DOI: [10.17323/2587-814X.2020.3.54.66](https://doi.org/10.17323/2587-814X.2020.3.54.66)

An approach to identifying bots in social networks based on the special association of classifiers

Vladimir N. Kuzmin

E-mail: kuzmindvn@ya.ru

Artem B. Menisov

E-mail: men.arty@yandex.ru

Ivan A. Shastun

E-mail: shastunivan1982@gmail.com

Space Military Academy named after A.F. Mozhaysky

Address: 13, Zhdanovskaya Street, Saint Petersburg 197198, Russia

Abstract

Currently the use of bots, i.e. auto-accounts in social networks which are managed with special programs but disguised as ordinary users, has serious consequences. For example, bots have been used to influence political elections, distort information on the Internet and manipulate prices on the stock exchange. Many research teams concerned with the detection of such accounts have made use of machine learning methods. However, the practical results of detecting social network bots indicate significant limitations because the methodological tools used have language limitation and ineffective criteria for detection. This article presents improved countermeasures in a methodological approach to develop a universal social network account classifier for minimizing the average risk of errors in bot detection. The application of an assembly of classifiers united by a data adaptation criterion and results from the variance of each model found the formation of a universal classifier for social network accounts. The main results obtained by the authors consist of the criteria system and the categorical (nominal) features transformation approach for the formation of the special ensemble of classifiers. In practice, use of the ensemble of classifiers allows us to increase the effectiveness of bot detection compared to other approaches.

Key words: bot detection; social networks; machine learning; ensemble of models; association of classifiers.

Citation: Kuzmin V.N., Menisov A.B., Shastun I.A. (2020) An approach to identifying bots in social networks based on the special association of classifiers. *Business Informatics*, vol. 14, no 3, pp. 54–66. DOI: [10.17323/2587-814X.2020.3.54.66](https://doi.org/10.17323/2587-814X.2020.3.54.66)

Introduction

The detection of bots on social networks has been a subject of research for over a decade [1] due to their active use for political propaganda purposes. However, to date, no unambiguous interpretation of the term “social network bot” has been formed [1]. In this research, a bot will be understood as a special page (account) of a social network disguised as an ordinary user, which automatically and/or according to a schedule performs actions to publish, promote and comment on materials aimed at achieving certain propaganda or political (economic) goals. Depending on the scope of application, several characteristic goals can be distinguished, as presented in *Table 1*.

Table 1.

**Informational purposes
for using social network bots**

No	Applications	Informational purposes
1.	political	promoting ideology; the imposition of political views; agitation; propaganda; to draw the attraction of the electorate
2.	economic	advertising of goods and services; increasing brand awareness
3.	social	increasing personality recognition; black, gray and white PR
4.	moral	propaganda and change of ideological stereotypes

In modern conditions, the use of bots of social networks has become a threat aimed at discrediting the legitimate government and worsening the controllability of social processes and organizations [2]. Currently, according to the number of bots and the consequences of their actions, the social network Twitter stands out: it has more than a million such accounts [3], and individual botnets contain up to half a million accounts [4].

Many organizations are interested in the development and improvement of means of countering the use of bots [5, 6], for identifying bots, evaluating the results of their actions, and neutralizing the consequences. Analysis of the published research results on this topic [7–22] shows that machine learning and neural networks are widely used to detect bots of social networks. There are two main approaches to detect bots of social networks: based on the processing of materials published by users (*Table 2*) and based on the processing of the quantitative and qualitative characteristics of the accounts themselves (*Table 3*).

Despite the achievement of good results of the approaches developed [1, 7–16], the following disadvantages can be identified:

- ♦ lack of a sufficiently complete set of data to check the quality of detection of social network bots;
- ♦ linguistic limitations of the methods used.

In connection with the above disadvantages, as well as to ensure universality, it is advisable to detect bots based on the results of the analysis of quantitative and qualitative characteristics of accounts, which are interpreted by some authors as “meta-features” [3, 6–8, 11].

As the algorithms that control social networking bots become more sophisticated, so too have algorithms for detecting bots. In studies [17–22], bot detection models range from the simplest [17–19], designed to analyze a single piece of meta-features, to models that use ensemble approaches to analyze large data sets, including a combination of meta-features, actions of social network accounts and content [20–22]. Model ensembles can detect new bot behavior that cannot be detected by individual models since the latter can only detect bots that are sufficiently similar to the data used for training [17].

Table 2.

**Results of analysis of approaches and effectiveness
of bots detection in social networks
(based on the processing of materials published by users)**

No	Authors	Approach	Effectiveness	Language	Note
1.	A. Bacciu, M. La Morgia, A. Mei, E. Nerio Nemmi, V. Neri, J. Stefa [7]	latent semantic analysis of tweets	accuracy greater than 0.9	English, Spanish	The research was conducted on data provided at the conference PAN 2019
2.	P. Gamallo, S. Almatarneh [8]	Bayesian classifier	accuracy for English – 0.81, for Spanish – 0.88	English, Spanish	PAN 2019
3.	I. Vogel, P. Jiang [9]	principal component analysis, N-gram	accuracy for English – 0.92, for Spanish – 0.91	English, Spanish	PAN 2019
4.	A. Mahmood, P. Srinivasan [10]	TF-IDF	accuracy – 0.91	English	
5.	M. Farber, A. Qurdina, L. Ahmed [11]	neural network (CNN)	accuracy – 0.9	English	PAN 2019

Table 3.

**Results of analysis of approaches and effectiveness
of bot detection in social networks
(based on processing the quantitative
and qualitative characteristics of the accounts)**

No	Authors	Approach	Effectiveness
1.	J. Lundberg, J. Nordqvist, M. Laitinen [12]	random forest	accuracy greater than 0.9
2.	S.R. Sahoo, B.B. Gupta [13]	Petri nets	accuracy – 0.9916
3.	J. Novotny [14]	random forest	accuracy – 0.9
4.	A. Davoudi, A. Z Klein, A. Sarker, G. Gonzalez-Hernandez [15]	Botometer	F-score – 0.7
5.	M. Mazza, S. Cresci, M. Avvenuti, W. Quattrociocchi, M. Tesconi [16]	neural network (LSTM)	F-score – 0.87

1. Statement of the research problem

The sets of accounts of the social network A and their states $Y = \{0, 1\}$ are given, and there is a statistical function $y^*: A \rightarrow Y$, the values of which $y_i = y^*(a_i)$ are known only on a finite

subset of objects $\{a_1, a_2, \dots, a_n\} \subset A$, and for $y_i = 1$ the account is a bot, and for $y_i = 0$ it is an ordinary user. The “object-state” pairs are use cases. The set of pairs (a_i, y_i) of precedents $A' = (a_i, y_i)_{i=1}^l$ is a training set for restoring the dependence y^* .

The problem of social networks bot detection is to construct a decision function $z: A \rightarrow Y$, as close as possible to $y^*(a)$, and not only on the objects of the training set but also on the entire set A . In other words, it is necessary to determine the state of an arbitrary social network account $a \in A$. In this case, the probability of correct classification and the probability of errors define the average risk of detecting bot error:

$$[E] = p_0 0 + p_1 E_1 + \dots + p_i E_i, \quad (1)$$

where H – the risk of detecting bot error;

$\mathbb{E} [E]$ – the mathematical expectation of detection errors;

E – the detection errors set;

$\langle E_1, \dots, E_i \rangle$ – the detection errors;

i – the number of states classes of social networks accounts;

p_0 – the probability of a correct decision;

$\langle p_1, \dots, p_i \rangle$ – the probability of errors.

Thus, the task of detecting social network bots is to form a decision on the state of a social networks account at the observed moment in time. In turn, the requirements for the quality of the decision are determined in the form of requirements to minimize the risk associated with making the wrong decision:

$$H = \mathbb{E} [E] = p_1 E_1 + p_2 E_2 \rightarrow \min, \quad (2)$$

where H – the risk of detecting bot errors;

$\mathbb{E} [E]$ – the mathematical expectation of detection errors;

p_1 – the probability of type 1 errors, i.e. when a user account is mistakenly classified as a bot;

p_2 – the probability of type 2 errors, i.e. when the bot skips.

2. Methods

2.1. Social networks bot detection criteria

It should be noted that it is not recommended to detect social network bots only by one indicator (for example, only by the number of pub-

lications or the number of subscribers) [22]. A combination of such features as the thematic relationship of accounts, activity, anonymity, and, in some cases, data inconsistency is important.

Thematic relationship of accounts. The presence of links, subscriptions or other actions by multiple accounts on a specific thematic cluster of accounts (or on one account) is a sign of using bots since one of the main tasks of bots is to “amplify the signal” of other users, not only by commenting and quoting them. In social networks, a ranking system is used that increases the degree of distribution of account materials, depending on the number of not only subscribers but also those who passively view the material or simply navigate to the pages.

Activity. The most obvious sign of a bot is its activity. This feature can be determined by open data (for example, the number and frequency of posts and subscriptions since the creation of the account).

Anonymity. The third important feature is the degree of account anonymity. In general, the less personal information an account has, the more likely it is a bot. Also, a bot feature is the configured privacy for the page.

The inconsistency of the data may lie in the inconsistency of the language of the materials and the place of the account foundation, the place of the account foundation and the time zone of the account.

2.2. Formation of an initial data set

The use of social network bots attracted particular public attention in South Africa when Bell Pottinger used social media to spread negative content. [23] Among the subscribers to the accounts of South African politicians there were bots that followed and moderated their tweets on their pages [24].

Analysis of data from two accounts of the South African politicians (Paul Mashatile, chairman of the African National Congress (ANC) in Gauteng province [25], and Ayanda Dlodlo, a member of the ANC [26]), collected in September 2018, showed that out of 12,000 active subscribers 863 subscribers are shared (*Figure 1*). Of this category of users, 121 accounts were clearly used to increase the rating of propaganda material. They were selected for the following indicators: the account distributed more than 100 messages per day and was also distinguished by high activity, anonymity, and data inconsistency.

2.3. Description of the methodology for using a special assembly of classifiers

The methodological sequence for identifying social network bots includes the following stages: searching and saving data from social network accounts, preprocessing data, select-

ing and training individual machine learning models, combining them (models), determining the state of social network accounts.

Stage 1. Search and save data from social network accounts. This phase is aimed at collecting all available information about social network accounts using the API (application program interface) methods of the social network Twitter [27]. The data description is presented in *Table 4*.

Stage 2. Preprocessing data. This stage includes filling in empty (incorrect) rows, as well as transforming categorical (nominal) features [28].

Feature f of the object $a \in A$ (where A is the set of social network accounts) is the result of evaluating some characteristic of the object [28]. Formally, a feature is a function $f: A \rightarrow D_f$, where D_f is the set of valid feature values. If D_f is a finite set, then f is a nominal feature, and $f_1(a), \dots, f_n(a)$ is a feature description of the object $a \in A$. For the research, we will assume that $A = D_{f_1} \times \dots \times D_{f_n}$.

Average number of publications

Fig. 1. Posting frequency analysis of common active subscribers of two members of the African National Congress

Table 4.

Accessible Twitter account information

No	Name	Definition	Type	Example
1	ID	identification number	string	4452841
2	Screen name	display name	string	sToneBirD
3	Date of creation	account creation date	date	2007-04-13 04:33:54
4	Favorites	pages that are followed	int	43
5	Followers	followers	int	386
6	Friends	friends	int	798
7	GEO	geographic settings	category	True
8	Lang	language	category	En
9	List	lists	int	18
10	Location	location	category	Doha, Qatar
11	Protected	privacy	category	False
12	Status count	number of posts	int	3770
13	Time zone	Time zone	category	Seoul
14	URL	address	string	http://pbs.twimg.com/ ...
15	Verified	verification	category	False

To transform categorical (nominal) features, we use the hot coding technique [28], which allows all categories to be represented as discrete values. Let us assume that the categorical variable $d_i \in D_{f(cat)}$, $i = 1, \dots, k$ for a given similarity of categories $sim(d_i, d_j): D_{f(cat)} \times D_{f(cat)} \rightarrow [0, 1]$ after that the set of values of the feature $f^{sim} \in R$ is determined in the form:

$$f^{sim} = [sim(d_i, d_1), \dots, sim(d_i, d_k)], \quad (3)$$

where $d_k \in D_{f(cat)}$ – the set of all categories.

The categorical feature transformation approach avoids performing one of the laborious stages of training machine learning models – normalizing records in databases [29].

Stage 3. Selection and training of models.

The approach developed is focused on the use of an ensemble of classifiers. Ensembles are well known for their effect of improving the accuracy and generalization of a solution, as

well as providing parallelism. They have been successfully used in various problems of binary classification [30]. Since the activity of social network bots includes categorical features, adaptation to the initial data is necessary to use ensembles of models (Stage 2). A particularity of the special association of classifiers is that the classifier union is a wrapper for many different models that operate in parallel (Figure 2) to take advantage of the different strengths of each model [31].

Creation of the special association of classifiers consists in performing the following actions:

1. Development of N separate models, each of which has its own values of detection accuracy. Since each model is trained with multiple partitions of the data sample for sliding control, the value of the number of models (N) depends on the statistical robustness of the results and the improvement of the solution.

Fig. 2. Functional diagram of the special association of classifiers
 $(D_{f1} \times \dots \times D_{fn})$ – the set of features,
 $\langle M_1, \dots, M_n \rangle$ – the set of models,
 (p_1, p_2) – the detection errors)

2. Train each model separately. The learning from the precedents of each model is reduced to the selection of the best value of the model's hyperparameters (controllers, external) [32]. For example, in a polynomial regression model, an attempt to optimize the degree of the polynomial over the training sample will result in the selection of the maximum possible degree and overfitting.

3. Combining models and improving the values in the final classifier by the following method: the vector of the error probability over the entire set of precedents for each predicted class (for all classifiers) is summed up and averaged. The value of the class is revealed while minimizing the average risk of detection error (expression (2)):

$$y^*(a) \approx z(a), \text{ while} \quad (4)$$

$$\arg \min \frac{1}{N} \sum_{i=1}^N \left[(H(p_1, p_2))_i \right],$$

where $z(a)$ – the class value;

N – the number of models in the final classifier;

$H(p_1, p_2)$ – the risk associated with errors of the first and second type.

Stage 4. Determining the state of social network accounts. The class value (bot or valid account) is chosen by opposing the models to each other, taking into account the weights. For example, suppose that there are three models in the ensemble, the output of which is the following values:

$$z_1(a) = 0, \text{ while } H = (0, 1; 0, 1),$$

$$z_2(a) = 0, \text{ while } H = (0, 5; 0, 5),$$

$$z_3(a) = 1, \text{ while } H = (0, 9; 0, 9).$$

In the output of classification presented, the model doing the union of the solutions can interpret this result as $y^*(a) = 0$, but assigning the weights to the models $\{0, 1, 0, 1, 0, 8\}$ will predict $y^*(a) = 1$. Note that the possibility of such a choice is a significant difference between the methods for combining solutions based on multi-tiered generalization from other approaches, for example, from methods in which the final solution is always chosen from the set of solutions proposed by the basic classifiers [33, 34].

The advantage of the proposed approach is that the ensemble of models can be more easily trained on small input datasets and will improve the bot detection performance compared to any single model.

3. Results

In scientific research on machine learning [35, 36], it is customary to present the results of testing the proposed new method in comparison with other methods on a representative set of problems. The comparison should be carried out under equal conditions using the same methodology (especially if it is a sliding control). Table 5 shows the result of comparing the developed approaches to machine learning models suitable for binary classification and input data into the final classifier.

Table 5.

**The comparison of model results
for detecting social network bots**

No	Model	Learning time	Prediction time	Accuracy	Precision	Recall	F1-score	AUC-ROC
1	The model proposed	0.003908	0.006760	0.994577	0.991750	0.986458	0.994471	0.999306
2	Logistic regression	0.003907	0.003127	0.949845	0.921635	0.872000	0.946988	0.872000
3	Random forest	0.010161	0.008214	0.941092	0.925735	0.815792	0.936246	0.964767
4	K-neighbors	0.000000	0.012502	0.939055	0.896358	0.864667	0.938098	0.936906
5	Linear discriminant analysis (LDA)	0.087689	0.012163	0.866311	0.433155	0.500000	0.804308	0.506250
6	Support vector machines (SVM)	0.002342	0.005471	0.858846	0.753099	0.841125	0.872463	0.908184
7	Multinomial naive Bayes (NB)	0.00424	0.01271	0.807822	0.750054	0.803255	0.842993	0.807844
8	Quadratic discriminant analysis (QDA)	0.001763	0.004887	0.740661	0.678199	0.809625	0.777933	0.933076

Such an increase in the classification accuracy can be explained by the fact that each model has a weight characterizing the importance of the contribution to the overall solution, which is calculated by the formula (3). The contribution of each classifier can be interpreted as an assessment of his competence used to scale the outputs (work results) of the classifiers, thereby increasing or decreasing the contribution of each classifier to the overall solution.

Figure 3 shows the variance of the classification results (accuracy) of the developed approach and individual models included in the final ensemble obtained with sliding control.

4. Discussion

To assess the quality of the approach we developed, we define the validation part as equal

to 0.3 of the total data set (clause 1.2), which includes 47 records for bots and 242 records for real accounts.

To assess the quality of the output data, let us construct the bot detection confusion matrix shown in Figure 4a.

The confusion matrix displays the number of correct and incorrect detections compared to the actual data:

- ◆ (0,0) – correctly identified real social networks accounts;
- ◆ (1,1) – correctly identified bots;
- ◆ (0,1) – for ordinary accounts, it was decided that they are bots;
- ◆ (1,0) – the decision was made for bots that they are real accounts.

These probabilities of the first and second type can be calculated as the probability of the random variable z falling into the range

Fig. 3. The contribution of classifiers to the overall solution

Fig. 4. The confusion matrix of bot detection among subscribers of the South African politicians accounts:
a) using the proposed approach; b) using the CatBoost

of acceptable values of the classes of social networks accounts, that is, $p_1 = P(0,1)$ and $p_2 = P(1,0)$. Substituting these values from the confusion matrix into formula (1), we get:

$$H = \mathbb{E}[E] = \frac{2}{241}2 + \frac{1}{47}1 \approx 0.038,$$

where H – the risk of detecting bots error;

$\mathbb{E}[E]$ – the mathematical expectation of detection errors.

Let us compare the average risk obtained with the results of the CatBoost model [37] developed by the Russian company Yandex (Figure 4b). It is based on gradient boosting with the implementation of the categorical (nominal) feature transformation approach:

$$H_{CatBoost} = \mathbb{E}[E_{CatBoost}] = \frac{1}{253}1 + \frac{11}{47}11 \approx 2.57,$$

where $H_{CatBoost}$ – the risk of detecting bots error of the CatBoost model;

$\mathbb{E}[E_{CatBoost}]$ – the mathematical expectation of detection errors of the CatBoost model.

We also compare this result with the result obtained on the basis of the average risk of a random choice:

$$H_{random} = \mathbb{E}[E_{random}] = \frac{126}{253}126 + \frac{24}{47}24 \approx 25.195,$$

where H_{random} – the risk of random detecting bots error;

$\mathbb{E}[E_{random}]$ – the mathematical expectation of random detection errors.

Thus, the proposed approach showed the best result, which characterizes an increase in the quality of detecting social network bots.

Conclusion

The development of new approaches to improve the security of government organizations and users of information web-systems is a constant and urgent task.

An element of the scientific novelty of the

approach we developed for identifying bots in social networks is the recommended combination of a number of features: thematic relationship of accounts, activity, anonymity and data inconsistency. A particularity of this approach is taking into account the growing trend of using one set of bots to achieve different information goals.

The practical significance of the study lies in the possibility of applying the proposed approach in the substantiation and development of technical solutions for information security.

The approach we developed to identifying bots in Twitter based on a special combination of classifiers has an advantage in terms of efficiency compared to modern machine learning algorithms and reduces errors in detecting bots. Since the activity of bots in social networks includes categorical features, adaptation to the original data is necessary to use the ensemble of models.

However, despite the advantages of machine learning, one of the main disadvantages of the developed approach may be its impracticality if there are too many unique records, for example, if the string representations of categorical features display typos or combinations of several data in the same records.

As directions for the further development of this study, the following can be distinguished:

- ◆ research into the collection of additional data on social network accounts;
- ◆ analysis of the impact of data imbalance on training models;
- ◆ research into the possibilities of improving the performance of detecting social network bots;
- ◆ development of technical solutions to improve services for detecting bots of different types. ■

References

1. Williamson W., Scrofani J. Trends in detection and characterization of propaganda bots. *Proceedings of the 52nd Hawaii International Conference on System Sciences. Honolulu, USA, 8–11 January 2019*, pp. 7118–7123. DOI: 10.24251/HICSS.2019.854.
2. Lukyanov R.V. (2018) Methodology for monitoring the state of information security of automated systems in the context of heterogeneous mass incidents. *Transactions of the Military Space Academy named after A.F. Mozhaysky*, no 660, pp. 111–115 (in Russian).
3. *As many as 48 million Twitter accounts aren't people, says study*. CNBC. Available at: <https://www.cnbc.com/2017/03/10/nearly-48-million-twitter-accounts-could-be-bots-says-study.html> (accessed 1 December 2019).
4. *Massive networks of fake accounts found on Twitter*. BBC. Available at: <http://www.bbc.co.uk/news/technology-38724082> (accessed 1 December 2019).
5. Terdimia D. *Here's how Facebook uses AI to detect many kinds of bad content*. Fast Company. Available at: <https://www.fastcompany.com/40566786/heres-how-facebook-uses-ai-to-detect-many-kinds-of-bad-content> (accessed 5 December 2019).
6. *Fighting disinformation online*. RAND. Available at: <https://www.rand.org/research/projects/truth-decay/fighting-disinformation.html> (accessed 5 December 2019).
7. Bacciu A., La Morgia M., Nemmi E., Neri V., Mei A., Stefa J. (2019) Bot and gender detection of Twitter accounts using distortion and LSA. *Proceedings of the Conference and Labs of the Evaluation Forum (CLEF 2019). Lugano, Switzerland, 9–12 September 2019*. Available at: http://ceur-ws.org/Vol-2380/paper_210.pdf (accessed 03 April 2020).
8. Gamallo P., Almatarneh S. (2019) Naive-Bayesian classification for bot detection in Twitter. *Proceedings of the Conference and Labs of the Evaluation Forum (CLEF 2019). Lugano, Switzerland, 9–12 September 2019*. Available at: http://ceur-ws.org/Vol-2380/paper_194.pdf (accessed 03 April 2020).
9. Vogel I., Jiang P. (2019) Bot and gender identification in Twitter using word and character N-grams. *Proceedings of the Conference and Labs of the Evaluation Forum (CLEF 2019). Lugano, Switzerland, 9–12 September 2019*. Available at: http://ceur-ws.org/Vol-2380/paper_65.pdf (accessed 03 April 2020). DOI: 10.13140/RG.2.2.28481.71528.
10. Mahmood A., Srinivasan P. (2019) Twitter bots and gender detection using tf-idf. *Proceedings of the Conference and Labs of the Evaluation Forum (CLEF 2019). Lugano, Switzerland, 9–12 September 2019*. Available at: http://ceur-ws.org/Vol-2380/paper_253.pdf (accessed 03 April 2020).
11. Farber M., Qurdina A., Ahmedi L. (2019) Identifying Twitter bots using a convolutional neural network. *Proceedings of the Conference and Labs of the Evaluation Forum (CLEF 2019). Lugano, Switzerland, 9–12 September 2019*. Available at: http://ceur-ws.org/Vol-2380/paper_227.pdf (accessed 03 April 2020).
12. Lundberg J., Nordqvist J., Laitinen M. Towards a language independent bot detection. *Proceedings of the 4th Conference on Digital Humanities in the Nordic Countries (DHN 2019). Copenhagen, Denmark, 5–8 March 2019*. P. 308–319.
13. Sahoo S.R., Gupta B.B. (2019) Hybrid approach for detection of malicious profiles in Twitter. *Computers & Electrical Engineering*, no 76, pp. 65–81. DOI: 10.1016/j.compeleceng.2019.03.003.
14. Novotny J. (2019) *Twitter bot detection & categorization – a comparative study of machine learning methods*. Master's thesis in Statistics. Lund: Lund University.
15. Davoudi A., Klein A.Z., Sarker A., Gonzalez-Hernandez A. (2019) Towards automatic bot detection in Twitter for health-related tasks. *Proceedings of the AMIA Joint Summits on Translation Science. 2 3–26 March 2020*, pp. 136–141.
16. Mazza M., Cresci S., Avvenuti M., Quattrociocchi W., Tesconi M. (2019) RTbust: Exploiting temporal patterns for botnet detection on Twitter. *Proceedings of the 10th ACM Conference on Web Science. Boston, MA, USA, 30 June – 3 July 2019*, pp. 183–192.

17. Beskow D.M., Carley K.M. (2019) Its all in a name: detecting and labeling bots by their name. *Computational and Mathematical Organization Theory*, pp. 1–12. DOI: 10.1007/s10588-018-09290-1.
18. Varol O., Ferrara E., Davis C.A., Menczer F., Flammini A. (2017) Online human-bot interactions: Detection, estimation, and characterization. Proceedings of the *Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017)*, Montreal, Quebec, Canada, 15–18 May 2017. Available at: <https://aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15587/14817> (accessed 3 April 2020).
19. Minnich A., Chavoshi N., Koutra D., Mueen A. (2017) BotWalk: Efficient adaptive exploration of Twitter bot networks. Proceedings of the *2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2017)*, Sydney Australia, 31 July – 03 August 2017, pp. 467–474. DOI: 10.1007/s10588-018-09290-1.
20. Chavoshi N., Hamooni H., Mueen A. (2016) DeBot: Twitter bot detection via warped correlation. Proceedings of the *IEEE 16th International Conference on Data Mining (ICDM 2016)*. Barcelona, Spain, 12–15 December 2016, pp. 817–822. DOI: 10.1109/ICDM.2016.0096.
21. Ferrara E., Varol O., Davis C., Menczer F., Flammini A. (2016) The rise of social bots. *Communications of the ACM*, vol. 59, no 7, pp. 96–104. DOI: 10.1145/2818717.
22. Davis C., Varol O., Ferrara E., Flammini A., Menczer F. (2019) BotOrNot: A system to evaluate social bots. Proceedings of the *25th International Conference Companion on World Wide Web, Montreal, Canada, 11–15 May 2016*, pp. 273–275. DOI: 10.1145/2872518.2889302.
23. Thamm M. (2019) *Analysis: Bell Pottinger more than just spin, its political interference in sovereign states*. Available at: <https://www.dailymaverick.co.za/article/2017-07-05-analysis-bell-pottinger-more-than-just-spin-its-political-interference-in-sovereign-states/#gsc.tab=0> (accessed 5 December 2019).
24. Featherstone C. (2019) South African bot behaviour post the July 2018 Twitter account cull. Proceedings of the *2019 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD)*. Winterton, South Africa, 5–6 August 2019, pp. 1–6. DOI: 10.1109/ICABCD.2019.8851039.
25. *Twitter account Paul Mashatile*. Available at: <https://twitter.com/PaulMashatile> (accessed 4 September 2019).
26. *Twitter account Ayanda Dlodlo*. Available at: <https://twitter.com/MinAyandaDlodlo> (accessed 4 September 2019).
27. *Twitter API Documentation*. Available at: <http://www.developer.twitter.com/docs> (accessed 4 April 2019).
28. Vorontsov K.V. (2019) *Mathematical methods of teaching by procedures (theory of machine learning)*. Available at: <http://www.machinelearning.ru/wiki/images/6/6d/Voron-ML-1.pdf> (accessed 4 December 2019) (in Russian).
29. Zhang W., Du T., Wang J. (2016) Deep learning over multi-field categorical data. Proceedings of the *38th European Conference on Information Retrieval Research (ECIR 2016)*. Padua, Italy, 20–23 March 2016, pp. 45–57.
30. Menisov A.B., Shastun I.A., Kapitsyn S.U. (2019) An approach to the identification of malicious Internet sites based on the processing of lexical signs of addresses (URLs) and an average ensemble of models. *Information Technologies*, vol. 25, no 11, pp. 691–697 (in Russian). DOI: 10.17587/it.25.691-697.
31. Vorontsov K.V. (2019) *Lectures on methods for evaluating and selecting models*. Available at: <http://www.ccas.ru/voron/download/Modeling.pdf> (accessed 5 December 2019) (in Russian).
32. Gorodetsky V.I., Serebryakov S.V. (2006) Collective recognition methods and algorithms: a review. *Transactions of SPIIRAS*, vol. 1, no 3, pp. 139–171 (in Russian). DOI: 10.15622/sp.3.8.
33. Niyogi P., Pierrot J.-B., Siohan O. (2000) Multiple classifiers by constrained minimization. Proceedings of the *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing*. Istanbul, Turkey, 5–9 June 2000, pp. 3462–3465. DOI: 10.1109/ICASSP.2000.860146.
34. Prodromidis A., Chan P., Stolfo S. (1999) Meta-learning in distributed data mining systems: Issues and approaches. *Advances in Distributed Data Mining*, no 3 pp. 81–114.

35. Gnidko K.O., Makarov S.A., Sergeev A.S. (2019) A model of an intellectual decision support system in order to identify the negative informational and psychological impact on students of educational organizations of the Ministry of Defense of Russia and to protect against it. *Transactions of the Military Space Academy named after A.F. Mozhaysky*, no 666, pp. 142–147 (in Russian).
36. Kachura Ya.O., Saprykin D.I., Faleev P.A. (2018) Modeling of the military-political activity of states by the methods of associative analysis in decision support systems. *Transactions of the Military Space Academy named after A.F. Mozhaysky*, no 660, pp. 19–29 (in Russian).
37. *Developer Documentation CatBoost*. Available at: <https://tech.yandex.ru/catboost/> (accessed December 2019).

About the authors

Vladimir N. Kuzmin

Dr. Sci. (Mil.), Professor;

Leading Researcher, Military Institute (Science and Researching),
Space Military Academy named after A.F. Mozhaysky,
13, Zhdanovskaya Street, Saint Petersburg 197198, Russia;

E-mail: kuzmindvn@ya.ru

ORCID: 0000-0002-6411-4336

Artem B. Menisov

Cand. Sci. (Tech.);

Researcher, Military Institute (Science and Researching),
Space Military Academy named after A.F. Mozhaysky,
13, Zhdanovskaya Street, Saint Petersburg 197198, Russia;

E-mail: men.art@yandex.ru

ORCID: 0000-0002-9955-2694

Ivan A. Shastun

Cand. Sci. (Tech.);

Lecturer, Military Institute (Science and Researching),
Space Military Academy named after A.F. Mozhaysky,
13, Zhdanovskaya Street, Saint Petersburg 197198, Russia;

E-mail: shastunivan1982@gmail.com

ORCID: 0000-0002-1086-5345