consists of two steps (E-step and M-step), which make it possible to iteratively recalculate the parameters θ and successively approach the locally maximum estimate at a certain O.

However, the classic implementation of the Baum–Welch algorithm does not take into account the features of the subject area and the specifics of the business processes functioning. Therefore, this paper proposes an improved modification of the algorithm for application to the studied task.

Let us define the set of the observed events V of the model as equal to the set of unique business process events from multiset L:

$$V = \bigcup_{\sigma \in L} \left\{\sigma(i),\ 1 \leq i \leq |\sigma|\right\}. \quad (20)$$

Since each element in (18) is a logical AND operator and any unique group of permutable events must be assigned to one hidden state of the model, we define a set of hidden states:

$$SPU = \left\{\sigma(2) :\ \sigma \in ANDGROUP\right\}. \quad (21)$$

$$SOU = V \setminus \bigcup \delta,\ \forall \delta \in SPU. \quad (22)$$

$$S = \left\{i : 1 \leq i \leq |SOU| + |SPU|\right\}. \quad (23)$$

The iterative nature of the implementation of the classical Baum–Welch algorithm allows matrices A and B to be specified with arbitrary values before starting its operation, since in the process of updating the parameters convergence to optimal values will be achieved. However, it is known that the HMM various organizational structures (ergodic, left-to-right, parallel left-to-right, etc.) can influence the nature of its behavior and the obtained outcome. Let each event from set V be consecutively numbered with natural number v_k from 1 to $|V|$. Let us define matrices A and B as follows:

$$a_{ij} = \frac{1}{|S|},\ 1 \le i, j \le |S|. \tag{24}$$

$$b_{ik} = \begin{cases} 1 : \forall i \in \{1 \le i \le |SOU|\} \wedge \forall k = \\ = v_k \in SOU\ |\forall i \ne j\ b_j \ne b_i \\ \frac{1}{|\sigma|} : \forall i \in \{|SOU| < i \le |S|\} \wedge \forall k = \\ = v_k \in \sigma\ |\forall i \ne j\ b_j \ne b_i, \forall \sigma \in SPU \\ 0 : \text{else}. \end{cases} \tag{25}$$

To set the initial distribution, we will assume that the start event of the business process is unique and for this event it is defined that $v_k = 1$:

$$\pi_i = \begin{cases} 1 : i = v_k \\ 0 : i \ne v_k. \end{cases} \tag{26}$$

If there are elements in multiset L that contain different start events, then it is always possible to add a new surrogate event to the beginning of all elements of the multiset in order to move to the unicity of the start event.

To reduce the number of operations and simplify calculations, the forward and backward pass method is used, which is based on the principles of dynamic programming. In this case, a matrix of intermediate values is formed, which makes it possible to estimate the probability at each step by summing up the calculations performed in the previous steps using auxiliary functions:

$$\alpha_t(i) = P(o_1, o_2, o_3, \ldots, o_t, q_t = s_i | \theta). \tag{27}$$

$$\beta_t(i) = P(o_{t+1}, o_{t+2}, o_{t+3}, \ldots, o_T | q_t = s_i, \theta). \tag{28}$$

The classical implementation (27) and (28) does not take into account the model features associated with the specifics of the problem being studied. It is necessary to take into account the restrictions imposed on transitions between hidden states associated with logical AND operator. To do this, we define auxiliary functions as follows:

$$\varphi(i) = \{v_k | b_{ik} > 0,\ 1 \le k \le |V|\}. \tag{29}$$

$$HFW_t(i) = \begin{cases} |\varphi(i)| \times \prod_{j=0}^{|\varphi(i)|-1} [o_{t+j} \in \varphi(i)] : T > t + |\varphi(i)| \\ 0 : T \le t + |\varphi(i)|. \end{cases} \tag{30}$$

Function (30) specifies the estimate that will be used to select the most appropriate hidden state when calculating α and β. For such a state, the value of function (30) will be maximum:

$$HFW_t^{\max} = \underset{1 \le i \le N}{\operatorname{argmax}}\, HFW_t(i). \tag{31}$$

Then we will obtain:

$$\alpha_1(i) = \pi_i b_i(o_1). \tag{32}$$

$$\alpha_{t+1}(j) = \begin{cases} b_j(o_{t+1}) \sum_{i=1}^{N} \alpha_t(i) a_{ij} : \big(j = HFW_{t+1}^{\max} \wedge \\ \wedge\, o_{t+1} \in PS \wedge o_t \notin PS \big) \vee (o_{t+1} \notin PS) \\ b_j(o_{t+1})\, \alpha_t(j) a_{jj} : o_{t+1} \in PS. \end{cases} \tag{33}$$

This definition (33) enables us to limit the transitions space between hidden states of the model, which contain a group of events that form a logical AND operator. In a similar way we define:

$$HBW_t(i) = \begin{cases} |\varphi(i)| \times \prod_{j=0}^{|\varphi(i)|-1} [o_{t-j} \in \varphi(i)] : t - |\varphi(i)| \ge 0 \\ 0 : t - |\varphi(i)| < 0 \end{cases} \tag{34}$$

$$HBW_t^{\max} = \underset{1 \le i \le N}{\operatorname{argmax}}\, HBW_t(i). \tag{35}$$

$$\beta_T(i) = 1. \tag{36}$$

$$\beta_t(i)=\begin{cases}\sum_{j=1}^{N}\beta_{t+1}(j)a_{ij}b_j(o_{t+1}) : \Big(i=HBW_t^{\max}\wedge \\ \wedge\, o_t\in PS\wedge o_{t+1}\notin PS\Big)\vee(o_t\notin PS) \\ \beta_{t+1}(i)\,b_i\,(o_{t+1})\,a_{ii}\;:o_t\in PS \\ 0\;:o_t\in PS\wedge o_{t+1}\notin PS\wedge i\neq HBW_t^{\max}.\end{cases} \tag{37}$$

Since for a business process the end event means its final completion and the impossibility of transition to any other states, as well as for compliance with (7), we will assume that such hidden states should transform into themselves forming a loop. Let us define ξ and γ as follows:

$$\xi_t(i,j)=\begin{cases}\dfrac{\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}\;:1\le t\le T-1 \\ 1\;:t=T\wedge i=j\wedge\alpha_{T-1}(i)>0.\end{cases} \tag{38}$$

$$\gamma_t(i)=\frac{\alpha_t(i)\beta_t(i)}{\sum_{j=1}^{N}\alpha_t(j)\beta_t(j)}. \tag{39}$$

Considering that updating the coefficients shall be performed across all elements of multiset L, we will use:

$$a_{ij}^*=\frac{\sum_{\sigma\in L}\sum_{t=1}^{T}\xi_t(i,j)}{\sum_{\sigma\in L}\sum_{t=1}^{T}\gamma_t(i)}. \tag{40}$$

$$b_{ik}^*=\frac{\sum_{\sigma\in L}\sum_{t=1}^{T}\left[o_t=v_k\right]\gamma_t(i)}{\sum_{\sigma\in L}\sum_{t=1}^{T}\gamma_t(i)}. \tag{41}$$

Iterative execution of E and M steps of the algorithm is carried out up to its convergence or until a specified limit of repetitions is reached.

5. Options for using the method to solve business tasks

5.1. Predicting the outcome of a business process and finding deviations

Let us assume that there are many instances of a business process which are divided according to some parameter into several non-crossing groups G_1, G_2, G_3, ..., G_N. For example, within the framework of the “sale of goods” business process, we can split its instances according to the transaction outcome. In this case, the following groups can be generated: “refused to purchase”, “postponed the purchase” and “the purchase is successful”. Each such group corresponds to its own multiset L_1, L_2, L_3, ..., L_N. Using the proposed method described above, we will build N hidden Markov models using these multisets as a training sample. As a result, θ_n will correspond to each L_n. Let us use the forward-backward pass algorithm and define (42):

$$\alpha_t(j)=\begin{cases}\pi_j b_j(o_1)\;:t=1 \\ b_j(o_t)\sum_{i=1}^{N}\alpha_{t-1}(i)a_{ij}\;:1<t\le T.\end{cases} \tag{42}$$

For a new instance of the business process O_x, using the built HMMs, we can predict its membership in one of the previously generated groups. The group G_n will be the target group for which the estimate $P(O_x\,|\,\theta_n)$ is the highest:

$$F(O_x)=\underset{\theta}{\operatorname{argmax}}\,P(O_x|\theta)=\underset{\theta}{\operatorname{arg\,max}}\sum_{i=1}^{N}\alpha_T(i). \tag{43}$$

As a result, instance O_x is most likely to have an outcome that matches group G_n. Such a prediction can also be obtained for incomplete instances of a business process, that is, for those cases when there is only part of the O_x sequence. Having the capacity to obtain such an evaluation, one can solve various practical tasks. So, for example, for the business process of selling a product, you can analyze transactions that are at some intermediate stage to predict a possible outcome. If a high probability of an undesirable outcome is determined, then corrective actions can be developed for such transactions aimed at correcting the path. In addition, having a reference model of a business process as input and data from the respective event log, we can, having received an evaluation of the membership of each sequence, identify deviant instances for the purpose to further examine the causes and make management decisions.

5.2. Representation of a business process as a dependency graph

Suppose that for a certain process, an event log is recorded and processed, on the basis of which multiset L is formed:

$$L=\{<a,z>^5,\ <a,e,b,c,z>^5,\ <a,e,c,b,z>^3,$$
$$<a,b,c,d,z>^5, <a,b,d,c,z>^5, <a,c,b,d,z>^5,$$
$$<a,c,d,b,z>^5, <a,d,c,b,z>^5, <a,d,b,c,z>^5,$$
$$<a,l,z>^{10}, <a,l,l,z>^5, <a,l,l,l,z>^1\}. \qquad (44)$$

If we build the HMM for the multiset (44) using the proposed method described above, we obtain the following matrices A and B:

$$A=\begin{pmatrix} 0 & 0 & 0.51 & 0.14 & 0.27 & 0.08 \\ 0 & 0.5 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.67 & 0 & 0 & 0.33 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.3 & 0\;7 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$B=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0.33 & 0.33 & 0.33 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

We assume that hidden states for which $a_{ii}>0$ and containing several events in matrix B correspond to the logical AND operator, and those having only one event generate a cycle. The dependency graph drawm for the multiset (44) is shown in *Fig. 5.*

When needed, this dependency graph can be converted into other representations: BPMN, Petri nets, Casual Net, etc. The resulting dependency graph can be used to study the actual execution of a business process, conduct a comparative analysis of implementation options between different structural units, search for deviations and identify their causes. If we supplement the model with data from the event log about the execution time of basic operations, we can calculate various performance indicators (processing and idle time, duration and effective time of one cycle, etc.). In addition, event logs may contain information about participants, costs incurred and resources used which will allow the model to be scaled to analyze other aspects of the business process.

Conclusion

The data-driven approach is not an alternative to traditional modeling using analysts and domain experts. However, the use of this approach makes it possible to improve the quality of analysis, modeling, design and reengineering of business processes through the study of actual data that have been accumulated in the enterprise's information systems. Detection of non-obvious connections, as well as the capability of

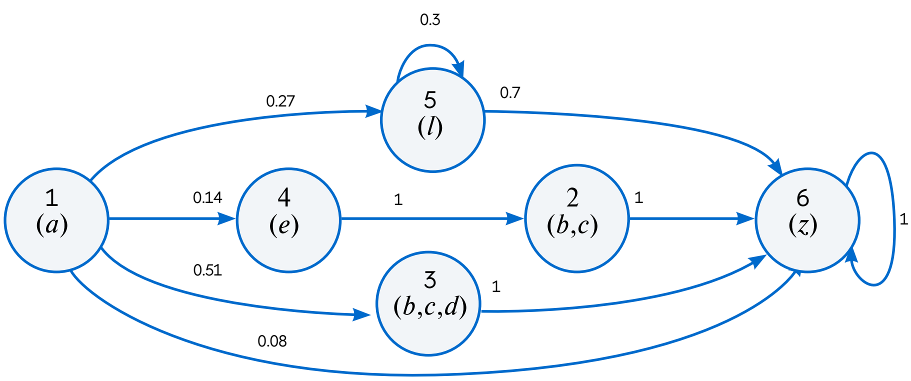

Fig. 5. Business process dependency graph.

impartial analysis, regardless of the value judgment of the process participants, help to minimize the likelihood of distortions and erroneous conclusions. The model so built can be used to monitor the execution of specific instances of a business process, identify deviations or abnormal behavior, and will also provide support for the implementation of key performance indicators (KPIs) in the company's activities, both at the level of individual employees and entire departments.

Unlike other algorithms described in this paper, the proposed method is based on a hidden Markov model, which allows the use of the apparatus of probability theory and mathematical statistics. In particular, a method for obtaining an evaluation of a business process future outcome is demonstrated, which enables implementation of proactive management influence in order to adjust the expected result. In addition, using HMM, you can perform clustering of business process instances and solve the classification task.

The identified shortcomings include: the lack of a guaranteed occurrence of all events that generate the logical AND operator (when using the model as a generator), as well as a narrow horizon for accounting the dependencies (due to the first-order assumption).

As a direction for the development of the method, it is advisable to consider the multi-level hierarchical organization of the model, the introduction of ensemble methods of machine learning and use of higher-order HMMs. ■

References

1. Lizano-Mora H., Palos-Sánchez P.R., Aguayo-Camacho M. (2021) The evolution of business process management: A bibliometric analysis. *IEEE Access*, vol. 9, pp. 51088–51105. https://doi.org/10.1109/ACCESS.2021.3066340

2. Fetais A., Abdella G.M., Al-Khalifa K.N., Hamouda A.M. (2022) Business process re-engineering: A literature review-based analysis of implementation measures. *Information*, vol.13, no. 4, 185. https://doi.org/10.3390/info13040185

3. Rosemann M. (2006) Potential pitfalls of process modeling: part A. *Business Process Management Journal*, vol. 12, no. 2, pp. 249–254. https://doi.org/10.1108/14637150610657567

4. Nambiar A., Mundra D. (2022) An overview of data warehouse and data lake in modern enterprise data management. *Big Data and Cognitive Computing*, vol. 6, no. 4, 132. https://doi.org/10.3390/bdcc6040132

5. Pegoraro M., van der Aalst W.M.P. (2019). Mining uncertain event data in process mining. *2019 International Conference on Process Mining (ICPM)*, pp. 89–96. https://doi.org/10.1109/ICPM.2019.00023

6. Andrews R., van Dun C.G.J., Wynn M.T., Kratsch W., Röglinger M.K.E., ter Hofstede A.H.M. (2020) Quality-informed semi-automated event log generation for process mining. *Decision Support Systems*, vol. 132, 113265. https://doi.org/10.1016/j.dss.2020.113265

7. Park G., van der Aalst W.M.P. (2021) Realizing a digital twin of an organization using action-oriented process mining. *3rd International Conference on Process Mining (ICPM)*, pp. 104–111. https://doi.org/10.1109/ICPM53251.2021.9576846

8. Kratsch W., Manderscheid J., Röglinger M. et al. (2021) Machine learning in business process monitoring: A comparison of deep learning and classical approaches used for outcome prediction. *Business & Information Systems Engineering*, vol. 63, pp. 261–276. https://doi.org/10.1007/s12599-020-00645-0

9. Teinemaa I., Dumas M., Rosa M.L., Maggi F.M. (2019) Outcome-oriented predictive process monitoring: Review and benchmark. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 13, no. 2, pp. 1–57. https://doi.org/10.1145/3301300

10. Leno V., Polyvyanyy A., Dumas M., La Rosa M., Maggi F.M. (2021) Robotic process mining: vision and challenges. *Business & Information Systems Engineering*, vol. 63, pp. 301–314. https://doi.org/10.1007/s12599-020-00641-4

11. Munoz-Gama J., Martin N., Fernandez-Llatas C. et al. (2022) Process mining for healthcare: Characteristics and challenges. *Journal of Biomedical Informatics*, vol. 127, 103994. https://doi.org/10.1016/j.jbi.2022.103994

12. Grisold T., Mendling J., Otto M., vom Brocke J. (2021) Adoption, use and management of process mining in practice. *Business Process Management Journal*, vol. 27, no. 2, pp. 369–387. https://doi.org/10.1108/BPMJ-03-2020-0112

13. Mehdiyev N., Fettke P. (2021) Explainable artificial intelligence for process mining: A general overview and application of a novel local explanation approach for predictive process monitoring. *Interpretable Artificial Intelligence: A Perspective of Granular Computing*, *Springer*, pp. 1–28. https://doi.org/10.1007/978-3-030-64949-4_1

14. van der Aalst W.M.P., Weijters T., Maruster L. (2004) Workflow mining: Discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 9, pp. 1128–1142. https://doi.org/10.1109/TKDE.2004.47

15. Mannhardt F., de Leoni M., Reijers H.A. (2017) Heuristic mining revamped: an interactive, data-aware, and conformance-aware miner. *15th International Conference on Business Process Management (BPM 2017)*, pp. 1–5.

16. van der Aalst W.M.P., Adriansyah A., van Dongen B. (2011) Causal Nets: A modeling language tailored towards Process Discovery. *International Conference on Concurrency Theory*, pp. 28–42. https://doi.org/10.1007/978-3-642-23217-6_3

17. van Dongen B.F., Busi N., Pinna G.M., van der Aalst W.M.P. (2007) An iterative algorithm for applying the theory of regions in process mining. Proceedings of the *Workshop on Formal Approaches to Business Processes and Web Services (FABPWS'07)*, pp. 36–55.

18. Bergenthum R., Desel J., Lorenz R., Mauser S. (2007) Process mining based on regions of languages. *Business Process Management: 5th International Conference (BPM 2007), Brisbane, Australia, September 24–28, 2007*, pp. 375–383. https://doi.org/10.1007/978-3-540-75183-0_27

19. Leemans S.J.J., Fahland D., van der Aalst W.M.P. (2013) Discovering block-structured process models from event logs: A constructive approach. *Application and Theory of Petri Nets and Concurrency: 34th International Conference* (*PETRI NETS 2013)*, *Milan, Italy, June 24–28, 2013*, pp. 311–329. https://doi.org/10.1007/978-3-642-38697-8_17

20. Leemans S.J.J., Fahland D., van der Aalst W.M.P. (2015) Scalable process discovery with guarantees. *Enterprise, Business-Process and Information Systems Modeling (BPMDS EMMSAD 2015). Lecture Notes in Business Information Processing,* vol. 214. Springer, Cham, pp. 85–101. https://doi.org/10.1007/978-3-319-19237-6_6

21. Pande S.D., Kanna R.K., Qureshi I. (2022) Natural language processing based on name entity with n-gram classifier machine learning process through ge-based hidden Markov model. *Machine Learning Applications in Engineering Education and Management*, vol. 2, no. 1, pp. 30–39.

22. Sagayam K.M., Hemanth D.J. (2019) A probabilistic model for state sequence analysis in hidden Markov model for hand gesture recognition. *Computational Intelligence*, vol. 35, no. 1, pp. 59–81. https://doi.org/10.1111/coin.12188

23. Srivastava R.K., Pandey D. (2022) Speech recognition using HMM and Soft Computing. *Materials Today: Proceedings*, vol. 51, pp. 1878–1883. https://doi.org/10.1016/j.matpr.2021.10.097

24. Du J., Wang C., Wang L. et al. (2023) Automatic block-wise genotype-phenotype association detection based on hidden Markov model. *BMC Bioinformatics*, vol. 24, article 138. https://doi.org/10.1186/s12859-023-05265-5

25. Rabiner L.R. (1990) A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286. https://doi.org/10.1109/5.18626

About the author

Artem Yu. Varnukhov

Assistant, Department of Business Informatics, Ural State University of Economics, 62, 8 Marta Str., Yekaterinburg 620144, Russia

E-mail: varnuhov_ayu@usue.ru

DOI: 10.17323/2587-814X.2024.3.56.69

Long-term investment optimization based on Markowitz diversification

Alexander V. Kulikov
E-mail: avkulikov15@gmail.com

Dmitriy S. Polozov
E-mail: polozov.ds@phystech.edu

Nikita V. Volkov
E-mail: nikita.v.volkov@phystech.edu

Moscow Institute of Physics and Technology, Dolgoprudny, Russia

Abstract

The article introduces a long-term investment algorithm that identifies optimal solutions in lower dimensional spaces constructed through principal component analysis or kernel principal component analysis. Portfolio weights optimization is carried out using the Markowitz method. Hyperparameters of the model include window size, smoothing parameter, rebalancing period and the fraction of explained variance in dimensionality reduction methods. The algorithm presented incorporates weights regularization taking into account portfolio rebalancing transaction costs. Hyperparameters' selection is based on the Martin coefficient, which allows us to consider the maximum drawdown for the suggested algorithms. The results demonstrate that the proposed algorithm, trained from 1990 to 2016, shows higher returns and Sharpe ratios compared to the S&P 500 benchmark from 2017 to 2022. This indicates that weights optimization can improve the algorithm's performance through rebalancing.

Keywords: PCA, Kernel PCA, window size, Markowitz algorithm, Grid Search, Bayesian optimization

Citation: Kulikov A.V., Polozov D.S., Volkov N.V. (2024) Long-term investment optimization based on Markowitz diversification. *Business Informatics*, vol. 18, no. 3, pp. 56–69. DOI: 10.17323/2587-814X.2024.3.56.69

Introduction

The rapid development of financial markets is one of the most significant global trends nowadays. In the first two decades of the 21st century, market capitalization increased more than three times [1]. In the United States alone, from 2017 to 2021, the number of client investment accounts at the three largest brokerage firms – Charles Schwab, Fidelity Investments and Robinhood – nearly doubled [2]. This drive by private investors to actively seek promising investment opportunities has spurred researchers' interest in finding the optimal investment portfolio to maximize returns while accounting for financial risk.

Considerable progress in portfolio optimization was achieved with the advent of Markowitz's portfolio theory [3]. The classical Markowitz optimization problem involves finding an asset allocation in a portfolio that minimizes risk at a fixed expected return level. More generally, the search for an optimal asset allocation occurs with a given balance coefficient between risk minimization and return maximization, corresponding to the investor's degree of risk aversion.

The algorithm for finding the optimal asset allocation in the Markowitz problem is based on estimating the average returns of stocks and their covariance matrix from historical data. However, this algorithm has several limitation; in particular, the assumption that historical returns and risks of stocks will maintain their distribution in the future is not always accurate [4]. Moreover, Markowitz diversification is vulnerable to outliers due to the instability of average stock returns [5]. Outliers can also occur in the covariance matrix of returns used for risk assessment.

One approach to addressing the issues mentioned, as discussed in the literature, involves dimensionality reduction methods such as Principal Component Analysis (PCA) and Kernel Principal Component Analysis (Kernel PCA). These methods enhance the Markowitz algorithm by eliminating outliers in the covariance matrix and identifying key components to focus on when assembling a portfolio. PCA helps us to identify the most significant directions in the data space (principal components) that explain the highest variance. By reducing dimensionality, outliers have less impact on the principal components, as PCA considers the overall variability of the data. These approaches help to reduce noise in the data and improve portfolio optimization quality, as demonstrated in works [6, 7]. An example of using Kernel PCA in the Markowitz problem is presented in [8].

This paper addresses a generalized Markowitz problem that includes the commission an investor pays for rebalancing the portfolio and proposes an algorithm to solve this optimization problem. The algorithm preprocesses asset returns to remove noise, then constructs a portfolio using exponentially smoothed stock returns, and optimizes a linear combination of risk and return while accounting for the commission.

To overcome the limitations of the classical solution for the Markowitz problem, this study examines four variations of the proposed algorithm: without dimensionality reduction, with PCA, and with polynomial and Gaussian kernels in Kernel PCA.

In the aforementioned studies, the historical time horizon for determining expected returns and the covariance matrix was empirically determined. Studies [9–17] discuss approaches to predicting expected returns and the covariance matrix using machine learning methods and statistical models like GARCH. In this paper, to optimize the performance of the Markowitz algorithm, we propose optimizing the model's hyperparameters to determine the optimal portfolio based on historical data: portfolio rebalancing frequency, window size, the number of selected companies for further analysis, and the dimensionality reduction parameter in Kernel PCA.

Given the substantial number of hyperparameters in the algorithm, traditional parameter tuning methods (GridSearch and RandomSearch) may not yield optimal results. Therefore, we also use the Bayesian optimization method based on Gaussian processes for comparison [18]. Bayesian optimization is an algorithm for finding the optimal parameter values of a function when only limited information about the function and its behavior is available. The essence of the method lies in an iterative sequence of selecting subsequent trial points based on a probabilistic model that approximates the unknown function. This method allows for finding the optimal solution using a relatively small number of iterations, as the algorithm actively adapts to information obtained from previous iterations. In the context of this study, the function to be optimized is the Martin index, introduced in [19]. Thus, for each of the four algorithm variations, parameter tuning is conducted using both random search and Bayesian optimization.

The proposed algorithms were trained on the return data of approximately 300 stocks from 1990 to 2016. To evaluate the effectiveness of each of the eight algorithm variations, the study compared various performance metrics of the resulting portfolios and benchmarked them against the S&P 500 index. This index is a classic benchmark for developing algorithms as it is considered the best indicator of large-cap company stocks [20] and reflects the overall state of the US economy [21, 22]. It has been shown in [23] that the weights in the classical Markowitz algorithm are also related to this benchmark.

The research results indicate that the most effective algorithm is the variation using the Gaussian kernel in Kernel PCA tuned through Bayesian optimization. The portfolio optimized using this method shows the highest return and the Sharpe ratio, due to the lowest drawdown during the COVID-19 pandemic outbreak, also exhibited the highest volatility compared to other variations. Additionally, it is shown that from 2017 to 2022, all proposed algorithms achieve higher returns than the S&P 500 index.

The paper is structured as follows: Section 1 presents the formulation of the optimization problem, Section 2 provides a description of the algorithm for solving this problem and its variations, Section 3 details the metrics for comparing algorithm variations, Section 4 describes the data used, Section 5 presents an analysis of the research results.

1. Optimization problem

Consider a financial market model with N assets (stocks), whose returns are random variables r_i, $i = 1, \ldots, N$. The vector of expected stock returns is denoted as $\vec{\mu} = \left(\mathsf{E}r_1, \ldots, \mathsf{E}r_N\right)^\top$, and the covariance matrix of stock returns is denoted as

$$\Sigma = \left(cov(r_i, r_j)\right)_{(i,j=1)}^N.$$

An investor forms a portfolio from these assets with stock weights $w_i \geq 0$, $i = 1, \ldots, N$ where the condition $w_i \geq 0$ corresponds to the prohibition of short sales. Let α be the investor's risk aversion coefficient. Then the transformed Markowitz optimization problem for this investor takes the form:

$$\begin{aligned} \min_{\vec{w}} \quad & \frac{1}{2}\vec{w}^\top \Sigma \vec{w} - \alpha \vec{\mu}^\top \vec{w} \\ s.t. \quad & \vec{1}^\top \vec{w} = 1, \\ & \vec{w} \geq \vec{0}. \end{aligned} \tag{1}$$

According to [24], this problem is equivalent to the classical Markowitz problem, except for the fixed parameters: in the classical problem, the expected return parameter is fixed, while in the transformed problem, the balance coefficient between risk minimization and return maximization is fixed. In the algorithm considered $\alpha = 0.05$ however, in general, this parameter can be adjusted, with smaller α emphasizing risk minimization and larger α emphasizing return maximization.

In this paper, we consider a more general Markowitz optimization problem, which includes a rebalancing commission λ. Thus, if P is the rebalancing period and $\vec{w}_p$ is the stock weight distribution from the previous period, the optimization problem (1) takes the form:

$$\begin{aligned} \min_{\vec{w},P} \quad & \frac{1}{2}\vec{w}^{\top}\Sigma\vec{w} - \vec{\alpha\mu}^{\top}\vec{w} - \frac{\lambda}{P}\left\|\vec{w}-\vec{w}_p\right\|^2 \\ s.t. \quad & \vec{1}^{\top}\vec{w} = 1, \\ & \vec{w} \geq \vec{0}. \end{aligned} \tag{2}$$

We define the weight distribution vector at time 0 as $\vec{w}_p = \vec{0}$.

2. Algorithm

2.1. Algorithm stages

The optimization algorithm receives input data on stock prices over a specified period. The output of the algorithm is the weights w_i – the proportions of the corresponding assets in the investor's portfolio, which are rebalanced at a specified frequency, provided to the algorithm as a hyperparameter.

The operation of the algorithm, as presented below, can be broadly divided into two parts: hyperparameter tuning of the portfolio and selection of optimal weights corresponding to these hyperparameters. The hyperparameter tuning involves selecting parameters for data preprocessing (portfolio rebalancing frequency, window size factor, the number of selected companies for further analysis) and dimensionality reduction parameters (explained variance ratio and kernel hyperparameters in Kernel PCA).

Weight selection is performed using the fit method on a training dataset of stocks over a certain period and includes the following stages.

Data filtering. In the first stage, we filter data to keep for the analysis only observations within the specified time window.

The window size is determined by two hyperparameters: *period_change_portfolio* – the portfolio rebalancing period, and *size_of_window_rank* – the window size factor. The window size is expressed as the product of these parameters. This functional relationship allows the window size to vary in conjunction with changes in the portfolio rebalancing period.

Typically, the window size factor ranges from two to five, corresponding to a window of approximately 2–5 rebalancing periods. For instance, in [25], a 5-year window is used to analyze an asset's annual beta, whereas in [26], a moving window of 1000 days is used to evaluate portfolio rebalancing parameters every 250 days. Applying window filtering allows for the consideration of the most recent and relevant data while excluding outdated information.

Stock selection for portfolio construction. At this stage, the profitability of stocks is evaluated for inclusion in a portfolio consisting of the most profitable ones.

To assign greater weight to the most recent data, we apply exponential smoothing with a parameter $\alpha \in (0, 1)$. In the algorithm, the exponential smoothing parameter is set to $\alpha = 0.99$. If the application of filtering in the previous stage is equivalent to multiplying all observations within the window by 1, and outside it by 0, then exponential smoothing can be represented as multiplying returns by $\alpha^{\Delta t}$, where Δt represents the number of trading days that have passed from the date under consideration to the present moment. This approach ensures a higher weight for recent data while preserving the influence of earlier periods, with diminishing weight as one moves further back in time.

To select stocks to be included in the portfolio, we calculate the weighted average return for each stock using exponentially smoothed weights. The calculation is performed according to the formula:

$$\overline{r}_{i\tau} = \frac{\sum_{t=\tau-T+1}^{\tau} r_{it}\alpha^{\tau-t}}{\sum_{t=\tau-T+1}^{\tau} \alpha^{\tau-t}},$$

where T is the window size, r_{it} – return of the stock i at time t.

After computing the weighted average returns, we sort the stocks in descending order, and select the top *n_top_companies* for inclusion in the portfolio. The number of selected stocks is a hyperparameter of the algorithm.

Dimensionality reduction. This step is applied only for variations of the algorithm that use PCA or Kernel PCA. As noted above, these methods enhance the Markowitz algorithm by removing outliers in the covariance matrix and isolating components. These components are subsequently used by the algorithm in portfolio formation.

When applying these dimensionality reduction methods, the number of components to include in the model is initially determined. This number is based on the explained variance of the first 1, 2, ..., *n_top_companies* components according to the hyperparameter *var_ratio*, which dictates the proportion of explained variance. The proportion of explained variance for d components is denoted as δ_d, which is the ratio of the sum of squared deviations of the observed data from their projection onto the principal components to the sample variance of the data. A high residual variance indicates that the principal components do not explain a sizable portion of the data variability, possibly due to additional unaccounted factors in the model. The number of components d is determined as the number where $\delta_d \geqslant$ *var_ratio*, but $\delta_{d-1} <$ *var_ratio*.

The identified number of components and hyperparameters are then passed to the dimensionality reduction method, which is trained on the training data. This process reduces the data dimensionality while preserving key characteristics and removing outliers.

Kernel PCA Hyperparameters:

- *kernel* – specifies the kernel type; the algorithm considered only the Gaussian kernel "rbf":

$$k\left(\vec{x},\vec{y}\right)=\frac{e^{-\left\|\vec{x}-\vec{y}\right\|^{2}}}{\gamma},$$

and polynomial kernel "poly":

$$k\left(\vec{x},\vec{y}\right)=\frac{\left(\vec{x}^{\top}\vec{y}+c\right)^{q}}{\gamma};$$

- *kerneldegree* – the degree of the polynomial kernel (q in the polynomial kernel formula);
- *kernelgamma* – the scale parameter (γ in the above formulas);
- *kernelcoef0* – the constant of the polynomial kernel (c in the polynomial kernel formula).

Stock allocation in the portfolio. After forming a set of stocks for inclusion in the portfolio, the algorithm determines their weights.

The algorithm solves the problem (2), where the last term essentially acts as regularization of the weights. We set the parameter λ, which accounts for the rebalancing commission of the portfolio, to be equal to 1%.

Additionally, if dimensionality reduction methods are used, it is necessary to transition from the latent space to the original space by inverse transformation $\vec{w}=\phi^{-1}\left(\vec{w}_q\right)$ for solving the problem. Let $\bar{\vec{r}}$ be the vector of average returns in the latent space, and C be the sample covariance matrix in the latent space. Then the transformed optimization problem takes the form:

$$\begin{aligned}\min_{\vec{w}_q}\quad & \frac{1}{2}\phi^{-1}\left(\vec{w}_q\right)^{\top}C\phi^{-1}\left(\vec{w}_q\right)-\alpha\bar{\vec{r}}^{\top}\phi^{-1}\left(\vec{w}_q\right)+\frac{\lambda}{P}\left\|\phi^{-1}\left(\vec{w}_q\right)-\vec{w}_p\right\|^{2}\\ s.t.\quad & \vec{1}^{\top}\phi^{-1}\left(\vec{w}_q\right)=1,\\ & \phi^{-1}\left(\vec{w}_q\right)\geq\vec{0}.\end{aligned}\tag{3}$$

In the case of PCA, the inverse transformation is performed by multiplying the weights by the component matrix $\phi^{-1}\left(\vec{w}_q\right)=Q\vec{w}_q$. This is a quadratic programming problem, which is effectively solved using the *cvxpy* package [27].

In the Kernel PCA method, to find the vector $\phi^{-1}\left(\vec{w}_q\right)$ in the original space we search for the approximate preimage of the vector $\vec{w}_q$ by solving a minimization problem with Ridge regression (see [28]):

$$\phi^{-1}\left(\vec{w}_q\right)=\arg\min_{\vec{z}\in\mathbb{R}^{D}}\left\|\phi\left(\vec{z}\right)-\vec{w}_q\right\|^{2}.$$

This problem is not a quadratic programming problem and is solved less accurately using the *scipy.optimize* library [29].

The inverse transformation of the weights $\vec{w}=\phi^{-1}\left(\vec{w}_q\right)$ follows the formulas above, along with truncating weights based on a threshold value: if $w_i <$ *treshold*, then $w_i = 0$. We set the threshold value around 10^{-6} to eliminate weights that are too small to be practically achievable, as whole stock lots must be purchased.

Portfolio return estimation. After training the algorithm on the training dataset, we can construct a portfolio using the obtained weights and evaluate its performance on the test dataset. However, a portfolio with constant weights performs worse over time because the distribution of stock returns changes. Therefore, it is periodically necessary to rebalance the portfolio. Every *period_change_portfolio* trading day, we retrain the algorithm on both the training data and the known test data. The portfolio change period in the algorithm was chosen to be relatively long, ranging from three to four months to several years. This is because the Markowitz method performs well over long periods with infrequent portfolio changes. For instance, in [30], minimum-risk portfolios, i.e., classic Markowitz portfolios, were rebalanced 2–3 times a year.

We evaluate the portfolio return using the predict method, which takes the trained algorithm and test dataset as inputs. The output is a dataframe indexed by time and containing the calculated returns as a column. This allows for visualizing the results as a time series of portfolio returns and calculating economic indicators using the score method.

2.2. Hyperparameter tuning

In this study, we consider four variations of the initial algorithm: no dimensionality reduction, dimensionality reduction using PCA, and Kernel PCA with polynomial and Gaussian kernels. For each method, we optimize hyperparameters using two approaches: *RandomizedSearchCV* [31] and Bayesian optimization. This resulted in eight algorithms with different sets of hyperparameters.

The following hyperparameters were tuned in the methods for finding optimal weights:

- for the method without dimensionality reduction, only the hyperparameters of the pre-processing of the returns' matrix were selected. These parameters include:
  - the portfolio change period (ranging from three months to two and a half years);
  - the window size factor (from 2 to 5);
  - the number of top companies selected based on average returns (from 50 to 200);
- for the PCA method, in addition to the above parameters, the explained variance ratio was also tuned (from 80% to 95%);
- for Kernel PCA with a polynomial kernel, the above hyperparameters were tuned along with the polynomial kernel parameters:
  - degree (from 2 to 4);
  - constant term (from 0.5 to 1);
  - order parameter (from 0.02 to 0.04);
- for Kernel PCA with a Gaussian kernel, apart from the pre-processing parameters and the explained variance ratio, the order parameter was tuned (from 0.001 to 0.1). It is also noteworthy that, according to article [8], a portfolio formed using Kernel PCA with a Gaussian kernel is more risky. Therefore, we decided to diversify the risk less and optimized the number of *n_top_companies* based on returns in the range from 5 to 30 (for RBF and Bayesian RBF).

3. Portfolio comparison metrics

To compare portfolio optimization methods, we consider several metrics.

Portfolio value. The ratio of the current portfolio price to the initial price:

$$V_{\mathcal{P}}(\tau)=\prod_{t=1}^{\tau}\left(1+\vec{r}_t^{\top}\vec{w}_t\right).$$

Average rate of return (AR). The average investment profit earned per year. This measure is used for comparing the returns of different investment instruments. Let L be the total investment period in years, and T be the total number of trading days. The average annual return is then

$$AR_{\mathcal{P}}=\left(V_{\mathcal{P}}\left(T\right)\right)^{\frac{1}{L}}-1.$$

VaR (value at risk) at level α. The inverse sample α-quantile of the portfolio returns

$$VaR_{\mathcal{P}}(\alpha) = -\widehat{q}_{\alpha}\left(\vec{r_t}^{\top}\vec{w_t}\right).$$

Risk premium (excess return). The gain from holding the portfolio compared to a risk-free asset. For a portfolio $\mathcal{P}$ with asset weights $\vec{w_t}$, relative returns $\vec{r_t}$, and risk-free rate rf_t at time $t = 1, ..., T$:

$$\mu_{\mathcal{P}} = \frac{1}{T}\sum_{t=1}^{T}\left(\vec{r_t}^{\top}\vec{w_t} - rf_t\right).$$

Portfolio standard deviation. The sample standard deviation of the excess return, calculated as

$$\sigma_{\mathcal{P}} = \frac{1}{T-1}\sqrt{\sum_{t=1}^{T}\left(\vec{r_t}^{\top}\vec{w_t} - rf_t - \mu_{\mathcal{P}}\right)^2}.$$

Sharpe ratio. The ratio of the risk premium to the standard deviation of the portfolio:

$$ShR_{\mathcal{P}} = \frac{\mu_P}{\sigma_{\mathcal{P}}}.$$

Relative drawdown. The relative difference between maximum portfolio value up to the current moment and its current value. If $V_{\mathcal{P}}(t)$ is the portfolio price ratio at time t to the initial price, then

$$RD_{\mathcal{P}}(t) = 1 - \frac{V_{\mathcal{P}}(t)}{\max\limits_{\tau=1,\dots,t} V_{\mathcal{P}}(\tau)}.$$

Martin ratio (Ulcer performance index, UPI) [10]. The ratio of the portfolio's excess return to the root mean square relative drawdown:

$$PF_{\mathcal{P}} = \frac{\mu_P}{\sqrt{\overline{RD_P^2}}}.$$

The Martin ratio is used for hyperparameter tuning during the validation stage.

4. Data

The study utilized data sourced from Yahoo's database using the *yfinance* library [32]. The analyzed assets included stocks of companies listed in the S&P 500 index, as well as 500 randomly selected stocks from Yahoo's database. This asset selection ensures that the weights obtained from the algorithm are independent of the weights built solely on the benchmark index. The constructed dataset includes the daily price dynamics of 1000 stocks over the period from 1990 to 2022.

We conducted preliminary data filtering to exclude stocks with more than 10% missing values across observations. For the remaining observations with missing values, the return was estimated as the average return for the entire preceding historical period. Additionally, we excluded from the analysis stocks with high volatility, whose price changed by more than twice compared to the previous day at least once.

After filtering, 300 stocks representing various sectors were selected. These included companies from the IT sector such as Apple and Microsoft; the financial sector including JPMorgan Chase and Citigroup; and the healthcare industry including Johnson & Johnson and Pfizer. Moreover, the list included consumer goods companies such as PepsiCo, The Coca-Cola Company, McDonald's, Procter & Gamble, and Walmart.

The data obtained was divided into two periods: train period (1990–2016) and test period (2017–2022). Hyperparameter tuning was performed using cross-validation for time series. This cross-validation method involves dividing the data into sequential blocks based on time. The model is trained on all preceding data blocks and validated on the subsequent block (see *Fig. 1*). We repeat this process multiple times and using the *TimeSeriesSplit* method from *sklearn* [33].

5. Results

The configurations of the selected hyperparameters for all methods are presented in *Table 1.* It is noteworthy that for most methods, the optimal portfolio rebalancing period is approximately six months (180 days), potentially indicating a global optimality for this time interval. Interestingly, the kernel scale factor for kernel models is roughly the same, around 0.04.

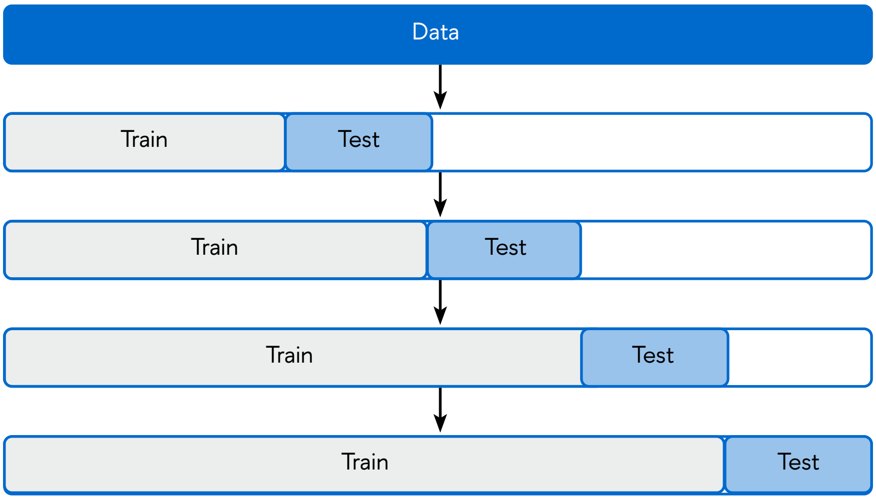

Fig. 1. Cross-validation scheme by *TimeSeriesSplit* from *sklearn*.

Table 1.

Selected hyperparameters for all 8 algorithms

	size_of_window_rank	period_change_portfolio	n_top_companies	var_ratio	kernel	kernelgamma	kerneldegree	kernelcoef0
No optimization	3.20	190	97	–	–	–	–	–
PCA	4.90	192	150	0.935	–	–	–	–
Poly KPCA	2.10	249	86	0.86	poly	0.04	4	0.77
rbf	4.90	170	27	0.91	rbf	0.04	–	–
Bayes	4.79	190	100	–	–	–	–	–
Bayes PCA	3.45	189	71	0.891	–	–	–	–
Bayes Poly KPCA	3.37	190	66	0.921	poly	0.034	2	0.844
Bayes rbf	2.44	188	17	0.764	rbf	0.044	–	–

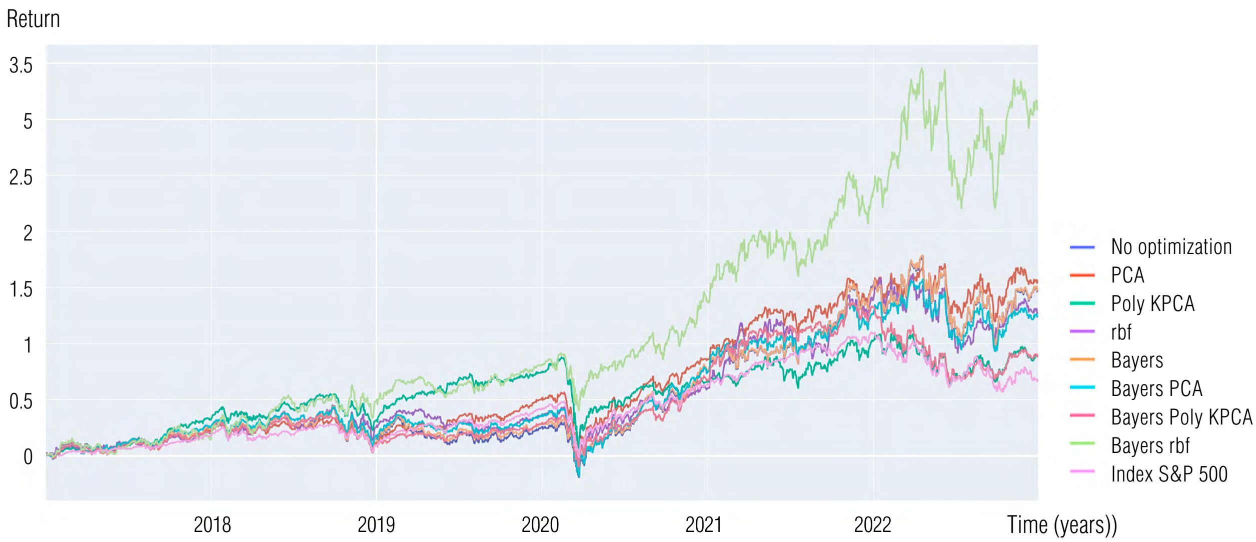

Fig. 2. Returns on the test dataset for all 8 algorithms and the S&P 500.

After the hyperparameter tuning processes using Bayesian optimization and random search for the four methods, we compute the return on the test sample. The return graphs on the test dataset for all eight algorithms and the baseline S&P 500 benchmark are presented in *Fig. 2* and are also available on an interactive HTML chart [34].

In addition to the graphs, we calculate economic indicators from section 3. These are presented in *Table 2* for each portfolio and the S&P 500.

The study results show that the Bayes RBF method, using a Gaussian kernel in Kernel PCA, tuned with Bayesian optimization, is the most efficient algorithm. The portfolio optimized with this method had the lowest drawdown of 28.5% during the period of February–April 2020, corresponding to the first outbreak of the COVID-19 pandemic, which is 5 percentage points lower than the drawdown level of the S&P 500. This observation confirms the findings of study [8], demonstrating that algorithms with a Gaussian kernel perform better in market crisis situations. At the same time, the Gaussian kernel RBF and Bayes RBF algorithms show the highest volatility (27.7% and 28.5%, respectively). This result is due to the limited number of stocks considered, and thus, lower portfolio risk diversification.

Portfolios with the lowest volatility, apart from the S&P 500 index, were those with a polynomial kernel: Poly KPCA and Bayes Poly KPCA, with 22.5% and 23.1% volatility, respectively. This indicates the polynomial kernel's capability to effectively filter noise. Additionally, these portfolios are characterized by the best 5% VaR values (around 2.1%), which are slightly higher than the value for the S&P 500. However, the average annual returns for these portfolios are among the lowest – 12.9% and 13.0%.

PCA and Bayes PCA portfolios, based on the principal component method, showed average results in terms of risk and return. This indicates that even such classical algorithms with additional optimization outperform benchmarks on most metrics.

At the same time, the study results do not allow us to establish that Bayesian optimization or random search is more effective in hyperparameter selection.

Table 2.

Economic indicators on the test set for all 8 algorithms and the S&P 500

	S&P 500	Without optimization	PCA	Poly KPCA	rbf	Bayes	Bayes PCA	Bayes Poly KPCA	Bayes rbf
Profit factor	1.09	1.1	1.13	1.09	1.09	1.1	1.1	1.09	1.15
Sharpe ratio	0.41	0.57	0.65	0.46	0.63	0.58	0.53	0.45	0.85
Martin ratio	0.26	1.05	1.9	0.71	1.47	1.11	1.09	0.54	3.64
Annualized return	10.87%	18.75%	18.11%	12.9%	17.17%	18.71%	16.54%	13.01%	26.98%
Annualized volatility	20.26%	28.14%	23.97%	22.5%	27.71%	27.83%	26.08%	23.13%	28.51%
5% VaR	1.935%	2.815%	2.4%	2.113%	2.715%	2.817%	2.442%	2.151%	2.897%
Max drawdown	33.92%	44.41%	38.04%	43.51%	37.03%	42.99%	43.29%	34.1%	28.27%
Winning days	54.5%	54.2%	55.3%	53.6%	53.6%	54.1%	54.1%	53.4%	54.0%

While the S&P 500 is characterized by the lowest risk, its return is also low, which is reflected in the ratio indicators. Thus, the portfolios constructed, although more volatile than the S&P 500, show better results in other metrics, demonstrating that optimal portfolio rebalancing can provide additional return at the expense of risk.

To align our portfolios with the S&P 500's risk level, i.e. the average annual volatility of the S&P index, we consider the results of the Bayes RBF algorithm if part of the money is invested in this algorithm and the remainder in a risk-free asset (average annual return results are presented in *Table 3*). Note that these results allow the use of the considered algorithm over a long period, achieving higher returns than the S&P while maintaining the same risk level.

Conclusion

The study explored a long-term investment algorithm based on solving the Markowitz problem with an initial transition to a lower-dimensional space using various principal component analysis (PCA) variations. Periodic portfolio rebalancing allows for flexible responses to market changes, which is evidenced by drawdowns comparable to the benchmark in some models, despite higher risk indicators. However, portfolio rebalancing was not performed too frequently to allow the Markowitz method to demonstrate its effectiveness over sufficiently long intervals.

Moreover, optimizing the Martin coefficient during hyperparameter selection showed that the portfolios

Table 3.

Comparison of the best algorithm and S&P 500

	S&P 500	Bayes rbf	S&P-risk Bayes rbf	0% Bayes rbf	50% Bayes rbf	75% Bayes rbf	90% Bayes rbf
Annualized return	10.87%	26.98%	19.93%	2.61%	14.8%	20.89%	24.54%
Annualized volatility	20.26%	28.51%	20.26%	0.0%	14.26%	21.38%	25.66%
Sharpe ratio	0.41	0.85	0.85	0.0	0.85	0.85	0.85

achieved high returns and relatively moderate drawdowns.

Portfolios with trajectories more profitable than the benchmark, albeit slightly riskier, were also obtained. It is important to note that portfolio risk can be managed through the risk aversion coefficient, allowing investors to choose portfolios based on their risk preferences. Risk can also be regulated by the selected methods: a method with Kernel PCA and a polynomial kernel selects a less risky but also less profitable portfolio, while a method without dimensionality reduction takes on more risk with the potential for higher returns. A balanced approach in this dilemma could be the simple PCA method, which maintains a balance between risk and return.

Notably, the algorithm using Kernel PCA with a Gaussian kernel exhibited the best economic indicators, mainly due to its modest drawdown during the pandemic compared to other algorithms. It can be concluded that during crisis periods, the application of a Gaussian kernel to stock returns for transitioning to a lower-dimensional space is most effective. ■

References

1. The World Bank Group (2023) *Market capitalization of listed domestic companies. World Federation of Exchanges database*. Available at: https://data.worldbank.org/indicator/CM.MKT.LCAP.CD (accessed 20 November 2023).

2. Abramov A.E., Kosyrev A.G., Radygin A.D., Chernova M.I. (2021) Behavior of private investors in the stock markets of Russia and the US. *Russian Economic Developments*, vol. 18, pp. 11–16 (in Russian).

3. Markowitz H. (1952) Portfolio selection. *The Journal of Finance*, vol. 7, no. 1, pp. 77–91. https://doi.org/10.2307/2975974

4. Durall R. (2022) Asset allocation: From Markowitz to deep reinforcement learning. *arXiv:2208.07158*. https://doi.org/10.48550/arXiv.2208.07158

5. Best M.J., Grauer R.R. (1991) On the sensitivity of mean–variance-efficient portfolios to changes in asset means: Some analytical and computational results. *The Review of Financial Studies*, vol. 4, no. 2, pp. 315–342. https://doi.org/10.1093/rfs/4.2.315

6. Stefatos G., Hamza A.B. (2007) Cluster PCA for outliers detection in high-dimensional data. Proceedings of the *2007 IEEE International Conference on Systems, Man and Cybernetics, Montreal, QC, Canada*, pp. 3961–3966. https://doi.org/10.1109/ICSMC.2007.4414244.

7. Saha B.N., Ray N., Zhang H. (2009) Snake validation: A PCA-based outlier detection method. *IEEE Signal Processing Letters*, vol. 16, no. 6, pp. 549–552. https://doi.org/10.1109/LSP.2009.2017477

8. Peng Y., Albuquerque P.H.M., do Nascimento I.F., Machado J.V.F. (2019) Between nonlinearities, complexity, and noises: An application on portfolio selection using kernel principal component analysis. *Entropy*, vol. 21, no. 4, 376. https://doi.org/10.3390/e21040376

9. Ma Y., Han R., Wang W. (2021) Portfolio optimization with return prediction using deep learning and machine learning. *Expert Systems with Applications*, vol. 165, 113973. https://doi.org/10.1016/j.eswa.2020.113973

10. Heaton J.B., Polson N., Witte J. (2016) Deep learning for finance: Deep portfolios. *Applied Stochastic Models in Business and Industry*, vol. 33, no. 1, pp. 3–12. https://doi.org/10.2139/ssrn.2838013

11. Chen K., Zhou Y., Dai F. (2015) A LSTM-based method for stock returns prediction: A case study of China stock market. Proceedings of the *2015 IEEE International Conference on Big Data (Big Data), Santa Clara, CA, USA*, pp. 2823–2824. https://doi.org/10.1109/BigData.2015.7364089

12. Yun H., Lee M., Kang Y.S., Seok J. (2020) Portfolio management via two-stage deep learning with a joint cost. *Expert Systems with Applications*, vol. 143, 113041. https://doi.org/10.1016/j.eswa.2019.113041

13. Chong E., Han C., Park F. (2017) Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, vol. 83, pp. 187–205. https://doi.org/10.1016/j.eswa.2017.04.030

14. Fischer T., Krauss C. (2018) Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, vol. 270, no. 2, pp. 654–669. https://doi.org/10.1016/j.ejor.2017.11.054

15. Hoseinzade E., Haratizadeh S. (2019) CNNpred: CNN-based stock market prediction using a diverse set of variables. *Expert Systems with Applications*, vol. 129, pp. 273–285. https://doi.org/10.1016/j.eswa.2019.03.029

16. Kim J., Lee M. (2023) Portfolio optimization using predictive auxiliary classifier generative adversarial networks. *Engineering Applications of Artificial Intelligence*, vol. 125, 106739. https://doi.org/10.1016/j.engappai.2023.106739

17. Siaw R., Ofosu-Hene E., Tee E. (2017) Investment portfolio optimization with GARCH models. *Elk Asia Pacific Journal of Finance and Risk Management*, vol. 8, no. 2. Available at: https://ssrn.com/abstract=2987932 (accessed 04 August 2024).

18. Bardenet R., Bengio Y., Bergstra J., Kégl B. (2011) Algorithms for hyper-parameter optimization. Proceedings of the *Advances in Neural Information Processing Systems 24 (NIPS 2011)*.

19. Martin P.G., McCann B.B. (1989) *The investor's guide to fidelity funds*. John Wiley & Sons.

20. S&P Global (2023) *S&P 500. S&P Dow Jones Indices*. Available at: https://www.spglobal.com/spdji/en/indices/equity/sp-500/#overview (accessed 20 November 2023).

21. Beneish M.D., Whaley R.E. (1997) A scorecard from the S&P game. *Journal of Portfolio Management*, vol. 16, no. 2, 23.

22. Latham S., Braun M. (2010) Does short-termism influence firm innovation? An examination of S&P 500 firms 1990–2003. *Journal of Managerial Issues*, vol. 22, no. 3, pp. 368–382.

23. Zhang Z. (2022) Study of portfolio performance under certain restraint comparison: Markowitz Model and Single Index Model on S&P 500. Proceedings of the *2022 7th International Conference on Social Sciences and Economic Development*, pp. 1930–1938. https://doi.org/10.2991/aebmr.k.220405.323

24. Lien G. (2002) Non-parametric estimation of decision makers' risk aversion. *Agricultural Economics*, vol. 27, no. 1, pp. 75–83. https://doi.org/10.1016/S0169-5150(01)00063-9

25. Fama E.F., MacBeth J. (1973) Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, vol. 71, pp. 607–636.

26. Paolella M.S. (2017) The univariate collapsing method for portfolio optimization. *Econometrics*, vol. 5, no. 2, 18. https://doi.org/10.3390/econometrics5020018

27. The CVXPY authors (2023) *CVXPY 1.4 Manual*. Available at: https://www.cvxpy.org/index.html (accessed 20 November 2023).

28. Bakir G., Weston J., Schölkopf B. (2003) Learning to find pre-images. Proceedings of the *Advances in Neural Information Processing Systems 16 (NIPS 2003)*.

29. The SciPy community (2023) *SciPy v1.11.4 Manual*. Available at: https://docs.scipy.org/doc/scipy/tutorial/optimize.html (accessed 20 November 2023).

30. Drenovak M., Rankovic V. (2014) Markowitz portfolio rebalancing with turnover monitoring. *Economic Horizons*, vol. 16, no. 3, pp. 207–217. https://doi.org/10.5937/ekonhor1403211D

31. scikit-learn. Machine Learning in Python (2023) *API Reference. sklearn.model_selection. RandomizedSearchCV*. Available at: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.RandomizedSearchCV.html (accessed 20 November 2023).

32. GitHub (2023) *yfinance documentation*. Available at: https://yfinance.readthedocs.io/en/documentation/ (accessed 20 November 2023).

33. scikit-learn. Machine Learning in Python (2023) *API Reference. sklearn.model_selection. TimeSeriesSplit*. Available at: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.TimeSeriesSplit.html (accessed 20 November 2023).

34. *Interactive chart of research results*. Available at: https://disk.yandex.ru/d/5nF7RiXKIWp1ig (accessed 20 November 2023).

About the authors

Alexander V. Kulikov

Dr. Sci. (Phys.-Math.);

Associate Professor, Moscow Institute of Physics and Technology, 9, Institutskiy Ln., Dolgoprudny 141701, Moscow Region, Russia;

E-mail: avkulikov15@gmail.com

ORCID: 0000-0002-3963-0814

Dmitriy S. Polozov

Master's Student, Moscow Institute of Physics and Technology, 9, Institutskiy Ln., Dolgoprudny 141701, Moscow Region, Russia;

E-mail: polozov.ds@phystech.edu

ORCID: 0009-0008-1108-4955

Nikita V. Volkov

Doctoral Student, Moscow Institute of Physics and Technology, 9, Institutskiy Ln., Dolgoprudny 141701, Moscow Region, Russia;

E-mail: nikita.v.volkov@phystech.edu

ORCID: 0009-0007-8434-9822

DOI: 10.17323/2587-814X.2024.3.70.86

Designing a multi-agent system for a network enterprise

Yury F. Telnov
E-mail: Telnov.YUF@rea.ru

Vasiliy A. Kazakov
E-mail: Kazakov.VA@rea.ru

Andrey V. Danilov
E-mail: Danilov.AV@rea.ru

Plekhanov Russian University of Economics, Moscow, Russia

Abstract

The necessity to enhance the efficiency of modern network enterprises based on digital platform technologies, Digital Twins, and Digital Threads determines the relevance of implementing dynamic multi-agent technologies in production practice. The architectural complexity of existing multi-agent systems (MAS) and the lack of scientific research in the field of justifying methods and tools for their creation motivate the goal of this study to develop a comprehensive MAS design technology. This technology should encompass all architectural levels and allow for the adaptation of reference and best design practices. This article analyzes the possibilities of applying Digital Twins and Digital Threads in the creation of network enterprises and proposes methods for their implementation using MAS. A design technology for MAS has been developed in accordance with the IIRA (Industrial Internet Reference Architecture) and RAMI (Reference Architectural Model Industrie 4.0) architectural frameworks, which enables the interconnected formation and display of design results across various architectural levels. At the business level, a method is proposed for formulating business requirements for MAS based on the selection and adaptation of business models and application scenarios. At the level of constructing manufacturing and business processes, a method for formulating functional requirements for MAS is presented, revealing the transition from value networks to manufacturing and business process structures. At the level of functional design of the network enterprise's multi-agent system, a method is proposed for forming key design solutions from the perspective of implementing various service categories using AAS (Asset Administrative Shells) and their specialization. At the technological implementation design level of MAS, a method for implementing software agents using a microservice software organization is proposed. The method presented for adapting reference and best MAS design models allows for the selection of appropriate design solutions from libraries of reference models and knowledge bases for subsequent refinement. This accelerates and improves the quality of the design process. The implementation of the developed technology for designing multi-agent systems will increase the adaptability of network enterprises to dynamically changing business needs, taking into account the interests and capabilities of all stakeholders.

Keywords: multi-agent systems, digital threads, digital twins, network enterprise, asset administrative shell, project ontology, microservice architecture

Citation: Telnov Yu.F., Kazakov V.A., Danilov A.V. (2024) Designing a multi-agent system for a network enterprise. *Business Informatics*, vol. 18, no. 3, pp. 70–86. DOI: 10.17323/2587-814X.2024.3.70.86

Introduction

To produce innovative products and services for the specific needs of customers, dynamically formed flexible network or Internet-distributed enterprises can be created, requiring the implementation of new management systems based on the use of modern digital and intelligent technologies. As a result of the creation of such enterprises, the life cycle of products and services provided should be shortened both at the stage of launching into the market and at the production stage, ensuring high quality and adaptability of product configurations for various categories of consumers [1].

The creation of network or virtual enterprises based on modern digital platforms, on the one hand, leads to an increase in the level of integration and cooperation of enterprises interacting within the overall network structure [2], and on the other hand, gives rise to new tasks of creating mechanisms for coordinating participants in network enterprises and selecting reliable partners, organizing joint ownership and determining the rights to use data, the unresolved nature of which can lead to a loss of confidence of potential participants in a network enterprise regarding the possibility of joint activities within the overall network structure [3]. Solving the problems of creating network enterprises in industry becomes more complicated, as a rule, due to a large number of cooperative connections, high resource intensity and large investment cycles.

The implementation of Industrial Internet of Things (IIoT) technologies, Digital Twins and Threads into the practice of digital transformation of enterprises based on the concept of the Industrie 4.0 creates objective prerequisites for increasing the efficiency of managing network interactions of enterprises carrying out joint activities through the creation of modern digital systems [4–7]. At the same time, Digital Twins and Threads are based on digital models which represent systems of mathematical and computer models that make it possible to display the information state, predict the behavior of simulated objects in real time and formulate decisions.

One of the effective approaches to the implementation of the listed technologies is the creation of digital systems for managing manufacturing and business processes of a network enterprise based on the use of multi-agent technologies – multi-agent systems (MAS) of a network enterprise. In works [8, 9] developed computer models based on the use of agent-based and discrete-event modeling methods which are built into the structure of Digital Twins to optimize production processes at various stages of the life cycle. To increase the efficiency of using computer models of manufacturing and business processes, it is necessary to ensure their interaction with systems for operational data collection using the IIoT, as well as their integration with other intelligent technologies based on decision-making rules, analysis of big data and machine learning [10–13].

The architectural complexity of the MAS determines the goal of the study to develop a comprehensive technology for designing a software implementation that would affect all levels of the network enterprise architecture in accordance with the IIRA and RAMI architectural frameworks [14, 15] and the use of a variety of decision support tools in computer models of agents. Existing works on the design of multi-agent implementation of digital systems mainly consider the functional level of design [16–19] and practically do not consider the design of MAS at other architectural levels.

This article solves the problems of analyzing the capabilities of Digital Twins and Digital Threads for creating network enterprises, their implementation using an MAS and developing MAS design technology at the levels of business and user requirements, functional design and implementation with mutual mapping between the levels of the results obtained. A feature of the proposed technology is the linking of MAS design stages by sequential mapping of design entities (categories) between architectural levels and adaptation of design solutions based on libraries of reference solutions and knowledge bases of the best usage precedents [20–22].

1. Analysis of the possibilities of using Digital Twins and Digital Threads to create network enterprises based on multi-agent technology

Digital twin technology is widely used in industry and allows you to manage enterprise assets (products, equipment, any resources) at different stages of their life cycle. At the same time, Digital twins not only reflect the current state of assets, but also allow, using a set of procedures, to model, predict and formulate decisions to optimize their behavior. From this point of view, Digital twins are an integrated system of data, models and tools for analysis and decision-making used throughout the entire life cycle of various assets [8, 23].

Due to the need to track and manage the behavior of not only individual assets, but also the dynamic processes in which they participate, there is an objective need to implement more complex production technologies based on digitalization which are reflected in the concept of Digital Threads. The Digital Thread concept involves the use of modern modeling and management tools that link the life cycle processes of interconnected assets and make it possible to improve the manufacturability, controllability and sustainability of production systems [24]. Digital Thread in the economic sense implements value chain management. In the RAMI architectural framework [15], Digital Thread is associated with interacting assets: production chains and supply chains.

The use of Digital Twin and Digital Thread provides the flexibility and adaptability needed to quickly develop and implement products while reducing risk. Thus, data obtained from existing or designed production systems can form the basis of improved models that will allow forecasting at both the component level and the asset level as a whole. Archiving digital asset descriptions can greatly facilitate any future required redesign of the production system. The combination of Digital Twins and Digital Thread constitutes a digital system for a specific production system or an entire network enterprise.

A unified information model of Digital Twin and Digital Thread can be represented using software open source, which allows for the implementation of digital technologies in complex projects [25]. This option ensures the implementation of Digital Twin and Digital Thread not only in digitalization projects of individual enterprises, but also in the creation of network enterprises through the integration of various software systems based on a single digital platform. Using Digital Twin and Digital Thread technologies, information is transferred from individual links in the value chain to the production system of a network enterprise, which makes it possible to monitor compliance with requirements and the impact of the results obtained on the efficiency of the entire network enterprise [26].

The dynamic nature of Digital Twin and Digital Thread technologies makes it natural to use multi-agent systems to organize the interaction of Digital Twins within the Digital Thread, while the tool for implementing Digital Twins are software agents, and the Digital Thread is the MAS as a whole. The issues of implementing production systems using multi-agent interaction technologies are quite well theoretically worked out [9, 10, 13, 16, 17, 27].

The modern development of Industrie 4.0 concepts, which has led to new architectures for Digital Twins organizing in the form of Asset Administrative Shells operating on common digital platforms, provides the opportunity to develop MAS on a new technological basis, primarily using microservice implementation of mechanisms for performing agent functions [18, 19, 28].

In the listed works, the main emphasis is on the functional implementation of MAS for creating digital and network enterprises and to a lesser extent devoted to the issues of their design technology. At the same time, the construction of multi-agent systems for interaction between participants in a network enterprise makes relevant the issues of creating an MAS design technology that takes into account the level of complexity of network enterprises being created. Previously developed MAS design technologies, for example, such as ASEME [29], RTMIAS [30], X-Machine [31], etc., are local in nature, based on component technology of object-oriented design, and in some cases use ontologies for developing interaction between agents, but for such complex systems as network enterprises, they are of little use.

From the point of view of applying the concept of Industrie 4.0, the design of the MAS of a network enterprise comes down to the design of Asset Administrative Shells (AAS), which correspond to agents that implement active and proactive modes of operation [21, 28, 32], as well as designing scenarios for their interaction within the unified Digital Thread through the exchange of messages between AAS in accordance with the protocols established by FIPA [33]. A distinctive feature of the implementation of AAS on the principles of multi-agent technology is the possibility of using knowledge bases to develop solutions, machine learning and simulation mechanisms for analyzing and interpreting events.

To implement software agents and their interaction, it is necessary to develop a set of functional AAS services, and for a digital platform that serves many interacting AAS, a set of infrastructure services that ensure the creation and registration of AAS and their users, commercial and information security of use and a number of other functions.

Taking into account the complexity of the process of creating an MAS of a network enterprise, this article proposes a design technology that is based on the consistent refinement of design solutions at the levels of architecture of a network enterprise in accordance with IIRA and RAMI frameworks [14, 15] using the method for adapting reference and best design models.

2. Stages of technology for designing an MAS for a network enterprise

The stages of designing an MAS for a network enterprise are well defined by points of view on the system architecture in accordance with IIRA [14]:

- Business Modeling – Business viewpoint;
- Design of manufacturing and business processes – Usage viewpoint;

- Functional design of MAS – Functional viewpoint;
- Design of implementation technology – Implementation viewpoint.

The listed points of view on architecture or architecture levels are interrelated: each subsequent level of architecture specifies the previous level of architecture in its own specific language and confirms the possibility of implementing the requirements formulated above. A feature of the proposed MAS design technology is the sequential decomposition and detailing of design solutions and iterative repetition of stages if the need arises.

The set of digital system entities used in the proposed MAC design technology at various levels of the IIRA architecture, as well as design tools, is presented in *Table 1.*

Table 1.

Entities and tools for MAS design at various levels of IIRA architectural framework

Entity	Levels of MAS architecture of a network enterprise			
	Business level	**Usage level**	**Functional level**	**Implementation level**
Subject	Participant, Role	Performer role	Interface (Boundary Class)	Client application
			Subject agent (AAS)	Composite microservice
Object	Assets	Product, Resource	Object agent (AAS)	Composite microservice
			Entity class	Database
Function	Activity / Activities	Process	Operation (method, service)	Microservice
			Entity class	Microservices registry
Modeling and Design Tools				
Model type	Value Network (St. Gallen); Use-Case diagram	Activity diagram; BPMN diagram; Usage Model (IIRA)	UML diagrams: Sequence; Class; State-chart	UML diagrams: Component; Deployment

By Value Network (network enterprise) participant we mean any enterprises or organizations participating in the value chain as subjects.

By Role we mean a specific behavior scenario performed by a value network participant in manufacturing or business processes. Moreover, the same participant can play different roles in the same process; for example, an enterprise can be a supplier of equipment and a service provider, and an operating enterprise can also be an operator of a private cloud infrastructure. Roles at the usage level can be refined to the level of specific performers.

At the functional level, the role is performed directly either by an actor (organizational unit) or by a software agent (subject). In the first case, an interface must be created for the subject, implemented in the form of a client application, through which it interacts with the MAS. In the second case, a Digital Twin is created for the subject which automates a number of functions of the actor, essentially replacing it. In this regard, it is proposed to represent the subject's Digital twin at the functional level in the same way as the asset's Digital twin – in the form of AAS, which is implemented by a composite microservice in the sense that a composite software component includes a registry of services corresponding to individual functions [34].

In accordance with RAMI, an Asset of a network enterprise will be understood as any physical or software objects (products – goods, services or their components, and resources – individual devices, equipment, production lines, production systems) [15], which are represented by AAS [33]. AAS of objects (products or resources) are proposed to be implemented using microservices [35, 36]. If any part of the information about assets does not require active management, then it is technologically implemented in the form of a passive database which is local for microservice or a database which is shared across multiple microservices.

From a business point of view, the Activity of a network enterprise will be understood as a certain function that produces flows of material assets, information and value (costs). Each activity is detailed as a process executed by the MAS, which consists of individual operations. The identified operations are specified as elements of the AAS structure which link through API to an implementation in the form of a microservice. In this case, the meta-description of the operation as an entity class is placed in the service registry [21, 34].

A combination of different tools is used to model and design MAS. This article uses the object-oriented modeling language UML as a comprehensive tool to present an end-to-end example of the MAS design process.

Let us consider the stages of technology for designing an MAS for a network enterprise, corresponding to IIRA viewpoints, in more detail.

Business modeling. Business Viewpoint determines the strategy for creating and operating a network enterprise. From this position, at the business modeling stage, stakeholders and their vision of the functioning of the enterprise in the context of the use of common digital platforms, as well as the values and goals of digitalization of manufacturing and business processes are determined. At this stage, business requirements for the designed multi-agent system are formed. The most important role at this stage is played by the design of business models for the functioning of network enterprises, specified in business scenarios [21, 37]. The St. Gallen value network framework is widely used as a notation for describing business models [38]. From the perspective of multi-agent implementation, the method of determining the main activities and their actors, which are subsequently supported by software agents and their processes, becomes of utmost importance. The core activities of value networks can be captured using UML use case diagrams.

The business modeling process begins with SWOT analysis of the proposed network enterprise organization. This identifies the strengths and weaknesses of digitalization from the perspective of using internal resources, as well as opportunities and threats from the perspective of the influence of the external environment. As a result of SWOT analysis, the company's vision, the main values being formed are determined and a tree of goals is built. For goals, sets of measures are formed to achieve them, including the creation of a new or customization of an existing software and hardware platform.

The selection of business models and corresponding application implementation scenarios is carried out according to the methodology described in [22, 39]. In accordance with this methodology, depending on the type of business model (the Industrial Internet of Things platform model, the Value adding services in operation model, the Data Trustee model) and the type of business process (product lifecycle management processes, production system lifecycle management processes, supply chain management, asset service management) applied scenarios for the implementation of a network enterprise are selected (Adaptive Factory, Value Creation Network "Innovative Product Development", Value Creation Network "Order-Controlled Production", Value-Based Scenario, etc.). The selection of an application scenario for the implementation of a network enterprise is carried out on the basis of a knowledge base of typical scenario models, organized using the ontology of digital transformation, and subsequent analysis of network effects.

The selected application scenario is presented in the form of a value network model and is adapted to the operating conditions of a particular enterprise. In the value network model, the composition of enterprise participants and their roles is first determined (parent enterprise, subcontractors, functional service providers, platform operators, software developers, system integrators, etc.). The performance of activities by participants in a network enterprise in accordance with their roles can be represented in the form of a use case diagram. An example of a use case diagram for a value-based application scenario [40] is presented in *Fig. 1*.

Design of manufacturing and business processes in accordance with the point of view of system use (Usage Viewpoint). At this stage, functional requirements for the organization of manufacturing and business processes are determined in terms of specifying participants in network enterprises and their roles in various activities (processes). In this case, activities are determined from the perspective of initiation conditions, task workflow, resulting effects and restrictions on the execution of processes, and executor roles are assigned to tasks [37, 40].

The transition from the value network model to models of manufacturing and business processes is carried out in accordance with the following method: each value flow and accompanying information and financial flows correspond to manufacturing or business processes, which can be modeled in the form of an activity diagram or BPMN diagram reflecting the assignment of roles process participants for performing specific operations (services). As a result of activity diagrams design, functional requirements are formed for the future composition of software agents implementing the allocated roles, and the composition of operations (services) of the corresponding AAS.

Functional design of MAS, reflecting a functional point of view on the system architecture (Functional viewpoint). At this stage, a method is presented for the formation of basic design solutions for building an MAS of a network enterprise from the position of implementing various categories of services using AAS and

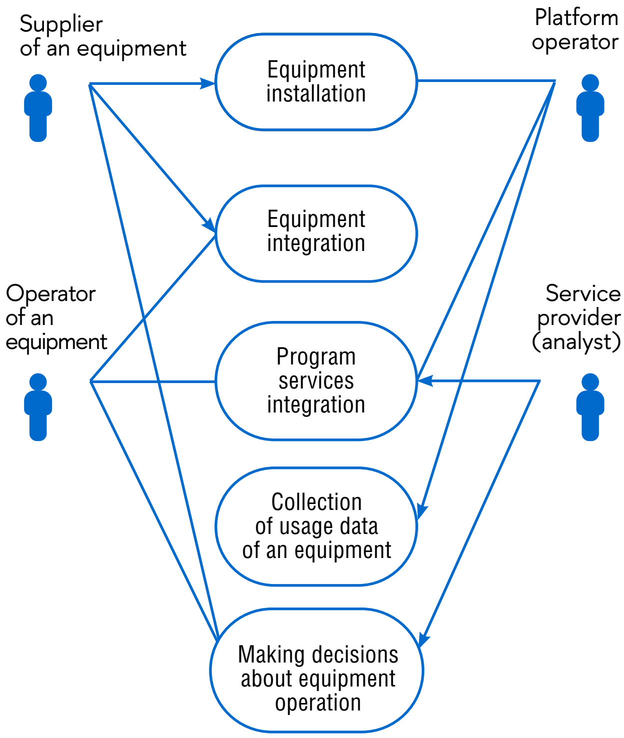

Fig. 1. Example of a use-case diagram for a value-based scenario (VBS).

their specialization: interaction with external business services (ERP, MES, PLM, etc.); applications (Application) – functional services of digital systems; system services (System Management) – platform services; asset integration services – physical devices, products (Control); communication services of digital system components (*Table 2*) [14, 34].

From a functional viewpoint, in accordance with the concept of the Platform Industrie 4.0, each component of a network enterprise represents an asset and its AAS [41, 42]. The structure of AAS includes a set of properties, operations and events which can also be divided into submodels [42]. At the same time, the information part (passive) of AAS in the form of a set of properties allows you to reflect the dynamic information model of the asset, and the operational (active) part allows you to interact with assets, other AAS and external applications, and to execute functional services. In accordance with the types of assets, software

Table 2.

Sublevels of the IIRA functional level

Functional level sublevels	Description
Business level (BUSINESS DOMAIN)	Services for exchanging information with external applications of various information systems: ERP, CRM, PLM, MES, HRM, etc.
Application level (APPLICATION DOMAIN)	Functional services performing the functions of monitoring, diagnostics, forecasting, coordination, optimization, support, management of manufacturing and business processes
Operating level (SYSTEM MANAGEMENT DOMAIN)	General infrastructure services for managing a software system (platform): deployment, configuration, monitoring, diagnostics and updating of its components, as well as orchestration for coordinating the operation of various system components
Information level (INFORMATION DOMAIN)	Services for collecting, cleaning, syntactic and semantic transformations, storing and issuing data for functional components. Auxiliary data management services: data security, data access control and data rights management, backup and recovery, etc. Services that implement universal methods of data analysis, incl. Big Data analysis methods, machine learning, simulation, knowledge extraction
Communication level	Considered between services (components) at all sublevels
Control level (CONTROL DOMAIN)	Services interaction of software components with physical devices (assets): collecting data using sensors from physical devices, monitoring operation and executing control commands on physical devices.
Physical systems	Physical devices: equipment parts, equipment, production lines, production systems (factories), network enterprises (Connection World) Software: software libraries, knowledge bases, ontologies, repositories

agents that are represented by AAS are divided into product agents and resource (equipment) agents [43]. A digital platform can be represented by AAS with a set of system services, the actor for which is the operator (administrator) of the platform. Digital Twins of subjects managing the production system are also represented by agents – AAS.

AAS as software agents are well represented by classes in the object-oriented paradigm. The interaction of software agents in the process of executing individual activities with the addition of the necessary interface classes (Boundary) and entity classes associated with the database (Entity) is displayed in the form of a sequence diagram (Sequence Diagram). *Figure 2* shows a fragment of a sequence diagram reflecting the process of generating recommendations for equipment operation in accordance with a Value-Based Scenario.

The process is launched through the boundary class – client application "Data Analysis" (Boundary), and the received recommendation is recorded in the local database through the entity class "Recommendation" (Entity).

On the generated class diagram, AAS "Analyst Agent" implements the functions of a software agent at the Functional Domain level. The service (operation) "Perform Forecast" using one of the machine learning methods, for example, a neural network, performs fault prediction, and the service "Form Recommendation" using a set of rules and/or a simulation model based on agent-based and discrete-event modeling methods. implemented, for example, in the AnyLogic system [9].

Equipment Agent AAS is a software agent that collects data into a historical dataset to predict the condition and updates the condition of the asset. In this sense, AAS performs the functions of the control level (Control Domain). Information services for converting data formats, checking access security, etc., associated with the execution of the "Data Analysis" and "Save Recommendation" services, are called within these services. Similarly, the communication service of AAS "Analyst Agent" and AAS "Equipment Agent" is called within the "Prepare Dataset" service.

Design of technological implementation (Implementation viewpoint). The point of view on the implementation of a digital system reflects the physical construction of the system from the components being created. Considering the autonomy of the main components – software agents, their distribution in a computer network, the need for independent access to infrastructure services of a common digital platform, the presence of local databases as part of the components, it is proposed to implement MAS based on microservices technology in a cloud containerization environment. Architectural patterns can be used to provide examples and references for conceptualizing real-life IIoT architectures.

The use of microservices is proposed to be carried out at two levels: at the level of AAS in the form of composite microservices and at the level of microser-

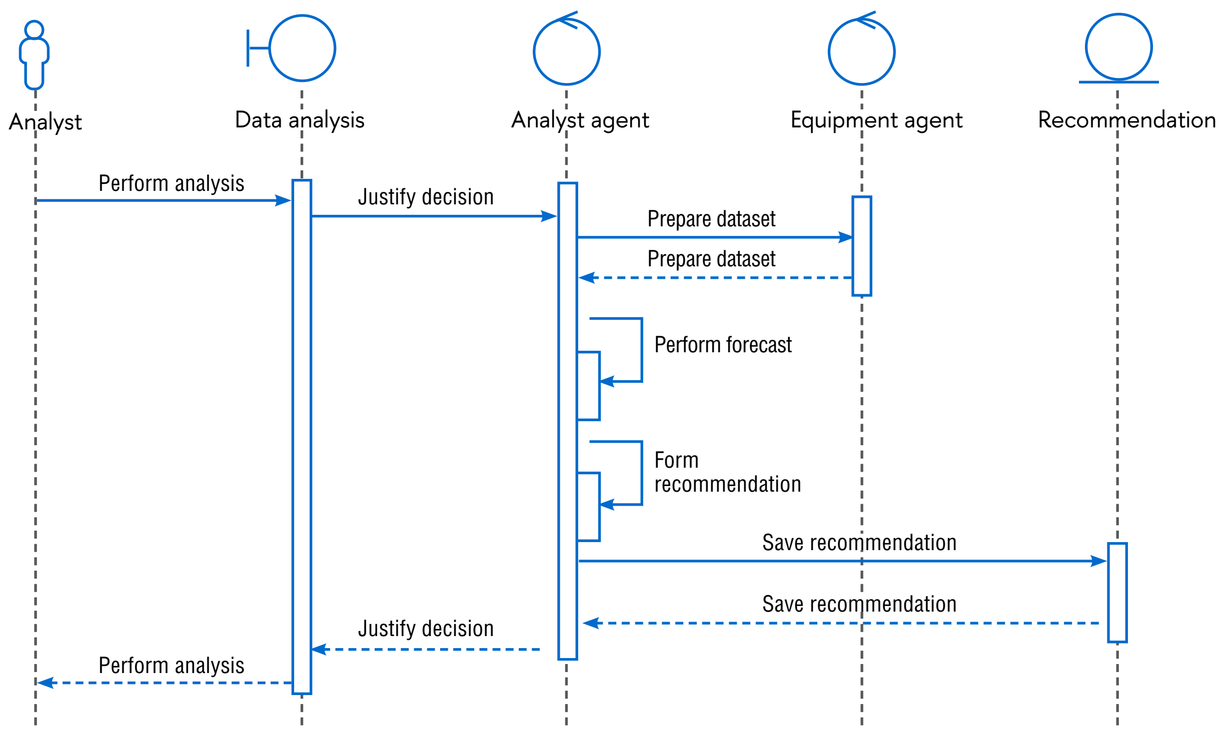

Fig. 2. Fragment of a sequence diagram for the process of recommendations formation for equipment operation.

vice implementation of operations (services, methods) of the AAS. In the second case, the operation in the information part of the AAS is connected via an API interface with a microservice which is stored in a dedicated library of the AAS, organized using a service registry (Registry) [34].

Integration and deployment of microservices is carried out on a technology platform for managing container microservices. An example of a deployment diagram of all software of a digital multi-agent system for the process of recommendation formation for equipment operation is presented in *Fig. 3*.

As an architectural pattern for implementing a network enterprise in the Industrial Internet concept (IIRA), we propose to use the Digital Twin pattern as an intermediate software layer (between the application and the physical world) [14]. This pattern implies the construction of industrial applications based on digital twins, which in turn are implemented based on standard services of the IIoT platform.

3. Method for adapting design models for a multi-agent system of a network enterprise

The MAS design technology is based on the use of a methodology for adapting design models (design patterns). For this purpose, knowledge bases (libraries) of design models are organized; they are systematized in accordance with a set of dictionaries (ontologies).

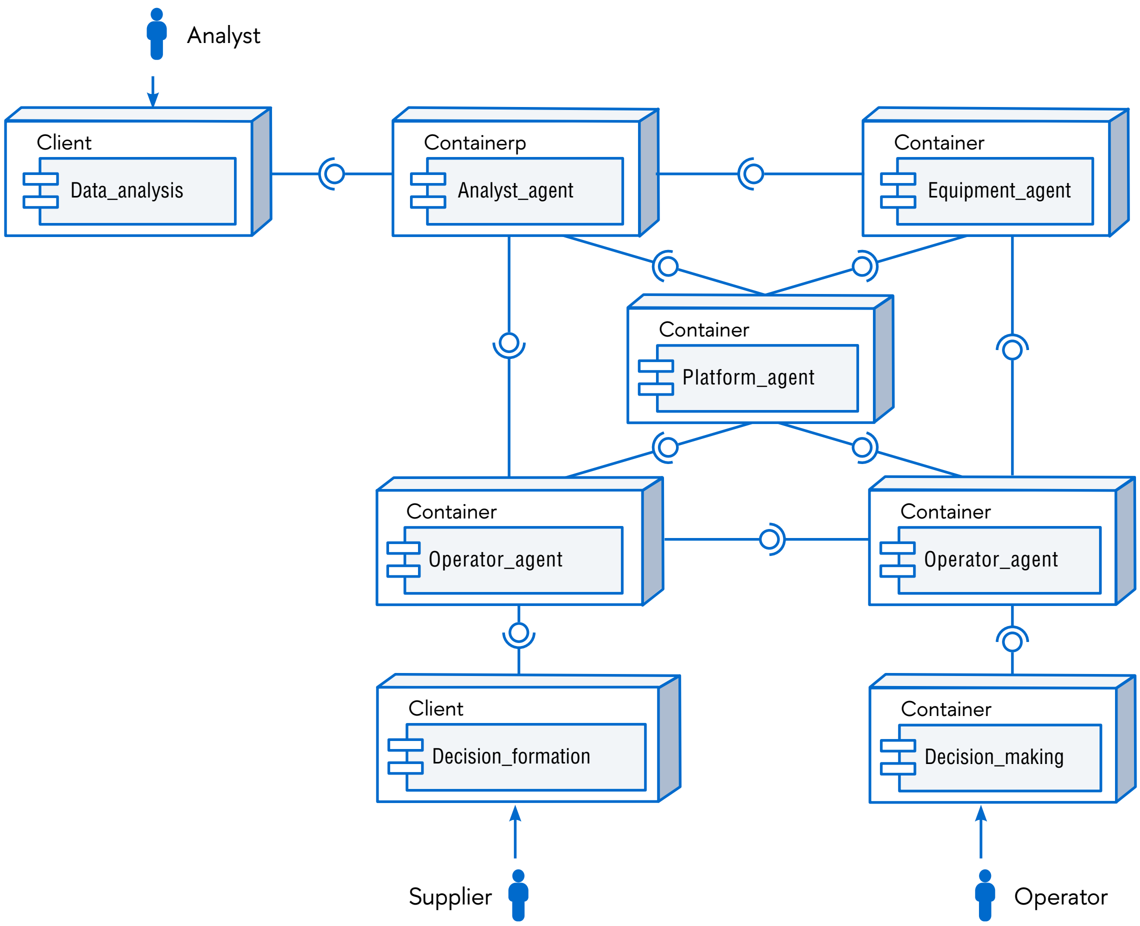

Fig. 3. MAS deployment diagram.

In the State Standard "Structure of the Digital Factory" [20], design patterns are specified in the asset class libraries of the digital factory, and a set of dictionaries is defined in accordance with which asset classes are built. In the materials of the Plattform Industrie 4.0 project [21], design models (patterns) are presented in a library of functional blocks. In [22], it was proposed to use not only reference models of application scenarios from the knowledge base, but also to accumulate models of precedents related to completed projects and use the hierarchy of ontologies to organize access to these models.

In this work, the approach to the use of MAS design patterns is developed from the point of view of their application at all stages of design technology. Method for adapting multi-agent system design models comes down to selecting precedents of appropriate models from libraries of reference models and knowledge bases and their subsequent refinement (*Fig. 4*).

To reflect the current state of the MAS creation project, a Project Repository is organized, which records the state of the project after each stage. At the same time, the Project Ontology is also clarified and developed.

Project organization at the initiation stage begins with defining the ontology of a network enterprise and selecting appropriate dictionaries from a set of ontologies that can be implemented by various standardization bodies, consortiums and research projects, and including upper ontologies, problem ontologies and domain ontologies. The selected ontologies form the prototype of the Project Ontology (or network enterprise ontology). In addition, as the project is implemented, the ontologies of external participants in the network enterprise are also connected to the existing ontology through links. A necessary condition for their unification is the alignment of external ontologies with the Project Ontology.

The subsequent development of the repository and ontology metadata involves versioning not only the data (a "snapshot" of the parameters of the system models is implemented as of a certain point in time), but also the ontology of the network enterprise (Project Ontology). This option allows you to use all the data accumulated in the historical perspective for the development of a project for the creation and operation of a network enterprise. Mechanisms that implement versioning, alignment and development of ontologies, their connection with each other, must be implemented as independent services within the platform that supports the functioning of a multi-agent system.

Each subsequent stage of designing a multi-agent system generates at the output a set of input parameters for the formation of the next stage, within which these input parameters serve as the basis for selecting new models from a library of reference models. According to the levels of model typification, which are stored in the library of reference models (design patterns), models can be both quite abstract, high-level, and specific to a particular subject area.

A distinctive feature of the proposed method for adapting MAS design models of a network enterprise is the use, along with a library of reference models (design patterns), also of a knowledge base aimed at preserving the formed and tested structures and descriptions of real business models, business scenarios, production and business processes, software agents (AAS), sets of services (precedent models).

Precedent models are presented as descriptions of cases that are stored in the system with reference to the description of the initial conditions (requirements), as well as the results of the work of the network enterprise, the quality characteristics of the products released (products) and the resulting economic effect. Thus, when searching for suitable models, not only reference models (design patterns of a certain type) are selected, but also adapted precedent models, taking into account the degree of proximity (maximum similarity value – S) of the corresponding characteristics of the models $M_i \in M$ and the available input and required output parameters that define the problem situation C^{In} [44]:

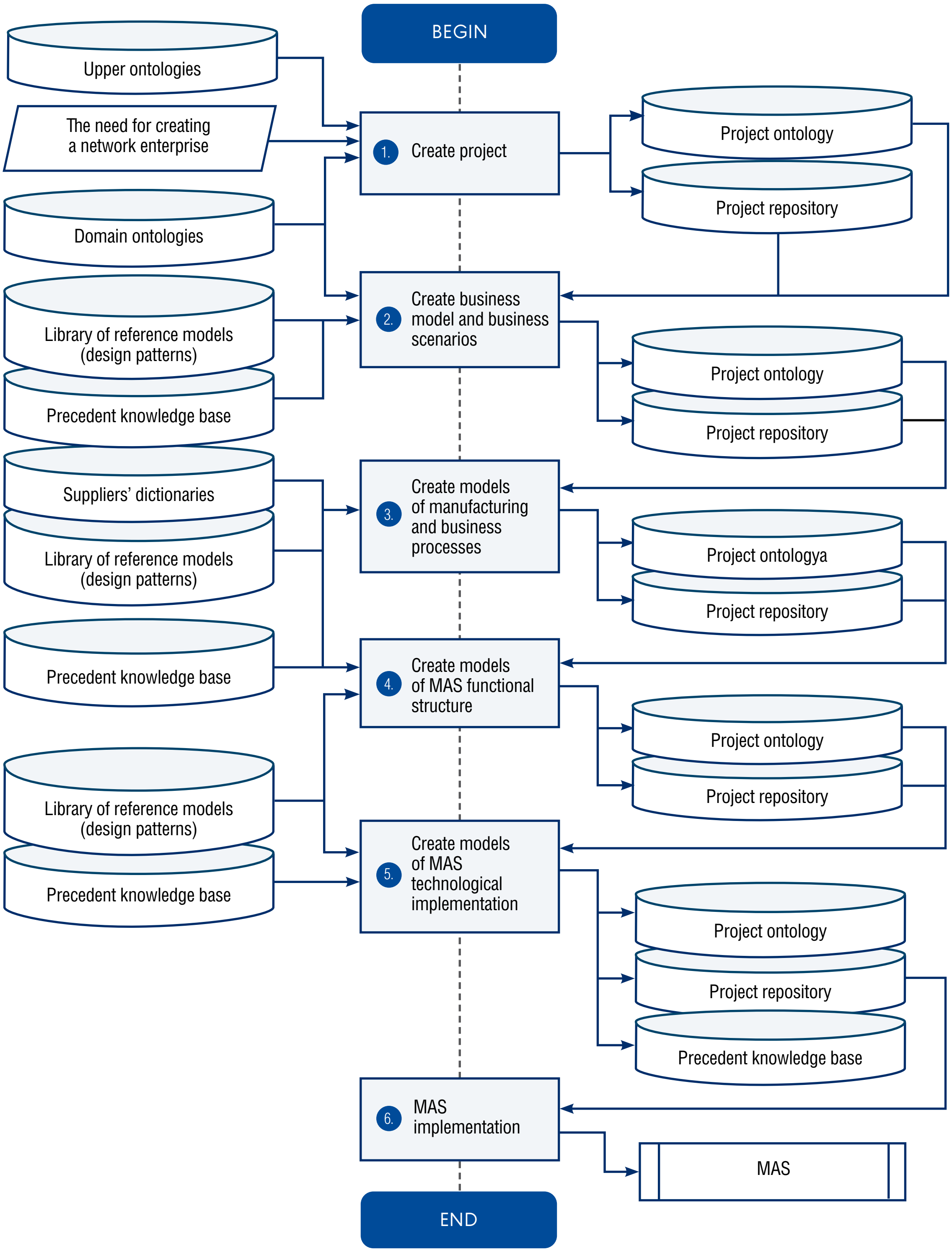

Fig. 4. Method for adapting MAS design patterns.

$$S(C^{In}, M) = \max_i \left(\frac{\sum_{j=1}^{n} w_j \cdot \mathrm{sim}(f_j^{In},\ f_{ij}^{M})}{\sum_{j=1}^{n} w_j} \right),$$

where

$i = \overline{1, N}$, where N is the total number of models available in the library and Precedent Knowledge Base;

$j = \overline{1, CR}$, where CR is a constant number of compared description elements (properties, relationships);

f_{ij}^{M} – j-th property (relationship) describing the i-th model,

f_j^{In} – j-th property (relationship), describing the set of required input or output parameters,

w_j – weight of the j-th property (relationship),

sim – similarity function f_j^{In}, f_{ij}^{M}.

After completing the design process of a multi-agent system, the design results reflected in the Project Repository are used to perform the following stages of creating an MAS: software development, testing and implementation. The result of the design process is also placed in the Precedent Knowledge Base for later use in other projects.

Conclusion

As a result of the study, we can conclude that the use of MAS fully ensures the creation of effective network enterprises based on the implementation of the principles of the Industrie 4.0 and the use of digital twins and digital thread, providing information collection, modeling and planning of asset behavior, organization and control of manufacturing and business processes.

The proposed methods and technology for designing MAS in accordance with IIRA and RAMI architectural frameworks provide, at the business and usage levels, for the construction of basic application scenarios for the use of MAS and the roles of actor-agents, the formation of structures of manufacturing and business processes; at the functional level – building a set of functional components in the form of AAS and models of their interaction in a common information space; at the technological level of implementation – adaptation of microservice implementation templates to the specific conditions of building a network enterprise.

A distinctive feature of the design technology so developed is the interconnected representation of all used categories (entities) at various levels of the architecture, which allows for a coordinated transition between the stages of MAS design.

The complex nature of the proposed technology for designing MAS interaction between participants in a network enterprise determines the organization of effective participation of all stakeholders in the creation of a network enterprise focused on the implementation of a business strategy, taking into account the adaptation of reference and best models of application scenarios, functional components and microservice structures using design patterns libraries, knowledge bases and ontologies. Consistent display of design results between different levels of architecture in the project repository allows you to fully implement functional and non-functional requirements, taking into account the available information and computing resources.

The proposed methodology for adapting design models for a multi-agent system of a network enterprise develops an approach to adapting MAS templates from libraries of reference solutions and knowledge bases of the best precedents for use at all stages of design technology.

The application of the presented technology for designing an MAS of a network enterprise will help improve the level of quality and reliability of the functioning of a network enterprise, adaptability to dynamically changing business needs and the capabilities of all stakeholders. ■

Acknowledgements

The research was supported by a grant from the Russian Science Foundation (project no. 22-11-00282[1]).

[1] https://rscf.ru/project/22-11-00282/

References

1. Matthyssens P. (2019) Reconceptualizing value innovation for Industry 4.0 and the Industrial Internet of Things. *Journal of Business & Industrial Marketing*, vol. 34, no. 6, pp. 1203–1209. https://doi.org/10.1108/JBIM-11-2018-0348

2. Feofanov A.N. Bondarchuk E.Yu., Tyasto S.A. (2018) Organization of a virtual enterprise – the future of production. *Bulletin of MSTU "Stankin"*, no. 3 (46), pp. 101–105 (in Russian).

3. Müller J.M. (2019) Antecedents to digital platform usage in Industry 4.0 by established manufacturers. *Sustainability*, vol. 11, no. 4, article 1121. https://doi.org/10.3390/su11041121

4. Golovin S.A., Lotsmanov A.N., Pozdneev B.M (2021) The Russian Federation Industry 4.0 program is a chance not to fall behind forever in the field of industrial production. *World of Information Technologies*, no. 1–2, pp. 38–40 (in Russian).

5. Borovkov A.I., Prokhorov A., Lysachev M. (2020) *Digital twin. Analysis, trends, world experience*. Moscow: Alliance Print (in Russian).

6. Rosstandart (2021) *National Standard of the Russian Federation GOST R 57700.37–2021. Computer models and simulation. Digital twins of products. General provisions* (in Russian).

7. National Institute of Standards and Technology (2018) *Digital thread for smart manufacturing*. Available at: https://www.nist.gov/programs-projects/digital-thread-smart-manufacturing (accessed 1 August 2024).

8. Makarov V.L., Bakhtizin A.R., Beklaryan G.L. (2019) Developing digital twins for production enterprises. *Business Informatics*, vol. 14, no. 1, pp. 7–16. http://doi.org/10.17323/1998-0663.2019.4.7.16

9. Makarov V.L., Bakhtizin A.R., Beklaryan G.L., Akopov A.S. (2021) Digital plant: methods of discrete-event modeling and optimization of production characteristics. *Business Informatics*, vol. 15, no. 2, pp. 7–20. http://doi.org/10.17323/2587-814X.2021.2.7.20

10 Gorodetsky V.I. (2019) Behavioral models of cyberphysical systems and group management. Basic concepts. *News of the Southern Federal University. Technical Sciences*, no. 1 (203), pp. 144–162.

11 Corsini R.R., Costa A., Fichera S., Framinan J.M. (2024) Digital twin model with machine learning and optimization for resilient production–distribution systems under disruptions. *Computers & Industrial Engineering*, vol. 191, article 110145.

12. Kabaldin Yu.G., Shatagin D.A., Anosov M.S., Kuzmishina A.M. (2019) Development of digital twin of CNC unit based on machine learning methods. *Vestnik of Don State Technical University*, vol. 19, no. 1, pp. 45–55 (in Russian). https://doi.org/10.23947/1992-5980-2019-19-1-45-55

13. Skobelev P., Mayorov I., Simonova E., Goryanin O., Zhilyaev A., Tabachinskiy A., Yalovenko V. (2020) Development of models and methods for creating a digital twin of plants within the cyber-physical system for precision farming management. *Journal of Physics: Conference Series*, vol. 1703, pp. 12–22. https://doi.org/10.1088/1742-6596/1703/1/012022

14. Industry IoT Consortium (2022) *The industrial internet reference architecture*. Available at: https://www.iiconsortium.org/wp-content/uploads/sites/2/2022/11/IIRA-v1.10.pdf (accessed 1 August 2024).

15. Plattform Industrie 4.0 (2018) *Plattform Industry 4.0. Reference architectural model Industry 4.0 (RAMI4.0) – An introduction*. Available at: https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/rami40-an-introduction.html (accessed 1 August 2024).

16. Seitz M., Gehlhoff F., Cruz Salazar L.A., Fay A., Vogel-Heuser B. (2021) Automation platform independent multi-agent system for robust networks of production resources in industry 4.0. *Journal of Intelligent Manufacturing,* vol. 32, pp. 2023–2041.

17. Karnouskos S., Leitao P., Ribeiro L., Colombo A.W. (2020) Industrial agents as a key enabler for realizing industrial cyber-physical systems: Multiagent systems entering Industry 4.0. *IEEE Industrial Electronics Magazine*, vol. 14, no. 3, pp. 18–32. https://doi.org/10.1109/MIE.2019.2962225

18. Vogel-Heuser B., Ocker F., Scheuer T. (2021) An approach for leveraging Digital Twins in agent-based production systems. *at – Automatisierungstechnik*, vol. 69, no. 12, pp. 1026–1039. https://doi.org/10.1515/auto-2021-0081

19. Telnov Yu.F., Kazakov V.A., Danilov A.V., Denisov A.A. (2022). Requirements for the software implementation of the Industrie 4.0 system for creating network enterprises. *Software & Systems*, vol. 35, no. 4, pp. 557–571 (in Russian).

20. Rosstandart (2022) *National Standard of the Russian Federation GOST R 70265.1–2022. Industrial-process measurement, control and automation. Digital factory framework. Part 1. Basic provisions* (in Russian).

21. Plattform Industrie 4.0 (2019) *Discussion Paper: Usage View of the Asset Administration Shell*. Available at: https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/2019-usage-view-asset-administration-shell.html (accessed 1 August 2024).

22. Telnov Yu.F., Kazakov V.A., Bryzgalov A.A., Fiodorov I.G. (2023) Methods and models for substantiating application scenarios for the digitalization of manufacturing and business processes of network enterprises. *Business Informatics*, vol. 17, no. 4, pp. 73–93. http://doi.org/10.17323/2587-814X.2023.4.73.93

23. Segovia M., Garcia-Alfaro J. (2022) Design, modeling and implementation of digital twins. *Sensors*, vol. 22, no. 14, article 5396. https://doi.org/10.3390/s22145396

24. Bajaj M., Hedberg T. (2018) System lifecycle handler – Spinning a digital thread for manufacturing. *INCOSE International Symposium*, vol. 28, no. 1, pp. 1636–1650. https://doi.org/10.1002/j.2334-5837.2018.00573.x

25. Idaho National Laboratory (2020) *Deep-Lynx*. Available at: https://github.com/idaholab/Deep-Lynx (accessed 1 August 2024).

26. Bonham E., McMaster K., Thomson E., Panarotto M., Müller J.R., Isaksson O., Johansson E. (2020) Designing and integrating a digital thread system for customized additive manufacturing in multi-partner kayak production. *Systems*, vol. 8, no. 4, article 43. https://doi.org/10.3390/systems8040043

27. Tarassov V.B. (2019) Enterprise total agentification as a way to Industry 4.0: Forming artificial societies via goal-resource networks. Proceedings of the *Fourth International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'19). Advances in Intelligent Systems and Computing (AISC)*, vol. 1156, pp. 26–40.

28. Sakurada L., Leitao P., de la Prieta F. (2022) Agent-based asset administration shell approach for digitizing industrial assets. *IFAC-PapersOnLine*, vol. 55, no. 2, pp. 193–198.

29. Spanoudakis N.I., Moraitis P. (2007) The agent systems methodology (ASEME): A preliminary report. *Computer Science*.

30. Julian V., Botti V. (2004) Developing real-time multi-agent system. *Integrated Computer-Aided Engineering*, vol. 11, no. 2, pp. 135–149. https://doi.org /10.3233/ICA-2004-11204

31. Eleftherakis G., Kefalas P., Kehris E. (2011) A methodology for developing component-based agent focusing systems on component quality. Proceedings of the *Federated Conference on Computer Science and Information Systems (FedCSIS 2011), Szczecin, Poland, 18–21 September 2011*, pp. 561–568.

32. Plattform Industrie 4.0 (2020) *Digital Twin and Asset Administration Shell Concepts and Application in the Industrial Internet and Industrie 4.0*. Available at: https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/Digital-Twin-and-Asset-Administration-Shell-Concepts.pdf (accessed 1 August 2024).

33. Foundation for Intelligent Physical Agents (2002) *FIPA ACL message structure specification*. Available at: http://www.fipa.org/specs/fipa00061/SC00061G.pdf (accessed 1 August 2024).

34. Plattform Industrie 4.0 (2021) *Functional view of the asset administration shell in an Industrie 4.0 system environment*. Available at: https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/Functional-View.html (accessed 1 August 2024).

35. Lewis J., Fowler M. (2014) *Microservices. A definition of this new architectural term*. Available at: https://martinfowler.com/articles/microservices.html (accessed 1 August 2024).

36. Richardson C. (2018) *Microservices Patterns: With examples in Java*. Manning Publications.

37. Plattform Industrie 4.0 (2017) *Exemplification of the Industrie 4.0 application scenario value-based service following IIRA structure*. Available at: https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/exemplification-i40-value-based-service.pdf (accessed 1 August 2024).

38. Gassmann O., Csik M., Frankenberg K. (2014) *The business model navigator: 55 models that will revolutionise your business*. FT Press.

39. Telnov Yu.F., Bryzgalov A.A., Kozyrev P.A., Koroleva D.S. (2022) Choosing the type of business model to implement the digital transformation strategy of a network enterprise. *Business Informatics*, vol. 16, no. 4, pp. 50–67. http://doi.org/10.17323/2587-814X.2022.4.50.67

40. Plattform Industrie 4.0 (2018) *Usage viewpoint of application scenario value-based service*. Available at: https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/hm-2018-usage-viewpoint.html (accessed 1 August 2024).

41. Rosstandart (2021) *National Standard of the Russian Federation GOST R 59799–2021. Smart manufacturing. Reference architecture model industry 4.0 (RAMI 4.0)* (in Russian).

42. Plattform Industrie 4.0 (2022) *Details of the asset administration shell – Part 1. The exchange of information between partners in the value chain of Industrie 4.0*. Available at: https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/Details_of_the_Asset_Administration_Shell_Part1_V3.html (accessed 1 August 2024).

43. Telnov Yu.F., Kazakov V.A., Danilov A.V., Bryzgalov A.A. (2023) Network enterprises: Production and business process models based on multi-agent systems. *Software & Systems*, vol. 36, no. 4, pp. 632–643.

44. Kolodner J. (1993) *Case-based reasoning*. Morgan Kaufmann.

About the authors

Yury F. Telnov

Dr. Sci. (Econ.), Prof.;

Head of the Department of Applied Informatics and Information Security, Plekhanov Russian University of Economics, 36, Stremyanny Lane, Moscow 117997, Russia;

E-mail: Telnov.YUF@rea.ru

ORCID: 0000-0002-2983-8232

Vasiliy A. Kazakov

Cand. Sci. (Econ.);

Associate Professor, Department of Applied Informatics and Information Security, Plekhanov Russian University of Economics, 36, Stremyanny Lane, Moscow 117997, Russia;

E-mail: Kazakov.VA@rea.ru

ORCID: 0000-0001-8939-2087

Andrey V. Danilov

Senior Lecturer, Department of Applied Information Technologies and Information Security, Plekhanov Russian University of Economics, 36, Stremyanny Lane, Moscow 117997, Russia;

E-mail: Danilov.AV@rea.ru

ORCID: 0000-0002-0433-9701

DOI: 10.17323/2587-814X.2024.3.87.107

The impact of artificial intelligence on re-purchase intentions: the mediation approach

Raed N. Alkaied
E-mail: Raedalkaied@bau.edu.jo

Shadi A. Khattab
E-mail: Shadikhattab@bau.edu.jo

Ishaq M. Al Shaar
E-mail: I.shaar@bau.edu.jo

Mohammed K. Abu Zaid
E-mail: Mohammed_abu_zaid@bau.edu.jo

Sakher A.I. Al-Bazaiah
E-mail: bazaiah1@bau.edu.jo

Al Balqa Applied University, Salt, Jordan

Abstract

Purchases made on online platforms have heavily incorporated artificial intelligence (AI) to shape consumer purchasing behavior. To investigate re-purchase intentions, this study combines AI, social media engagement, conversion rate optimization, brand experience and brand preference. A survey was conducted with a questionnaire sent to 355 people who had at least once purchased or used services offered online from any site associated with aviation. The questionnaire was analyzed using structural equation modeling. Utilizing Amos V.22, the study hypotheses were assessed. The empirical results show that social media engagement, brand experience, brand preference and conversion rate optimization were all impacted by AI. Conversion rate optimization and social media interaction also have an impact on brand preference and experience. Re-purchase intention is influenced by brand preference and brand experience. Additionally, the association between AI and re-purchase intention was mediated by social media engagement, brand experience, conversion rate optimization and brand preference. The study will support airline companies in developing AI and creating more effective branding and marketing campaigns to increase customer intention to re-purchase. This study discovered that the use of AI in marketing significantly improved brand preference, which subsequently affected consumers' desire to make additional purchases. Furthermore, to improve long-term commercial performance and brand attractiveness, the airline should focus brand-building efforts on AI. Thus, the airline ought to make greater investments in AI and booking service technology, both to draw in new business and to strengthen existing ones.

Keywords: artificial intelligence, conversion rate optimization, social media engagement, brand experience, brand preference, re-purchase intention

Citation: Alkaied R.N., Khattab S.A., Al Shaar I.M., Abu Zaid M.K., Al-Bazaiah S.A.I. (2024) The impact of artificial intelligence on re-purchase intentions: the mediation approach. *Business Informatics*, vol. 18, no. 3, pp. 87–107. DOI: 10.17323/2587-814X.2024.3.87.107

Introduction

Every facet of our real lives – both individually and collectively – has been impacted by technology, including both the real and virtual worlds [1]. Among its most crucial concerns was the various methods for increasing public awareness. According to predictions made by [2], the economic impact of artificial intelligence (AI) would increase from $20.82 billion in 2020 to $15 trillion in 2030.

One of the most significant technical advancements is AI, whose applications have completely transformed a wide range of societal fields [3, 4]. AI technologies are described as "natural genetic predisposition genetic inheritance or learned skills that form the essence of individual personalities" [5]. AI technologies rely on pre-defined computer programs, algorithms, and function similarly to the human mind in making decisions [6, 7].

Businesses may improve the customer experience by identifying innovation, developing strategies and identifying long-term solutions through business automation that leverages AI. Critical decisions can also be made with AI in a corporate environment that is surprisingly competitive and unstable [2].

According to [8], AI chatbot content recommendations are now a part of marketing AI activities. They also help to boost customer engagement on social media platforms, give online users a personalized experience and increase the likelihood that suggested goods and services will be purchased [2, 9]. For instance, Amazon is leading the way in utilizing AI technology and extending its use beyond object recognition, language understanding and conversation to include search and suggestion. This increases the conversion rate toward product purchases by personalizing and refining the recommendations of related and complementary products in real-time [10]. AI-driven digital businesses are attempting to interact with clients on social media in an effort to build, maintain and nurture enduring client relationship [11]. As a result, it is important to recognize the growing significance of online shopping. By 2025, the worldwide e-commerce market is expected to reach $1.2 trillion, growing at a rate three times faster than traditional commerce [12]. Marketers use customer engagement to get customers' attention by providing them with valuable knowledge [13]. Through their customer experience, marketers want to keep their goods or services at the front of consumers' minds. Social networking is one of the best platforms for connecting with customers. Customers should be able to use social media to interact with businesses [14]. Clients that are happy with the goods and services will write content for social media platforms. Companies may alter their current goods and services in response to negative customer feedback on social media [15]. These days, it's commonplace to see creative marketing. Marketers can expedite intelligent marketing by utilizing AI [9]. In order to boost the rate of conversion (from user to customer), businesses can also track customer opinions on social media and utilize that data to tailor marketing initiatives for each individual customer [16].

Social media consumer conversion is a continuous process rather than a one-time occurrence. The association between conversion rate and customer purchase intentions is not well understood [9]. Businesses are utilizing AI to forecast customer behavior as more and more consumers make purchases online. With a major influence on consumer decision-making, AI has been a crucial part of the digital transformation. AI technologies can be leveraged to entice consumers to make exciting new purchases [17].

AI technology has the potential to enhance user experiences in interactive environments and foster faster response times for products and services [18].

According to [19], intelligent service bots have become increasingly prevalent in gauging customer experiences with products and services in recent times. One of the primary motivations for implementing AI is to enhance customer experience, as AI technologies are becoming a more significant aspect of our daily lives and form the foundation for novel value propositions and unique consumer experiences. Delivering improved customer experiences is therefore essential for strengthening the bond between consumers and brands as well as for promoting brand distinctiveness [20].

Thus, businesses employ hardware, software, networks and AI for a range of objectives, including improving customer experience and fostering continuous harmonization and collaboration among stakeholders [21]. Based on whether the AI offers the services that customers have asked for, the consumer experience will differ. According to [22], customers felt more intellectual and sensory experiences when AI offered the services, and vice versa, when humans supplied the services, customers felt more emotional experiences. Even though AI and human services differ in how they are experienced, AI services are crucial to giving clients an enjoyable journey [23].

If customers receive the proper experience from the company, they feel content and joyful thanks to AI procedures [24]. A pleasant customer experience will create positive value for the company in terms of brand preference, helping companies achieve excellence and competitive advantage [25]. Research on how AI affects brands is scarce and dispersed, despite the significance of AI for consumer-brand connections. While improvements in technology may save customers time and effort during transactions, errors and a lack of human support can still generate dissatisfaction [20]. Therefore, it's still unknown how AI will affect branding.

Most marketers lack a strong knowledge of AI and how it may assist both organizations and consumers, despite expanding research in this field [26]. A road map for effective AI initiatives is necessary, according to comprehensive AI frameworks and empirical research, particularly in the area of digital marketing [27]. Furthermore, the literature that has already been written has not looked at AI in the setting of interactive marketing, where sellers and buyers work together to impact marketing choices that encourage active consumer participation, communication and interaction [18]. Accordingly, the real value of AI is not in the technology per se, but rather in the way it is applied to build robust, interactive buyer-seller interactions that are based on generating value together and keeping commitments [28]. There are few studies assessing how AI and digital innovations affect social media customer involvement. More investigation into the ways AI-powered marketing tools affect consumer views, opinions, and actions is advised by [29].

1. Literature review and hypothesis development

1.1. Artificial intelligence

Nowadays, online platforms are used for the majority of the shopping process. When it comes to buying things online, trust and client awareness are crucial factors. Organizations are working to get the most out of the enhanced trust and intent that customers have toward specific products and services as a result of the experience that AI has given them [30]. Studying client habits, purchasing patterns, behaviors and choices are only a few of the activities and functions where AI in marketing has been demonstrated to be widely applied [7]. Personalizing advertising messaging [31] in addition to tailoring items and other offerings to fit client demands and managing and altering prices in real-time in response to customer demand, rivals and supply chains [32].

AI provides virtual experiences to customers who are sitting in the right places, helping them make a final purchase decision. Because AI is a cutting-edge technology that selects the best choice from a range of options provided with a variety of facts through exchange and combinations, it saves customers money [7]. AI-powered augmented reality applications let consumers view things in new ways and facilitate enhanced decision-making [33]. Businesses have integrated most AI-enabled technology to offer clients the best and most customized solutions [34].

Thanks to AI's cutting-edge technologies, customers can easily understand their purchase preferences. Previous research [30, 35, 36] indicates that AI aims to develop software with human-like problem-solving capabilities that enhances ability to make decisions about purchase intention. Studies show that people who visit websites with the integration of AI feel more confident when making judgments about what to buy, which lowers the risk [37].AI is a lightweight technology that helps consumers make informed purchasing decisions. Because consumers are more interested in AI's promise and capabilities, they are using it widely [30]. AI's ability to manage the massive amount of relevant, high-quality data that consumers may access, and that is tied to their purchase activities, determines both its usefulness and efficacy [38, 39].

Because AI is an advanced technology, customers often discover the best virtual experience when they make purchases from online retailers. Customers' virtual experience is important whenever it concerns their purchasing intentions, and studies have indicated that positive virtual experiences affect consumers' intentions to purchase [33].

1.2. Artificial intelligence and social media engagement

Engaging customers and building customer loyalty is critical for providers who value face-to-face communication with consumers. Few studies have examined customer engagement from a technological perspective, even though several social and technological factors have been demonstrated to support customer engagement [40]. Research [40, 41] suggests that AI can enhance moral consumer behavior. Businesses that use AI may change the social media buying experience to give shoppers a social media platform experience, as AI has greatly changed consumer behavior [42].

By using AI to forecast consumer behavior and interact with customers on social media platforms, these businesses may increase the effectiveness of their online marketing campaigns and use it to make more analytical judgments [2, 17]. For instance, AI closes the gap between companies and customers by gathering and evaluating data about goods and services [43]. This changes the online buying experience. AI also offers solutions for a range of problems pertaining to social networking. For instance, evaluating the vast volume of data produced by social media platforms may cause stress for sales staff [2].

To address these issues, businesses utilizing AI may employ a range of AI-based methods for predictive analysis in marketing [9]. The use of AI in social media platforms is one of the outside variables that is believed to motivate consumers to engage with these platforms more. For instance, users can join any community of interest on e-commerce platforms like Facebook, Taobao and Etsy. From there, they can engage with other users, follow other buyers and sellers who share their interests, look up information about products and/or share their own related buying experiences [12]. According to [27], there is a connection across AI and dynamic marketing when real-time technologies are utilized to build personalized, response-focused relationships between buyers and sellers. If businesses provide several options for evaluating the qualities of their products or services, integrating AI increases the likelihood that customers will engage on social media platforms [44]. Thus, it can be assumed that:

H1: AI has a positive effect on social media engagement.

1.3. Artificial intelligence and conversion rate optimization

AI has been used in purchasing procedures to give customers more dependable, individualized services [10]. Wang and Lei [18] claim that artificial intelligence AI technology can manage interactions between consumers and goods or services as well as quickly responding to client demands in interactive environments. As part of AI marketing efforts, chatbots, content features and buyer sales recognition are artificially becoming autonomous [8].

Online social networking platforms give businesses the ability to interact with a wide variety of customers and customize their products to meet their needs [45]. To improve digital marketers' ability to use AI to raise visitor conversion rates, it is crucial to study the buying habits of customers on social networking sites [9]. Businesses on social networking platforms also use AI to entice consumers and win them over as devoted customers [2]. In order to elevate the rate of conversion (from user to customer), businesses can also track consumer behavior on social media and use that data to craft customized promotional campaigns for each individual customer [9, 16]. AI encompasses more than just conversation, language comprehension and object identification; it also includes consumer recommendation and research. This increases the rate at which products are purchased by improving the suggestion of related and complementary products in a more customized and real-time manner [10].

The relationship between the client and the business may be strengthened by AI-based social media initiatives that increase consumer involvement, feedback, and conversion [15]. AI could encourage people to buy goods and services by improving their interaction with social media adverts [2]. Because AI on social networking platforms allows users to examine items or services through the platform, it also encourages prospective consumers to buy a certain product or service. Companies can utilize AI to differentiate their goods or services from competitors' offerings and entice consumers to purchase them. The [46] have reported that prior research has furnished empirical proof of the affirmative correlation between social networking sites and consumer conversion rate. Thus, it can be assumed that:

H2: AI has a positive effect on conversion rate optimization.

1.4. Social media engagement and brand experience

According to [47], there are two types of consumer engagement: uncontrolled (word-of-mouth) and controlled (corporate-sponsored). By sharing knowledge with others, such as through sharing across online platforms, consumers can contribute to the improvement of customer experiences with brands [14]. Satisfactory brand buying experiences can be facilitated by customer engagement [48, 49]. According to research by [20], mobile applications for customer interaction have a favorable impact on customer equity and increase the likelihood that current consumers will make another purchase. A variety of studies [50–52] have also investigated the connection between brand experiences and customer engagement, concluding that there is a substantial impact from consumer involvement. Based on the explanation provided by [51] regarding how customer engagement functions as a mediator to enhance the brand experience and encourage repeat purchases, thus, it can be assumed that:

H3: Social media engagement has a positive effect on brand experience.

1.5. Conversion rate optimization and brand experience

Interaction with customers is important to every business. The shopping experience for customers in a virtual environment is mediated by technology. With the introduction of augmented reality, mixed reality and virtual reality technology, a new environment integrating virtual and physical elements at various levels has emerged. The customer experience environment is changing into new kinds of hybrid experiences as a result of the growth of mobile and wearable devices as well as highly dynamic physical-virtual interactions [9, 53]. Since shoppers of these businesses are more likely to express their favorable experiences with the brand in question, marketers must engage customers and offer a unique social media experience. This is because the most satisfied customers are those who are more involved on social media [2, 11].

72% of businesses prioritize improving the customer experience and appealing to customers throughout the buying process is a marketing trend. Businesses are concentrating on offering value-added ideas to create the greatest possible customer experiences in the digital age [53].

The consumer experience is being drastically changed by emerging technologies including the Internet of Things (IoT), chatbots, bots, augmented reality (AR), virtual reality (VR), mixed reality (MR) and virtual assistants, which are usually powered by AI. Concerns about privacy for clients who would rather buy goods and services online and through social media are crucial when trying to find a consistent way to incorporate the client experience. Instead, marketers must comprehend how digital technologies affect the customer experiences [54].

The cognitive component ingrained in the customer's relation with the brand is satisfaction. Positive remarks affect other users' cognitive processes [55]. Key clients may be drawn to interactive involvement and end up giving products or services favorable reviews [56]. Different client categories will likely require different approaches to customer interaction; after all, in the digital age, consumer engagement is essential [57]. For marketers to comprehend customer segmentation, they must create strong social media analytics. These analyses' findings show how marketers may use social media platforms to sway consumers and raise conversion rates, which in turn have a big impact on customer happiness. For businesses to increase sales, it is essential to comprehend users' attitudes regarding digital media [58]. Thus, it can be assumed that:

H4: Conversion rate optimization has a positive effect on brand experience.

1.6. Brand experience and brand preference

According to [50, 59], brand experience is defined as "the consumer's subjective responses (sensations, feelings and perceptions) and behavioral responses elicited by brand-related stimuli that are part of the brand's design, identity, packaging, communications and surrounding environment." Four categories can be used to categorize brand experience: sensory, intellectual, emotional and behavioral. The stimulation that a brand provides through the senses of sight, sound, smell, taste and touch is known as the sensory brand experience [60]. The emotions evoked by a brand are known as the emotional brand experience. According to [61], behavioral brand experience encompasses actual experiences, behaviors and brand interactions, whereas intellectual brand experiences refer to a company's capacity to elicit thought from consumers.

Perceptions of brand qualities by consumers influence their preferences, which in turn affect their intent and brand selections. As a result, according to [62], brand preference is a pattern of behavior that represents customers' views about the brand. Customers like a specific brand whenever they have positive thoughts regarding it, and their perceptions of a brand's features impact their preferences, which in turn determine their intentions and selection of brands [20].

Brand experience has an impact on brand loyalty and affection, according to [59].

Positive brand experiences help customers form strong bonds with brands and grow to love them [63]. Brand experience has an impact on brand preferences, according to [20]. The notion of brand experience was validated by [64], who found that brand experience is a key indicator of brand preference. It has been noted that brand preference and memorable brand experiences are related. A memorable brand experience positively influences brand preference, which then positively impacts usage intentions, word-of-mouth and readiness to pay more, according to the findings of a study conducted by [65]. Therefore, it can be assumed that:

H5: Brand experience has a positive effect on brand preference.

1.7. Brand preference and re-purchase intention

When comparing a company's products to those of other companies, consumers' preferences toward certain products determine their brand preference. In terms of capturing customers' hearts so they will re-purchase the company's brand, it can be said that this preference for a brand is the first phase of branding [66].

Re-purchase intention is the consumer's plan to carry out the behavioral act of purchasing a brand again [67, 68]. It is the process by which clients choose to re-purchase services or goods from the same company [69]. This probably happens because customers can buy the same thing again. Re-purchase intentions, according to [70], refers to a customer's willingness to make additional purchases from the same merchant or supplier, whereas re-purchases are, in theory, actual actions. According to Sullivan and Kim [71], reactionary intention to re-purchase can be understood as a consumer's wish to reevaluate the brand in light of their present circumstances. The intention to re-purchase is of particular relevance to marketers since it may result from the influence of prior purchases. Re-purchase intention is likely to be lower if consumers' perceptions of price, experience, brand and fulfillment differ from what they paid and received [70].

Customers are more likely to repeat purchases when they have a preferred brand. Only when consumers feel positive about a brand will they choose to re-purchase it and replicate their experience [23]. Additionally, research indicates that customers' decisions to buy a product are influenced by their information processing, which is reflected in their choice of a brand [70]. According to [62], re-purchase intention was positively impacted by brand preference. The [66] claim that a product's identity as a brand and preference are responsible for its resurgence in popularity. Research from Ho & Chow [20] has shown that brand preference affects consumers' likelihood to make more purchases. According to [64], brand preference and re-purchase intentions were positively correlated statistically. The [65] investigation confirmed that brand preference affects re-purchase intentions. Thus, it can be assumed that:

H6: Brand preference has a positive effect on repurchase intentions.

1.8. Brand experience and re-purchase intentions

Consumer brand experience precedes actual purchase because favorable brand experiences have a positive and significant impact on consumer purchase intentions, and prior experiences become memorable during brand purchase [72, 73]. The positive feelings that consumers have for a brand can impact their intention to make a purchase if they are feeling good about it [74]. This suggests that consumers' behavioral intentions may grow as a result of their brand experience. According to [62], a favorable brand experience can affect the propensity to re-purchase. According to [70], the re-purchase intention is positively impacted by brand experience. Therefore, it can be assumed that:

H7: Brand experience has a positive effect on repurchase intentions.

2. Methodology

2.1. Procedures and respondents

Data was gathered from Jordanian users of the internet who have at least once made a purchase or used services available online from any website relevant to aviation. An additional eight weeks were added to the data gathering period.

The data was gathered using convenience sampling, which is a non-random sampling technique [75]. Based on their actual usage of the web services for the websites of the aviation companies, the study's respondents were selected. A non-probability sampling design was used in this investigation, meaning that there are no odds associated with any member of the study population being selected as a sample subject [76]. The questionnaire was created in English, therefore with the assistance of two multilingual specialists, we translated it first into Arabic before translating it back into English. The individuals who participated were informed that they might opt out of the study at any moment and that participation in it was entirely optional. Pens were used by the participants to rate the questions. The participants' answers to the surveys were gathered directly. A survey was disregarded and the next one was chosen if it was not completed correctly. Surveys filled out by participants who had no prior e-commerce experience were disqualified. An amount of 800 questionnaires together with cover letters were given out, and 387 respondents brought the completed forms back. Thirty-two questionnaires were eliminated due to incomplete information. In the end, 355 replies were considered for study. 44.4% of respondents responded.

2.2. Measures

A five-point Likert scale, with 1 denoting "strongly disagree" and 5 denoting "strongly agree," was used to rate each scale item.

Artificial Intelligence: [2, 45] produced a 6-item scale that we used to gauge technological advancements in AI. "Multiple types of data about customers, such as sales, purchases, or demographic and behavioral data," was the sample item.

Social Media Engagement: An instrument consisting of five items was created by [2, 14] to gauge consumer participation on social media. "Using social networking websites sparks my curiosity about brands" was the sample item.

Conversion Rate Optimization: Utilizing a 4-item scale created by [2, 46], conversion rate optimization was examined. "I am influenced to buy products and services by web-based promotions and messages on social networking sites" was the sample item.

Brand Experience: We used a 5-item measure that was created by [20, 24]. The sample item was, "the experience of using AI".

Brand Preference: A 5-item scale created by [20, 77] was used to measure brand preference. The sample item was, "preferred brand over any other brand."

Re-purchase Intention: We used a 4-item measure that was created by [2, 20]. To gauge technological advancements in AI. To gauge the re-purchase intentions of the consumer. "I plan to keep using the site that I frequently use for booking flights" was the sample item.

2.3. Reliability and validity

Factor analysis, average variance extracted (AVE) and composite reliability (CR) are computed using AMOS. One method for condensing a large range of variables into a smaller number of factors is a confirmatory factor analysis. Using this method, all variables' highest common variance is extracted, and the results are combined into a single score. Confirmatory factor analysis also known as a CFA, was carried out to examine the multiple-item measures' discriminant validity, convergent validity and reliability. According to [78] recommendation, the analysis's findings supported each measuring scale's convergent validity. The theoretical constructs appear to have convergent validity, as indicated by the statistically significant ($p < 0.05$) factor loadings of 0.60 to 0.90 for all indicators in their respective constructs, as presented in *Table 1* [79]. Furthermore, every construct's average variance extracted (AVE) is greater than the minimal value of 0.5 that is advised [80]. The average variance extracted (AVE) results were used to evaluate the discriminant validity. According to the results of *Table 1*, the square roots of AVE are greater than correlations, which suggests that the discriminant validity is satisfactory [80]. The [81] suggested using composite reliability (CR) to assess the dependability of the measures. *Table 1* shows the discriminant and convergent validity as well as the reliability across all reflection measures based on CR values that are over the 0.70 threshold.

Cronbach's alpha is calculated using the average correlations between the concept-measuring items. The internal consistency reliability increases with Cronbach's alpha's proximity to 1 [76]. The Cronbach's coefficient is employed to assess the reliability of each concept. It is a metric for assessing a multi-item scale's internal consistency. According to the SPSS results, all alpha coefficient values are higher than 0.7, indicating that the measuring scales' reliability is sufficient [82].

3. Results

3.1. Descriptive statistics

The mean, standard deviations, and correlation matrix are the primary descriptive statistics that were utilized to characterize the study constructs. The study model constructs' mean scores ranged from 3.46 to 3.80, as shown in *Table 2.* Furthermore, the correlations showed that the research variables had a strong link and ranged from 3.46 to 3.80.

3.2. Research model and hypotheses

The links between the constructs were estimated by the application of structural equation modeling (SEM). Amos V.22 was used to calculate SEM estimates. Regarding the proposed connections, the path

Table 1.

Confirmatory factor analysis & reliabilities

Construct	Items	Loading factor	*t*-value	CR	AVE	α
Artificial Intelligence	AI1	0.719		0.881	0.557	0.898
	AI2	0.868	15.523			
	AI3	0.892	15.833			
	AI4	0.668	11.994			
	AI5	0.633	11.345			
	AI6	0.654	11.739			
Social Media Engagement	SME1	0.777		0.850	0.532	0.858
	SME 2	0.803	14.679			
	SME 3	0.716	13.066			
	SME 4	0.63	11.32			
	SME 5	0.71	13.013			
Conversion Rate Optimization	CRO1	0.813		0.856	0.598	0.855
	CRO2	0.82	16.009			
	CRO3	0.724	13.986			
	CRO4	0.732	14.171			
Brand Experience	BE1	0.603		0.861	0.557	0.852
	BE2	0.735	10.932			
	BE3	0.876	12.181			
	BE 4	0.822	11.766			
	BE 5	0.663	10.159			
Brand Preference	BR1	0.814		0.897	0.635	0.897
	BR2	0.726	14.915			
	BR3	0.816	17.39			
	BR4	0.792	14.723			
	BR5	0.833	17.783			
Re-purchase Intention	RI1	0.773		0.827	0.546	0.930
	RI2	0.788	15.291			
	RI3	0.698	13.308			
	RI4	0.691	13.145			

Table 2.

Means, standard deviations, and correlations for the study variables

Study variables	Mean	Std. dev	1	2	3	4	5	6
Artificial Intelligence	3.80	0.665	**0.746**					
Social Media Engagement	3.53	0.715	0.315*	**0.729**				
Conversion Rate Optimization	3.46	0.833	0.225*	0.313*	**0.773**			
Brand Experience	3.83	0.687	0.282*	0.307*	0.366*	**0.746**		
Brand Preference	3.56	0.709	0.282*	0.411*	0.433*	0.611*	**0.796**	
Re-purchase Intention	3.65	0.657	0.335*	0.392*	0.333*	0.620*	0.754*	**0.740**

Notes: *$p < 0.01$; square root of AVE is on the diagonal

that leads from AI to Social Media Engagement has a coefficient of 0.449 ($p < 0.01$) regarding the linkages. Therefore, the positive correlation implies that H1 is validated. Furthermore, the results confirm hypothesis H2 by demonstrating that the relationship among AI and Conversion Rate Optimization ($\beta = 0.305$, $p > 0.01$) follows the expected direction. Furthermore, the findings demonstrate that Brand Experience is positively and significantly impacted by Social Media Engagement ($\beta = 0.317$, $t = 5.519$, $p < 0.01$) and positively and significantly impacted by Conversion Rate Optimization ($\beta = 0.346$, $t = 5.592$, $p < 0.01$). As a result, theories H3 and H4 are validated. Additionally, according to the findings, Brand Experience significantly and favorably influences Brand Preference ($\beta = 0.684$, $t = 9.419$, $p < 0.01$). Re-purchase Intention is favorably correlated with both Brand Preference and Brand Experience, supporting hypotheses H6 and H7.

3.3. Mediating test

5000 bootstrap samples were chosen, with a 95% confidence level. According to the study model, there are four ways that indirect impacts can manifest.

- H8 AI → Social Media Engagement → Brand Experience → e-purchase Intention

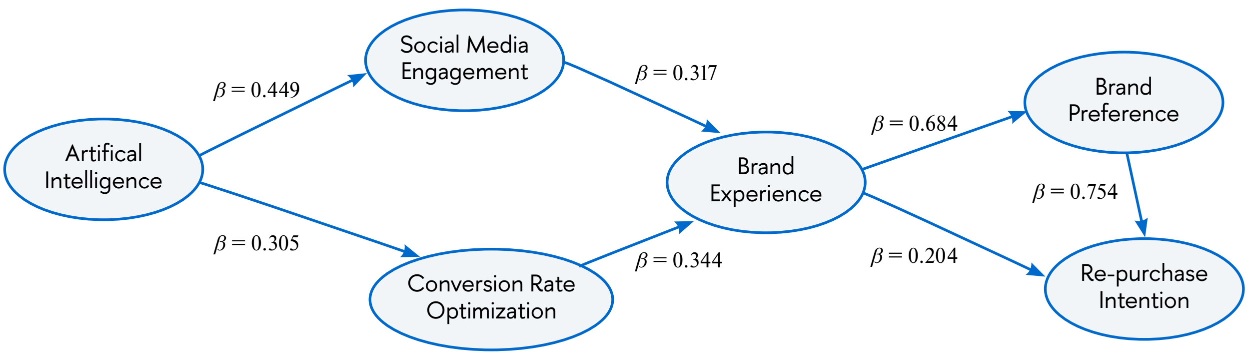

Fig. 1. Structural model with parameter estimates [2, 9, 51].

Table 3.

Path analysis for the constructs of the study

	Relation		Coefficients	*t*-value*	Support/ nonsupport
Path	Artificial Intelligence	Social Media Engagement	0.449	7.108*	Support
	Artificial Intelligence	Conversion Rate Optimization	0.305	4.998*	Support
	Social Media Engagement	Brand Experience	0.317	5.159*	Support
	Conversion Rate Optimization	Brand Experience	0.346	5.592*	Support
	Brand Experience	Brand Preference	0.684	9.419*	Support
	Brand Experience	Re-purchase Intention	0.204	3.405*	Support
	Brand Preference	Re-purchase Intention	0.754	10.597*	Support
Explained variance proportion R^2 of Conversion Rate Optimization				0.093	
Explained variance proportion R^2 of Social Media Engagement				0.202	
Explained variance proportion R^2 of Brand Experience				0.251	
Explained variance proportion R^2 of Brand Preference				0.467	
Explained variance proportion $R^2$2 of Re-purchase Intention				0.82	

- H9 AI → ocial Media Engagement → Brand Experience → Brand Preference → Re-purchase Intention
- H10 AI → Conversion Rate Optimization → Brand Experience → Re-purchase Intention
- H11 AI → Conversion Rate Optimization → Brand Experience → Brand Preference → Re-purchase Intention

The product of the route coefficients between AI and Re-purchase Intention was used to determine the indirect effect of AI on Re-purchase Intentions. Significant indirect effects of AI on Re-purchase Intentions were discovered from the study model for the four paths. To be more precise, there is an indirect effect by means of Conversion Rate Optimization and Brand Experience (β = 0.020, p < 0.01), Social Media Engagement and Brand Experience (β = 0.027, p < 0.01), and Social Media Engagement, Brand Experience and Brand Preference (β = 0.068, p < 0.01). Lastly, there is an indirect effect through Conversion Rate Optimization, Brand Experience and Brand Preference (β = 0.050, p < 0.01). Therefore, the influence of AI on Re-purchase Intentions was mediated by Social Media Engagement, Brand Experience, Brand Preference and Conversion Rate Optimization.

Bootstrap approaches were used to evaluate the indirect effect of AI on Re-purchase Intention.

4. Discussion

The [45] reported that the results showed how AI technology affects social media participation. This implies that in order for businesses to remain competitive in the present business environment, they have used social media campaigns to transform their offline operations into online ones and generate website traffic that eventually converts into actual customers. Businesses that use AI technology have a good correlation with social media user engagement [83]. This result supports

Table 4.

Indirect effects of SCC on NPP through KS and IC

Indirect effect	β	95% Bootstrap CI		p
		Lower limit	Upper limit	
AI → SME → BE → RI	0.027	0.011	0.069	0.003
AI → SME → BE → BR → RI	0.068	0.032	0.132	0.001
AI → CRO → BE → RI	0.020	0.008	0.047	0.003
AI → CRO → BE → BR → RI	0.050	0.020	0.105	0.001

(AI) Artificial Intelligence; (SME) Social Media Engagement; (CRO) Conversion Rate Optimization; (BR) Brand Experience; (RI) Re-purchase Intention.

the hypothesis that social media integration of AI could enable marketers to interact with prospective clients to promote the products and services they offer [84]. The study discovered an effect on businesses' adoption of AI to raise conversion rates. This lends credence to the idea that social networking sites might boost a business's amount of sales. Social media marketing powered by AI may increase consumer feedback and engagement as well as the customer-business relationship's conversion rate. AI improves user engagement with social media marketing, which encourages users to buy goods or services. This outcome is in line with the findings of [2, 85].

Customer conversion rates have been shown to affect the brand experience. Businesses may better understand client segmentation and increase conversion rates on online platforms with the help of social media analytics. Social media marketers can tailor their commercial practices and professional activities by leveraging AI technologies to learn customer attitudes. To provide a positive experience and increase sales volume; this outcome is in line with [2, 85].

According to [70], this study demonstrated the relationship between social media engagement and brand experience, indicating that consumers who are more active in social networks are more likely to interact with the brand. Our results align with earlier studies that have demonstrated positive consumer evaluations of goods and services on social networking platforms are given by those who are happy with their purchases [2].

Consistent with earlier research [20, 62, 70], the results show that brand experience influences brand choice and that a favorable brand experience can boost consumer-based brand preference. As a result, customers' brand preference may be increased by a favorable brand experience. Consumers mostly base their brand choice on their experiences. Customers are more inclined to like a brand when they have had numerous positive interactions with it. According to an earlier study [62], consumers will exhibit positive behavioral intentions when they perceive a high degree of brand experience. This highlights the significance of an unforgettable brand experience in the context of consumer behavior.

As previously said in the literature, brand experience aids in encapsulating a brand's behavioral, emotional, social, pragmatic, sensory, intellectual and lifestyle elements [70]. The consumer will develop preferences and make judgments about what to buy through this interactive experience [74, 86].

The findings indicate that the re-purchase intention is significantly influenced by brand preference and brand experience. Because consumer preferences and brand experiences are sustainable ideas that represent unreasonable elements related to the customer who engages with the brand and goes over the limits of rational assumptions, this means that if customers like the product and have a stimulating interaction with it, they are more likely to intend to re-purchase it. Customers will thus have a strong desire to buy without using reason [70].

The findings showed that the impact of AI on the intention to re-purchase is mediated by brand experience. The findings further indicate that the impact of AI on intention to re-purchase was mediated by brand preference. This is understandable given that AI offers a novel consumer experience that increases brand preference, customer satisfaction and product re-purchases [20].

The findings showed that the impact of AI on brand experience and propensity to re-purchase is mediated by social media engagement and conversion rate optimization. The majority of the research on social networking sites engagement has been on online brand communities and social media [51]. This demonstrates how crucial consumer engagement with AI technology is to brand marketing. Using AI technology for marketing brands necessitates higher social media engagement since it makes it easier for brand marketers to convey their experiences to consumers, which leads to the formation of positive brand experiences and repurchase intentions [51]. Using AI, marketers can also raise visitor conversion rates [9]. Organizations on social networking sites also employ AI as a technique to entice consumers and win them over as devoted customers [2].

5. Theoretical implications

According to this study, businesses may assess the relative value of each element of their products and services and how it affects customer satisfaction on social media platforms by utilizing AI-enabled solutions. The rising use of social media platforms has resulted in a huge increase in consumer interaction, suggesting that social media sites are becoming a new marketplace for establishing relationships with customers to sell products and services. According to [45], AI tracks and analyzes user habits on social media platforms. Technology is playing a bigger role in customer engagement. To improve consumers' repurchase intention, AI and consumer behavior should be taken into account while implementing a customer engagement plan.

According to this study, AI can improve social media platforms' capacity to attract new clients for Airline companies. Acknowledge the importance of AI, which efficiently manages data processing for specialized services through automation. Studies on customer contact on social media platforms and AI-powered automated business responses are still in their early stages. However, there is a gap in the way companies use real-time data to offer personalized customer service when interacting with clients [2].

Because AI can successfully generate brand preference and purchasing commitment, this study validates AI's overall effectiveness. As a result, by offering a thorough framework that clarifies the connections between AI and branding, our research adds to the body of knowledge on marketing and branding. This study's discussions will be beneficial to marketing academics who wish to apply this approach to other domains.

The study validates the impact of AI marketing methods on brand experience and preference, as well as the correlation among brand preference and re-purchase intention. A few studies have been conducted on AI preference for brands. By elucidating the function of AI in customer interactions within the setting of airline services, this study seeks to close this knowledge gap. By showing the predictive power of AI marketing techniques on brand preference, this work advances the field of service research and offers researchers with an interest in deploying AI to customer decision-making and behavior valuable information. Furthermore, the partial mediation arrangement of brand experience between AI marketing tactics and brand preference shows how these strategies predict brand preference and repurchase intentions in airlines both directly and indirectly via brand encounters [20].

6. Practical implications

There are various implications of this study for academics and professionals. First, social media platforms are gradually being used by customers in Jordan. These days, the majority of consumers would rather buy goods and services via the Internet and on social networking platforms than leave the comfort of their homes. Additionally, social media and online channels let companies boost sales. Companies may easily keep an eye on what their clients are doing on social media [10]. As a result, they implement efficient communication strategies that aid in raising the conversion rate. According to [2], organizations must deal directly with their clients to ascertain their requirements and expectations. Although it is essential to businesses, customer involvement on social media is insufficient to help them. Managers will benefit from the current research's understanding of the technological context and its effects on behavior and society. Third, incorporating social media platforms facilitates the analysis of client feedback and the conversion of those responses into actual sales [1]. Social media networks with AI capabilities can assist executives in forecasting customer behavior patterns within the aviation sector. The suggested structure enables managers to influence consumers on social media platforms, hence increasing sales capacity. This report encourages managers to increase social networking conversion rates by utilizing the newest digital technology.

Fourth, companies may boost sales and develop a computerized digital system that assesses and analyzes the social media user experience by integrating AI into their social media marketing campaigns. Social media sites are a useful tool for marketing goods and services internationally. The study's findings also demonstrate that after returning customers get used to online buying, they alter their decision-making processes [2]. Therefore, by providing vouchers for savings and promotions and cultivating customer re-purchase behaviors, airlines may entice frequent travelers.

Fifth, this study discovered that the use of AI in marketing significantly improved brand preference, which subsequently turn affected consumers' desire to make additional purchases. According to these findings, AI marketing initiatives should be seen as a vital instrument for enhancing the brand image in addition to being a means of improving the consumer experience [20]. To improve long-term commercial performance and brand attractiveness, the airline should focus brand-building efforts on AI. Thus, the airline ought to make greater investments in AI and booking service technology, both to draw in new business and to strengthen existing ones.

An airline that is reluctant to use AI may need to reevaluate its investment strategies because the first-mover advantage is still quite significant. Since this study shows how consumers appreciate AI activities after realizing their values, verification of AI campaigns is a crucial sign of the return on investment. Therefore, managers must make sure AI can provide accurate, dependable and efficient airline-related services. Managers can use AI to send clients tailored marketing communications about services and goods at the right time. In order to meet customer requests, AI assistants and agents should be designed with the capacity to provide knowledgeable customer support and guidance. Airlines' practitioners can also consider improving the AI interface.

7. Limitations and future research

This study has several limitations.

First, the 355 valid samples were obtained by online questionnaires from individuals who had made travel reservations through websites, suggesting that our knowledge of AI brand interactions may be restricted.

Secondly, a cross-sectional approach was employed to gather data from the participants. A longitudinal study might be pertinent to evaluate the suggested model to investigate consumers' intention to re-purchase because habits are amassed over time. To increase validity, future research might employ a bigger sample size.

Third, the study's focus is on the airline business. The results could apply to other businesses or philosophies, even if they are probably most helpful in the context of airlines. This study could be repeated in the future and expanded to include different sectors or nations. This study is quantitative in style; future research may use mixed or qualitative methodologies.

Lastly, producing a response rate that is higher than 44.4%.

To generalize these findings, researchers can also analyze consumer behavior across national borders and industry sectors through cross-cultural studies.

Conclusion

The study explores the mediating role of social media engagement, conversion rate optimization, brand experience and brand preference in the relationship between AI and re-purchase intentions. It also empirically evaluates a model for the implementation of AI in re-purchase intentions and its role in improving conversion rate optimization. The study discovered that through social media interaction, brand experience, brand preference and conversion rate optimization, AI significantly influences re-purchase intentions indirectly. In a similar vein, it has been discovered that social media interaction and conversion rate optimization significantly affect brand experience. Additionally, the re-purchase intention is significantly impacted by brand experience. ■

References

1. Dwivedi Y.K., Wang Y. (2022) Guest editorial: Artificial intelligence for B2B marketing: Challenges and opportunities. *Industrial Marketing Management*, vol. 105, pp. 109–113. https://doi.org/10.1016/j.indmarman.2022.06.001
2. Nazir S., Khadim S., Asadullah M.A., Syed N. (2023) Exploring the influence of artificial intelligence technology on consumer repurchase intention: The mediation and moderation approach. *Technology in Society*, vol. 72, 102190. https://doi.org/10.1016/j.techsoc.2022.102190
3. Collins C., Dennehy D., Conboy K., Mikalef P. (2021) Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, vol. 60, 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383
4. Chen P., Kim S. (2023) The impact of digital transformation on innovation performance – The mediating role of innovation factors. *Heliyon*, vol. 9, no. 3, e13916. https://doi.org/10.1016/j.heliyon.2023.e13916
5. Vishnoi S.K., Bagga T.E.E.N.A., Sharma A.A.R.U.S.H.I., Wani S.N. (2018) Artificial intelligence enabled marketing solutions: A review. *Indian Journal of Economics and Business*, vol. 17, no. 4, pp. 167–177.
6. Fan J., Fang L., Wu J., Guo Y., Dai Q. (2020) From brain science to artificial intelligence. *Engineering*, vol. 6, no. 3, pp. 248–252. https://doi.org/10.1016/j.eng.2019.11.012
7. Uzir M.U.H., Bukari Z., Al Halbusi H., et al. (2023) Applied artificial intelligence: Acceptance-intention-purchase and satisfaction on smartwatch usage in a Ghanaian context. *Heliyon*, vol. 9, no. 8, e18666. https://doi.org/10.1016/j.heliyon.2023.e18666
8. Overgoor G., Chica M., Rand W., Weishampel A. (2019) Letting the computers take over: Using AI to solve marketing problems. *California Management Review*, vol. 61, no. 4, pp. 156–185. https://doi.org/10.1177/0008125619859318
9. Bag S., Srivastava G., Bashir M.M.A., Kumari S., Giannakis M., Chowdhury A.H. (2022) Journey of customers in this digital era: Understanding the role of artificial intelligence technologies in user engagement and conversion. *Benchmarking: An International Journal*, vol. 29, no. 7, pp. 2074– 2098. https://doi.org/10.1108/BIJ-07-2021-0415
10. Yin J., Qiu X. (2021) AI technology and online purchase intention: Structural equation model based on perceived value. *Sustainability*, vol. 13, no. 10, 5671. https://doi.org/10.3390/su13105671
11. Majeed M., Asare C., Fatawu A., Abubakari A. (2022) An analysis of the effects of customer satisfaction and engagement on social media on repurchase intention in the hospitality industry. *Cogent Business & Management*, vol. 9, no. 1, 2028331. https://doi.org/10.1080/23311975.2022.2028331
12. Busalim A., Hollebeek L.D., Lynn T. (2023) The effect of social commerce attributes on customer engagement: an empirical investigation. *Internet Research*, vol. 34, no. 7, pp. 187–214. https://doi.org/10.1108/INTR-03-2022-0165
13. Thakur R. (2016) Understanding customer engagement and loyalty: a case of mobile devices for shopping. *Journal of Retailing and Consumer Services*, vol. 32, pp. 151–163. https://doi.org/10.1016/j.jretconser.2016.06.004
14. Hollebeek L.D., Glynn M.S., Brodie R.J. (2014) Consumer brand engagement in social media: Conceptualization, scale development and validation. *Journal of Interactive Marketing*, vol. 28, no. 2, pp. 149–165. https://doi.org/10.1016/j.intmar.2013.12.002

15. Dwivedi Y.K., Ismagilova E., Hughes D.L., et al. (2021) Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, vol. 59, 102168. https://doi.org/10.1016/j.ijinfomgt.2020.102168

16. Al-Natour S., Turetken O. (2020) A comparative assessment of sentiment analysis and star ratings for consumer reviews. *International Journal of Information Management*, vol. 54, 102132. https://doi.org/10.1016/j.ijinfomgt.2020.102132

17. Duan Y., Edwards J.S., Dwivedi Y.K. (2019) Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management*, vol. 48, pp. 63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021

18. Wang X., Lei S. (2018) Research on the impact of artificial intelligence on consumption and shopping experience in the new retail environment – Based on the perspective of commercial retail reform and the reconstruction of the human-goods-scene system. *Journal of Commercial Economics*, vol. 37, no. 17, pp. 5–8.

19. Murphy J., Gretzel U., Pesonen J. (2019) Marketing robot services in hospitality and tourism: The role of anthropomorphism. *Journal of Travel and Tourism Marketing*, vol. 36, no. 7, pp. 784–795. https://doi.org/10.1080/10548408.2019.1571983

20. Ho S.P.S., Chow M.Y.C. (2024) The role of artificial intelligence in consumers' brand preference for retail banks in Hong Kong. *Journal of Financial Services Marketing*, vol. 29, pp. 292–305. https://doi.org/10.1057/s41264-022-00207-3

21. Stylos N., Zwiegelaar J., Buhalis D. (2021) Big data empowered agility for dynamic, volatile, and time-sensitive service industries: the case of tourism sector. *International Journal of Contemporary Hospitality Management*, vol. 33, no. 3, pp. 1015–1036. https://doi.org/10.1108/IJCHM-07-2020-0644

22. Chan A.P.H., Tung V.W.S. (2019) Examining the effects of robotic service on brand experience: the moderating role of hotel segment. *Journal of Travel and Tourism Marketing*, vol. 36, no. 4, pp. 458–468. https://doi.org/10.1080/10548408.2019.1568953

23. Kim Y.J., Park J.S., Jeon H.M. (2021) Experiential value, satisfaction, brand love, and brand loyalty toward robot barista coffee shop: The moderating effect of generation. *Sustainability*, vol. 13, no. 21, 12029. https://doi.org/10.3390/su132112029

24. Trivedi J. (2019) Examining the customer experience of using banking chatbots and its impact on brand love: The moderating role of perceived risk. *Journal of Internet Commerce*, vol. 18, no. 1, pp. 91–111. https://doi.org/10.1080/15332861.2019.1567188

25. Kumar V., Rajan B., Venkatesan R., Lecinski J. (2019) Understanding the role of artificial intelligence in personalized engagement marketing. *California Management Review*, vol. 61, no. 4, pp. 135–155. https://doi.org/10.1177/0008125619859317

26. Mishra S., Ewing M.T., Cooper H.B. (2022) Artificial intelligence focus and firm performance. *Journal of the Academy of Marketing Science*, vol. 50, no. 6, pp. 1176–1197. https://doi.org/10.1007/s11747-022-00876-5

27. Peltier J.W., Dahl A.J., Schibrowsky J.A. (2024) Artificial intelligence in interactive marketing: A conceptual framework and research agenda. *Journal of Research in Interactive Marketing*, vol. 18, no. 12, pp. 54–90. https://doi.org/10.1108/JRIM-01-2023-0030

28 Manser Payne E.H., Peltier J., Barger V.A. (2021) Enhancing the value co-creation process: artificial intelligence and mobile banking service platforms. *Journal of Research in Interactive Marketing*, vol. 15, no. 1, pp. 68–85. https://doi.org/10.1108/JRIM-10-2020-0214

29. Vlačić B., Corbo L., e Silva S.C., Dabić M. (2021) The evolving role of artificial intelligence in marketing: A review and research agenda. *Journal of Business Research*, vol. 128, pp. 187–203. https://doi.org/10.1016/j.jbusres.2021.01.055

30. Bhagat R., Chauhan V., Bhagat P. (2023) Investigating the impact of artificial intelligence on consumer's purchase intention in e-retailing. *Foresight*, vol. 25, no. 2, pp. 249–263. https://doi.org/10.1108/FS-10-2021-0218

31. Huang M.H., Rust R.T. (2021) A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, vol. 49, pp. 30–50. https://doi.org/10.1007/s11747-020-00749-9

32. Dekimpe M.G. (2020) Retailing and retailing research in the age of big data analytics. *International Journal of Research in Marketing*, vol. 37, no. 1, pp. 3–14. https://doi.org/10.1016/j.ijresmar.2019.09.001

33. Pantano E., Rese A., Baier D. (2017) Enhancing the online decision-making process by using augmented reality: A two country comparison of youth markets. *Journal of Retailing and Consumer Services*, vol. 38, pp. 81–95. https://doi.org/10.1016/j.jretconser.2017.05.011

34. Reinartz W., Wiegand N., Imschloss M. (2019) The impact of digital transformation on the retailing value chain. *International Journal of Research in Marketing*, vol. 36, no. 3, pp. 350–366. https://doi.org/10.1016/j.ijresmar.2018.12.002

35. Astawa I.G.N.M.W., Sukawati T.G.R.S. (2019) The role of perceived value mediate the effect of utilitarian and hedonic shopping value on intent to online repurchase. *International Journal of Management and Commerce Innovations*, vol. 6, no. 1, pp. 1232–1242.

36. Qian M., Xu Z. (2019) A study of dynamic recognition of consumer brand decision-making preference based on machine learning method. *Nankai Business Review*, vol. 22, no. 1, pp. 66–76.

37. Haenlein M., Kaplan A., Tan C.W., Zhang P. (2019) Artificial intelligence (AI) and management analytics. *Journal of Management Analytics*, vol. 6, no. 4, pp. 341–343. https://doi.org/10.1080/23270012.2019.1699876

38. Sohn K., Kwon O. (2020) Technology acceptance theories and factors influencing artificial Intelligence-based intelligent products. *Telematics and Informatics*, vol. 47, 101324. https://doi.org/10.1016/j.tele.2019.101324

39 Attig C., Franke T. (2020) Abandonment of personal quantification: A review and empirical study investigating reasons for wearable activity tracking attrition. *Computers in Human Behavior*, vol. 102, pp. 223–237. https://doi.org/10.1016/j.chb.2019.08.025

40. Yan Y., Chen H., Shao B., Lei Y. (2023) How IT affordances influence customer engagement in live streaming commerce? A dual-stage analysis of PLS-SEM and fsQCA. *Journal of Retailing and Consumer Services*, vol. 74, 103390. https://doi.org/10.1016/j.jretconser.2023.103390

41. Sun Y., Shao X., Li X., Guo Y., Nie K. (2019) How live streaming influences purchase intentions in social commerce: An IT affordance perspective. *Electronic Commerce Research and Applications*, vol. 37, 100886. 1https://doi.org/10.1016/j.elerap.2019.100886

42. De Oliveira Santini F., Ladeira W.J., Pinto D.C., Herter M.M., Sampaio C.H., Babin B.J. (2020) Customer engagement in social media: a framework and meta-analysis. *Journal of the Academy of Marketing Science*, vol. 48, no. 6, pp. 1211–1228. https://doi.org/10.1007/s11747-020-00731-5

43. Garg P., Gupta B., Dzever S., Sivarajah U., Kumar V. (2020) Examining the relationship between social media analytics practices and business performance in the Indian retail and IT industries: The mediation role of customer engagement. *International Journal of Information Management*, vol. 52, 102069. https://doi.org/10.1016/j.ijinfomgt.2020.102069

44. Busalim A.H., Ghabban F. (2021) Customer engagement behaviour on social commerce platforms: An empirical study. *Technology in Society*, vol. 64, 101437. https://doi.org/10.1016/j.techsoc.2020.101437

45. Capatina A., Kachour M., Lichy J., Micu A., Micu A.E., Codignola F. (2020) Matching the future capabilities of an artificial intelligence-based software for social media marketing with potential users' expectations. *Technological Forecasting and Social Change*, vol. 151, 119794. https://doi.org/10.1016/j.techfore.2019.119794

46. Di Fatta D., Patton D., Viglia *G.* (2018) The determinants of conversion rates in SME e-commerce websites. *Journal of Retailing and Consumer Services*, vol. 41, pp. 161–168. https://doi.org/10.1016/j.jretconser.2017.12.008

47. Roy S.K., Eshghi A., Sarkar A. (2013) Antecedents and consequences of brand love. *Journal of Brand Management*, vol. 20, pp. 325–332. https://doi.org/10.1057/bm.2012.24

48. Chang C.W., Huang H.C., Wang S.J., Lee H. (2021) Relational bonds, customer engagement, and service quality. *The Service Industries Journal*, vol. 41, no. 5–6, pp. 330–354. https://doi.org/10.1080/02642069.2019.1611784

49. Khattab S.A., Al Shaar I.M., Zaid M.K.A., Qutaishat F.T. (2023) The effect of relational bonds on e-commerce use, the mediating effect of customers' online trust: evidence from Jordan. *International Journal of Business and Systems Research*, vol. 17, no. 5, pp. 483–503.

50. Brakus J.J., Schmitt B.H., Zarantonello L. (2009) Brand experience: what is it? How is it measured? Does it affect loyalty? *Journal of Marketing*, vol. 73, no. 3, pp. 52–68. https://doi.org/10.1509/jmkg.73.3.052

51. Hsu C.L. (2023) Enhancing brand love, customer engagement, brand experience, and repurchase intention: Focusing on the role of gamification in mobile apps. *Decision Support Systems*, vol. 174, 114020. https://doi.org/10.1016/j.dss.2023.114020

52. Islam J.U., Hollebeek L.D., Rahman Z., Khan I., Rasool A. (2019) Customer engagement in the service context: An empirical investigation of the construct, its antecedents and consequences. *Journal of Retailing and Consumer Services*, vol. 50, pp. 277–285. https://doi.org/10.1016/j.jretconser.2019.05.018

53. Flavián C., Ibáñez-Sánchez S., Orús C. (2019) The impact of virtual, augmented and mixed reality technologies on the customer experience. *Journal of Business Research*, vol. 100, pp. 547–560. https://doi.org/10.1016/j.jbusres.2018.10.050

54. Hoyer W.D., Kroschke M., Schmitt B., Kraume K., Shankar V. (2020) Transforming the customer experience through new technologies. *Journal of Interactive Marketing*, vol. 51, no. 1, pp. 57–71. https://doi.org/10.1016/j.intmar.2020.04.001

55. Chen Y., Fay S., Wang Q. (2011) The role of marketing in social media: How online consumer reviews evolve. *Journal of Interactive Marketing*, vol. 25, no. 2, pp. 85–94. https://doi.org/10.1016/j.intmar.2011.01.003

56. Choi H., Kandampully J. (2019) The effect of atmosphere on customer engagement in upscale hotels: An application of SOR paradigm. *International Journal of Hospitality Management*, vol. 77, pp. 40–50. https://doi.org/10.1016/j.ijhm.2018.06.012

57. Eigenraam A.W., Eelen J., Van Lin A., Verlegh P.W. (2018) A consumer-based taxonomy of digital customer engagement practices. *Journal of Interactive Marketing*, vol. 44, no. 1, pp. 102–121. https://doi.org/10.1016/j.intmar.2018.07.002

58. He W., Wu H., Yan G., Akula V., Shen J. (2015) A novel social media competitive analytics framework with sentiment benchmarks. *Information & Management*, vol. 52, no. 7, pp. 801–812. https://doi.org/10.1016/j.im.2015.04.006

59. Bae B.R., Kim S.-E. (2023) Effect of brand experiences on brand loyalty mediated by brand love: the moderated mediation role of brand trust. *Asia Pacific Journal of Marketing and Logistics*, vol. 35, no. 10, pp. 2412–2430. https://doi.org/10.1108/apjml-03-2022-0203

60. Andreini D., Pedeliento G., Zarantonello L., Solerio C. (2019) A renaissance of brand experience: Advancing the concept through a multi-perspective analysis. *Journal of Business Research*, vol. 96, pp. 355–365. https://doi.org/10.1016/j.jbusres.2018.05.046

61. Bapat D., Thanigan J. (2016) Exploring relationship among brand experience dimensions, brand evaluation and brand loyalty. *Global Business Review*, vol. 17, no. 6, pp. 1357–1372. https://doi.org/10.1177/0972150916660401

62. Ebrahim R., Ghoneim A., Irani Z., Fan Y. (2016) A brand preference and repurchase intention model: the role of consumer experience. *Journal of Marketing Management*, vol. 32, no. 13–14, pp. 1230–1259. https://doi.org/10.1080/0267257X.2016.1150322

63. Singh D., Bajpai N., Kulshreshtha K. (2021) Brand experience – brand love relationship for Indian hypermarket brands: The moderating role of customer personality traits. *Journal of Relationship Marketing*, vol. 20, no. 1, pp. 20–41. https://doi.org/10.1080/15332667.2020.1715179

64. Hwang J., Choe J.Y.J., Kim H.M., Kim J.J. (2021) The antecedents and consequences of memorable brand experience: Human baristas versus robot baristas. *Journal of Hospitality and Tourism Management*, vol. 48, pp. 561–571. https://doi.org/10.1016/j.jhtm.2021.08.013

65. Hwang J., Kim H., Kim H.M. (2023) Relationships among memorable brand experience, brand preference, and behavioral intentions: focusing on the difference between robot servers and human servers. *Journal of Hospitality and Tourism Technology*, vol. 14, no. 3, pp. 430–443. https://doi.org/10.1108/JHTT-09-2021-0254

66. Dias R.P., Kusuma N.I. (2023) The effect of brand identity and brand preference on Starbucks repurchase interest in Bekasi City. *Jurnal Ekonomi dan Bisnis Digital*, vol. 2, no. 3, pp. 1031–1054. https://doi.org/10.55927/ministal.v2i3.4207

67. Can Y., Erdil O. (2018) Determining antecedent of re-purchase intention: The role of perceived value and consumer's interest factor. *International Business Research*, vol. 11, no. 4, pp. 17–31. https://doi.org/10.5539/ibr.v11n4p17

68. Ibzan E., Balarabe F., Jakada B. (2016) Consumer satisfaction and repurchase intentions. *Developing Country Studies*, vol. 6, no. 2, pp. 96–100.

69. Langga A., Kusumawati A., Alhabsji T. (2021) Intensive distribution and sales promotion for improving customer-based brand equity (CBBE), re-purchase intention and word-of-mouth (WOM). *Journal of Economic and Administrative Sciences*, vol. 37, no. 4, pp. 577–595. https://doi.org/10.1108/JEAS-03-2019-0041

70. Yasri Y., Susanto P., Hoque M.E., Gusti M.A. (2020) Price perception and price appearance on repurchase intention of Gen Y: do brand experience and brand preference mediate? *Heliyon*, vol. 6, no. 11, e05532. https://doi.org/10.1016/j.heliyon.2020.e05532

71. Sullivan Y.W., Kim D.J. (2018) Assessing the effects of consumers' product evaluations and trust on repurchase intention in e-commerce environments. *International Journal of Information Management*, vol. 39, pp. 199–219. https://doi.org/10.1016/j.ijinfomgt.2017.12.008

72. Tynan C., McKechnie S. (2009) Experience marketing: a review and reassessment. *Journal of Marketing Management*, vol. 25, no. 5–6, pp. 501–517. https://doi.org/10.1362/026725709X461821

73. Diallo M.F., Siqueira J.R. (Jr.) (2017) How previous positive experiences with store brands affect purchase intention in emerging countries: A comparison between Brazil and Colombia. *International Marketing Review*, vol. 34, no. 4, pp. 536–558. https://doi.org/10.1108/IMR-07-2014-0224

74. Moreira A.C., Fortes N., Santiago R. (2017) Influence of sensory stimuli on brand experience, brand equity and purchase intention. *Journal of Business Economics and Management*, vol. 18, no. 1, pp. 68–83. https://doi.org/10.3846/16111699.2016.1252793

75. Bell E., Bryman A. (2007) The ethics of management research: an exploratory content analysis. *British Journal of Management*, vol. 18, no. 1, pp. 63–77. https://doi.org/10.1111/j.1467-8551.2006.00487.x

76. Sekaran U., Bougie R. (2016) *Research methods for business: A skill building approach*. Wiley.

77. Amoako G.K., Anabila P., Asare Effah E., Kumi D.K. (2017) Mediation role of brand preference on bank advertising and customer loyalty: A developing country perspective. *International Journal of Bank Marketing*, vol. 35, no. 6, pp. 983–996. https://doi.org/10.1108/IJBM-07-2016-0092

78. O'Leary-Kelly S.W., Vokurka R.J. (1998) The empirical assessment of construct validity. *Journal of Operations Management*, vol. 16, no. 4, pp. 387–405. https://doi.org/10.1016/S0272-6963(98)00020-5

79. Anderson J.C., Gerbing D.W. (1988) Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, vol. 103, no. 3, 411423. https://doi.org/10.1037/0033-2909.103.3.411

80. Fornell C., Larcker D.F. (1981) Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, vol. 18, no. 1, pp. 39–50. https://doi.org/10.2307/3151312

81. Henseler J., Ringle C.M., Sarstedt M. (2015) A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, vol. 43, no. 1, pp. 115–135. https://doi.org/10.1007/s11747-014-0403-8

82. Nunnally J.C. (1978) *Psychometric Theory (2nd edition)*. New York: McGraw-Hill Book Company.

83. Prentice C., Dominique Lopes S., Wang X. (2020) The impact of artificial intelligence and employee service quality on customer satisfaction and loyalty. *Journal of Hospitality Marketing and Management*, vol. 29, no. 7, pp. 739–756. https://doi.org/10.1080/19368623.2020.1722304

84. Huang M.H., Rust R.T. (2018) Artificial intelligence in service. *Journal of Service Research*, vol. 21, no. 2, pp. 155–172. https://doi.org/10.1177/1094670517752459

85. McDowell W.C., Wilson R.C., Kile C.O. (Jr.) (2016) An examination of retail website design and conversion rate. *Journal of Business Research*, vol. 69, no. 11, pp. 4837–4842. https://doi.org/10.1016/j.jbusres.2016.04.040

86. Chang P.L., Chieng M.H. (2006) Building consumer–brand relationship: A cross-cultural experiential view. *Psychology & Marketing*, vol. 23, no. 11, pp. 927–959. https://doi.org/10.1002/MAR.20140

About the authors

Raed Naser Alkaied

Instructor and Head of department, Faculty of Business, Management Information System Department, Al Balqa Applied University, Salt 19117, Jordan, PO Box 206;

E-mail: Raedalkaied@bau.edu.jo

ORCID: 0000-0002-6288-7503

Shadi Ahmed Khattab

Associate Professor, Faculty of Business, Management Information System Department, Al Balqa Applied University, Salt 19117, Jordan, PO Box 206;

E-mail: Shadikhattab@bau.edu.jo

ORCID: 0000-0002-0824-1437

Ishaq M. Al Shaar

Professor, Faculty of Business, Department of Business Administration, Al Balqa Applied University, Salt 19117, Jordan, PO Box 206;

E-mail: I.shaar@bau.edu.jo

ORCID: 0000-0001-6036-4189

Mohammed Khair Abu Zaid

Professor, Faculty of Business, Planning and Project Management Department, Al Balqa Applied University, Salt 19117, Jordan, PO Box 206;

E-mail: Mohammed_abu_zaid@bau.edu.jo

ORCID: 0000-0002-6687-1285

Sakher A.I. Al-Bazaiah

Associate Professor, Faculty of Business, Department of Business Administration, Al Balqa Applied University, Salt 19117, Jordan, PO Box 206;

E-mail: bazaiah1@bau.edu.jo

ORCID: 0000-0002-6648-8091