

Применение метода комитетов к анализу технических индикаторов фондового рынка

Н.П. Чернавин 

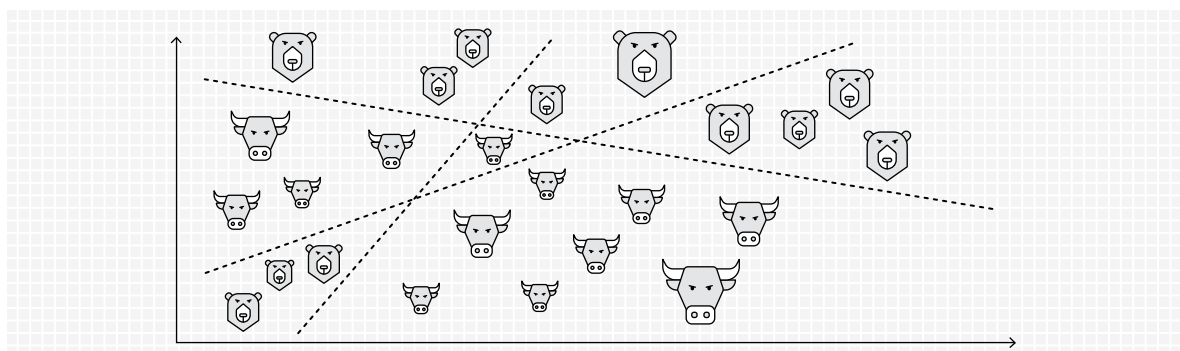
E-mail: ch_k@mail.ru

Институт экономики, Уральское отделение Российской академии наук
Адрес: 620014, г. Екатеринбург, ул. Московская, д. 29

Аннотация

В статье рассматриваются проблемы применения метода комитетов для принятия решений в условиях поступления различных сигналов от технических индикаторов фондового рынка. Метод комитетов является методом классификации данных с учетом нелинейных зависимостей и предусматривает построение группы линейных классификаторов. В рамках данного исследования основой для построения комитетов служит единая модель частично-целочисленного программирования, в рамках которой реализованы различные логики комитетных конструкций. В качестве предмета исследования выступает взаимосвязь показателей технических индикаторов фондового рынка с ценами биржевых финансовых инструментов. Цель исследования – показать эффективность построения комитетных конструкций для решения задач прогнозирования стоимости финансовых инструментов, котирующихся на фондовых рынках. Для достижения указанной цели были собраны основные биржевые данные по акциям ПАО «Сбербанк» (Московская фондовая биржа) за период с 2010 по 2019 годы, на основании которых были рассчитаны технические индикаторы и ряд взаимосвязанных с ними параметров. Эти показатели были использованы в качестве данных для комитетных моделей с различным числом членов комитета и логик голосования. В результате были получены решающие правила, применение которых при ведении спекулятивных торгов на фондовой бирже способно приносить стабильную прибыль. Для сравнения также приведены решения аналогичной задачи классическими методами классификации. Проведенное сравнение показало, что результаты, схожие с комитетным решением по качеству классификации, могут быть получены в рамках методов, способных работать с нелинейными зависимостями данных. Исследование может представлять интерес для профессиональных трейдеров, инвестиционных аналитиков, специалистов по анализу данных и студентов математических и финансовых специальностей.

Графическая аннотация



Ключевые слова: метод комитетов; биржа; технический анализ; машинное обучение.

Цитирование: Чернавин Н.П. Применение метода комитетов к анализу технических индикаторов фондового рынка // Бизнес-информатика. 2019. Т. 13. № 4. С. 73–86. DOI: 10.17323/1998-0663.2019.4.73.86

Введение

В современном мире методы технического анализа могут быть по праву названы наиболее популярными инструментом прогнозирования котировок ценных бумаг на фондовых и валютных биржах. Технический анализ представляет собой совокупность методов прогнозирования динамики биржевых котировок ценных бумаг через изучение внутренних факторов рынка, таких как изменения цены, объемов торгов, открытого интереса и т.д. Популярность методов технического анализа во многом обусловлена наличием, с одной стороны, простых аналитических методов, не требующих специальных знаний, а с другой стороны – возможностью получения экономической выгоды от результата анализа для практически любого человека, обладающего минимальными средствами для открытия биржевого счета.

Для представления общей картины методов технического анализа на *рисунке 1* представлена классификация методов по сложности обработки исходных данных (составлено автором на основе [1, 2]).

В соответствии с представленной схемой, в рамках технического анализа различают формально-аналитические и визуально-графические методы. Первые изучают изменение биржевых показате-



Рис. 1. Классификация методов технического анализа по сложности обработки данных

лей ценных бумаг за прошедший период и на этой основе прогнозируют рост или падение биржевых цен, вторые – предполагают анализ через изучение графических моделей, построенных по биржевым показателям. В данной статье изучается применение метода комитетов для оценки результатов технического анализа цен акций.

Метод комитетов относится к алгоритмам машинного обучения для решения задач бинарной классификации. Термин «комитетная конструкция» был впервые упомянут в 1965 году, в статье по распознаванию образов С.М. Эйблоу и Д.Дж. Кейлора [3]. Значимые результаты по методу комитетов в 1970-е годы были представлены в работах М.Л. Осборна и Р.А. Такиямы [4, 5]. Отечественные теоретические и практические разработки в данной области проводились, главным образом, Екатеринбургской школой распознавания образов Института математики и механики им. Н.Н. Красовского УрО РАН, в работах Вл.Д. Мазурова и М.Ю. Хачая [6–16].

Также достойны внимания исследования по оптимизации комитетных конструкций по степени уверенности в принятом решении, которые были проведены Б.М. Кувшиновым и О.В. Ширяевым [17, 18], а также сведение комитета к проблеме частично-целочисленного программирования в работах Ф.П. Чернавина и О.И. Никонова [19–22].

1. Методология исследования

Комитетные конструкции используются для решения задач нелинейного дискриминантного анализа. Дискриминантный анализ представляет собой группу методов машинного обучения, позволяющих выявлять различия между группами и дающих возможность классифицировать объекты по принципу максимального сходства. Необходимым условием для применения дискриминантного анализа является наличие совокупности признаков объектов (переменных) [23], включая классифицирующую (зависимую) переменную, представляющую собой значение класса объекта, и дискриминационные (независимые) переменные, по которым выявляются различия между объектами с разной классовой принадлежностью.

Итогом дискриминантного анализа является некоторая дискриминантная функция. Каноническая дискриминантная функция линейного вида имеет следующее математическое представление [23]:

$$d = \beta_0 + \sum_{i \in I} (x_i \cdot \beta_i), j \in J, \quad (1)$$

где d — классифицирующая переменная;

I — множество признаков;

J — множество всех изучаемых объектов в пространстве i -го числа признаков;

β_0 — свободный член, обеспечивающий выполнение требуемых условий;

β_i — коэффициент дискриминантной функции для i -го признака;

x_{ij} — дискриминационные переменные для i -го признака j -го объекта.

В рамках комитетной конструкции используется несколько дискриминационных функций, называемых членами комитетов. В зависимости от принимаемого значения классифицирующей переменной, говорят о том, что член комитета голосует «за» или «против» определенного решения. Итоговое решение принимается на основании решения каждого члена с помощью их обработки с использованием логики комитета. Существует три основных логики комитета: единогласия, большинства и старшинства (далее по тексту — КЕ, КБ и КС соответственно).

Для пояснения разницы между различными логиками комитета рассмотрим графические примеры для класса объектов «кружки» и «звездочки», характеризуемыми двумя признаками, представленными в виде значений по осям X_1 и X_2 . Соответственно, выполняются необходимые условия для проведения дискриминантного анализа в виде наличия классифицирующей переменной (бинарного разделения на классы: допустим, «кружки» = 0, а «звездочки» = 1), а также наличия дискриминационных переменных, выраженных значениями каждого объекта по осям X_1 и X_2 . Результаты дискриминационного анализа отображаются в виде линий, представляющих собой отдельные члены комитета (линейные дискриминантные функции), а стрелками у каждой линии обозначается направление голосования соответствующего члена комитета.

В случае КЕ требуется, чтобы все члены комитета проголосовали единогласно за принадлежность объекта к одному классу. Соответственно, все объекты, за которые члены комитета проголосовали не единогласно, будут относиться к другому классу.

Графический пример КЕ из трех членов представлен на *рисунке 2*.

На *рисунке 2* видно, что «кружки» находятся в области графика, за которую все члены комитета проголосовали единогласно. Все наблюдения в остальных областях будут содержать «звездочки».

В свою очередь, в рамках КБ решение о принадлежности объекта к классу принимается при наличии большинства голосов членов комитета. Графический пример такого комитета представлен на *рисунке 3*.

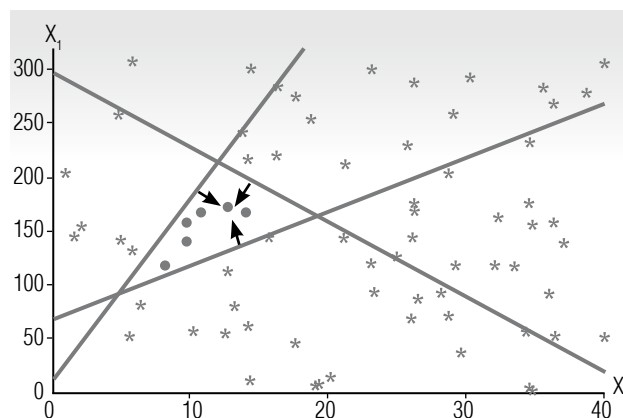


Рис. 2. Пример КЕ в пространстве двух признаков

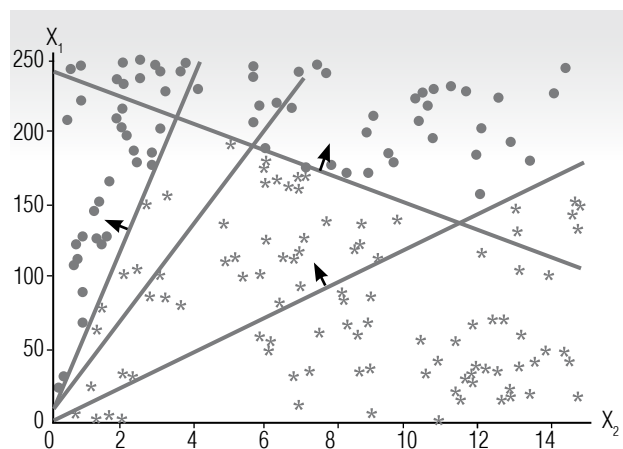


Рис. 3. Пример КБ в пространстве двух признаков

На *рисунке 3* видно, что «кружки» находятся в областях графика, за которую проголосовали хотя бы два (или более) члена комитета. Все объекты в остальных областях будут содержать «звездочки».

В последней из рассматриваемых логик используется принцип старшинства, когда разные члены

комитета могут обладать различными весами, определяющими их значимость при голосовании. Соответственно, решение о принадлежности к классу принимается при достижении в рамках комитета определенного уровня суммы весов за решение. Графический пример КС представлен на *рисунке 4*, где, в дополнение к предыдущим обозначениям, линиям соответствуют веса соответствующих членов комитета.

На *рисунке 4* видно, что «кружки» находятся в областях графика, где сумма весов голосования больше трех. Все объекты в остальных областях будут содержать «звездочки».

Представленные примеры отражают случаи, когда комитет позволяет абсолютно точно разделить объекты на классы. Однако надо понимать, что в реальной жизни, как правило, не все объекты удастся точно отнести к определенному классу. Поэтому комитет должен разделить объекты таким образом, чтобы качество разделения на классы было наилучшим.

Допустим, что существует некоторое исходное разделение объектов на классы кружков и крестиков. На *рисунке 5* представлен графический пример такой классификации.

На *рисунке 5* представлен случай КБ из трех членов в пространстве двух признаков, когда при классификации ряд объектов был отнесен не к тому классу. Для наглядности такие объекты обведены на графике. Наличие ошибок классификации комитетом является естественным явлением, которое может быть связано с множеством факторов, таких как неверная исходная разметка объектов на классы, недостаток исходной информации для более точного разделе-

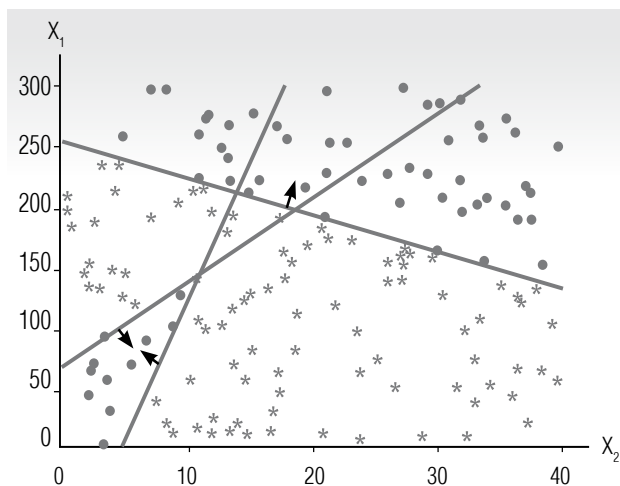


Рис. 4. Пример КБ в пространстве двух признаков

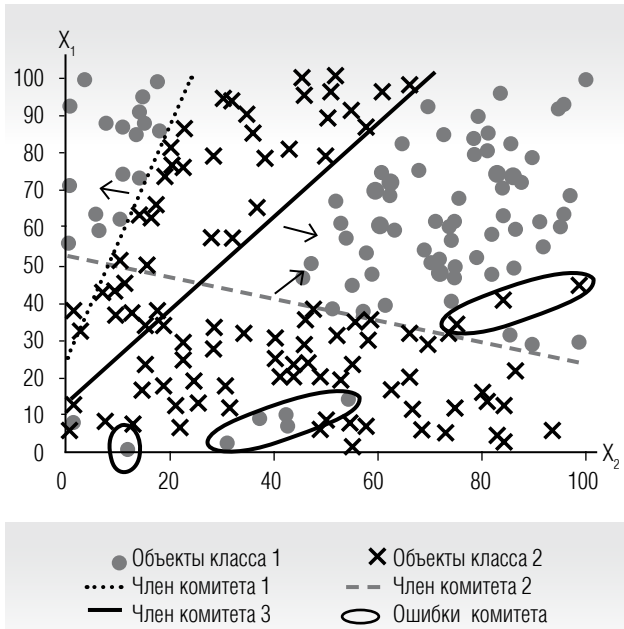


Рис. 5. Пример КБ в пространстве двух признаков при наличии ошибок классификации

ния, недостаточное количество членов комитета, неверно выбранная логика и т.д.

Математически комитет может быть представлен, как модель частично-целочисленного программирования. Ниже представлен пример такой модели:

$$\begin{aligned}
 & \sum_{i \in I} (x_{ij} \cdot \beta_i^t) + \beta_0^t - L \cdot z_j^t \leq -\varepsilon, j \in J_1, t \in T, \\
 & \sum_{i \in I} (x_{ij} \cdot \beta_i^t) + \beta_0^t + L \cdot z_j^t \geq \varepsilon, j \in J_2, t \in T, \\
 & \sum_{t \in T} (z_j^t \cdot V^t) \leq m + L \cdot d_j, j \in J_1, \\
 & \sum_{t \in T} (z_j^t \cdot V^t) \leq sv - m - 1 + L \cdot d_j, j \in J_2, \\
 & \sum_{j \in J_1} (d_j) \leq K_1 \cdot \varphi, \\
 & \sum_{j \in J_2} (d_j) \leq K_2 \cdot \varphi, \\
 & \min \varphi.
 \end{aligned} \tag{2}$$

где J_1 – множество объектов в пространстве i -го числа признаков, принадлежащих к классу 1;

J_2 – множество объектов в пространстве i -го числа признаков, принадлежащих к классу 2;

I – множество признаков;

T – множество членов комитета;

i, j, t – индексы соответствующих множеств;

x_{ij} – i -й признак j -го наблюдения (дискриминационные переменные);

β_i^t – коэффициент для i -го признака t -го члена комитета (коэффициент дискриминантной функции);

β_0^t – свободный член t -го члена комитета (коэффициент дискриминантной функции);

z_j^t – булева переменная, определяющая направление голосования t -й гиперплоскости;

d_j – булева переменная для фиксации ошибок классификации комитета для j -го наблюдения;

L – некоторое большое число (константа), относительно размерности параметров в решаемой задаче (введение данной константы необходимо для того, чтобы $L \cdot z_j^t$ и $L \cdot d_j$, могли выполнять функции невязки в неравенствах);

ε – некоторое малое число (константа), близкое 0 (введение данной константы необходимо для того, чтобы используемые в модели нестрогие неравенства были аналогичны строгим неравенствам $>$ или < 0 ; таким образом удастся исключить решение, при котором все коэффициенты и свободный член равны нулю);

V^t – вес t -й гиперплоскости (константа, соответствующая степени 2);

φ – величина Чебышевской аппроксимации (минимакс). Использование данной величины для минимизации позволяет искать оптимальное решение таким образом, чтобы одновременно минимизировалась доля ошибок классификации для каждого класса объектов;

sv – соответствует сумме весов всех комитетов ($\sum_{t \in T} V^t$);

m – меньшинство для расчета комитета (переменная в диапазоне $0 \leq m \leq sv-1$).

Модель (2) может быть использована для любой из описанных логик комитета. Если $V = 1$, то в зависимости от m может быть построен КЕ ($m = 0$ или $m = sv-1$) или КБ ($0 < m < sv-1$). Если $V = 2^t$, то в зависимости от m может быть построен КЕ ($m = 0$ или $m = sv-1$) или КС ($0 < m < sv-1$).

2. Описание анализируемых признаков

Основной проблемой методов технического анализа является наличие множества ложных сигналов, отсеивание которых требует опыта и интуиции. Попробуем провести такое отсеивание с применением комитетных конструкций. Для исследования технического анализа методом комитетов было выбрано 10 индикаторов рынка (таблица 1).

За основу выбора индикаторов взят набор индикаторов, описанный А. Элдером в его системе «тройного выбора» [24]. При этом сама философия системы «тройного выбора», связанная с тройным просеиванием сделки перед принятием решения, не была сохранена, так как предполагается, что метод комитетов на основе композиции индикаторов способен выявить решающее правило, обладающее большей потенциальной прибыльностью.

Рассмотрим более подробно расчет показателей, представленных в таблице 1.

Для расчета индикаторов P_1 и P_2 необходимо рассчитать гистограмму MACD следующим образом:

1. Вычислить 12-дневное ЕМА на основе цен закрытия;
2. Вычислить 26-дневное ЕМА на основе цен закрытия;
3. Вычесть 26-дневное ЕМА из 12-дневного ЕМА и таким образом получить «быструю» линию, или линию MACD;
4. Вычислить 9-дневное ЕМА «быстрой» линии. Это и будет «медленная», или сигнальная линия;
5. Вычислить разность между «быстрой» линией MACD (п. 3) и «медленной» сигнальной линией (п. 4). Полученное значение будет величиной гистограммы MACD.

Индикатор P_1 будет показывать, сколько дней подряд растет или падает гистограмма MACD. При этом если индикатор падает, то количество дней берется со знаком «минус».

Индикатор P_2 показывает, сколько дней подряд гистограмма MACD была больше или меньше нуля. При этом если индикатор меньше нуля, то количество дней берется со знаком «минус».

В расчете индикаторов P_3 и P_4 показатель SSO используется для определения моментов перекупленности и перепроданности рынка. Он рассчитывается следующим образом:

$$SSO = Mean_n \left(100 \cdot \frac{\sum_s (C_c - Min_r)}{\sum_s (Max_r - Min_r)} \right), \quad (3)$$

где r – временное окно стохастического осциллятора (в рамках данного исследования $r = 5$);

s – индекс временного интервала сглаживания стохастического осциллятора (в рамках данного исследования рассматривается интервал в 3 дня);

Max_r – максимум в интервале r дней;

Таблица 1.

Анализируемые индикаторы

Индикатор	Характеристика	Диапазон значений
P_1	Сколько дней подряд уменьшается или увеличивается гистограмма индикатора схождения–расхождения скользящих средних (moving average convergence divergence, MACD) ¹	Целые числа
P_2	Сколько дней подряд гистограмма MACD больше или меньше нуля	Целые числа
P_3	Сигнал от медленного стохастического осциллятора (slow stochastic oscillator, SSO) ²	от 0 до +1
P_4	Сколько дней подряд SSO подает сильный сигнал	Целые числа
P_5	Сигнал от индекса относительной силы (relative strength index, RSI) ³	от 0 до +1
P_6	Сколько дней подряд RSI подает сигнал	Целые числа
P_7	Тренд, рассчитанный по динамике изменения экспоненциальной скользящей средней (exponential moving average, EMA) от максимума цен	-1, 0, +1
P_8	Сколько дней подряд индикатор Чайкина ⁴ больше или меньше нуля	Целые числа
P_9	Сколько дней подряд индикатор Чайкина уменьшается или увеличивается	Целые числа
P_{10}	Является ли максимум цены более высоким, чем в предыдущий день (1) или минимум – более низким, чем в предыдущий день (-1). Если ни одно из условий не выполняется, то параметр равен нулю	-1, 0, +1

Min_r – минимум в интервале r дней;

C_c – цена закрытия на момент расчета SSO ;

$Mean_n$ – функция расчета среднего значения в интервале n дней (в рамках данного исследования $n = 5$).

Есть различные мнения по поводу интервалов перекупленности и перепроданности рынка. В данном исследовании будем считать, что рынок перепродан, когда $SSO \leq 0,2$ и перекуплен при $SSO \geq 0,8$. Следовательно, не имеет смысла использовать в модели данный индикатор напрямую, потому что в этом случае не будет полностью отражена логика. С этой целью следует преобразовать параметр с помощью следующей формулы:

$$P_3 = (2 \cdot (SSO - 0,5))^3 \quad (4)$$

Индикатор P_4 используется для более детального описания параметра P_3 . Он говорит о том, сколько дней подряд индикатор находится в зоне перекуплен-

ности (берется со знаком «минус») или перепроданности (берется со знаком «плюс»). Данный индикатор служит для примерной оцифровки графика SSO глазами трейдера, когда оцениваются не только текущие значения, но и предыдущая динамика.

Показатель P_5 отражает значение индикатора RSI , который рассчитывается следующим образом:

$$RSI = 1 - \frac{1}{1 + RS}, \quad (5)$$

где RS – отношение среднего значения повышения цен закрытия за 7 дней к среднему значению понижения цен закрытия за 7 дней.

В данном исследовании будем считать, что рынок перепродан, когда $RSI \leq 0,3$ и перекуплен при $RSI \geq 0,7$ (соответственно, сигнал индикатора равняется +1 и -1). Данный параметр рассчитывается по аналогии с P_3 , по формуле:

$$P_5 = (2 \cdot (RSI - 0,5))^3 \quad (6)$$

¹ MACD (moving average convergence/divergence) – индикатор схождения–расхождения скользящих средних. Подробнее с методикой расчета и применения можно ознакомиться в работе [24].

² SSO (slow stochastic oscillator) – медленный стохастический осциллятор. Характеризует соотношение между каждой из цен закрытия и недавним диапазоном максимумов и минимумов. Подробнее с методикой расчета и применения можно ознакомиться в работе [24].

³ RSI (relative strength index) – индекс относительной силы. Характеризует силу рынка через изменения цен закрытия. Подробнее с методикой расчета и применения можно ознакомиться в работе [24].

⁴ Индикатор Чайкина показывает, что происходит с трендом, исходя из отношения разницы цен закрытия и открытия к разности максимальной и минимальной цен, умноженному на объем торгов. Подробнее с методикой расчета и применения можно ознакомиться по ссылке: <https://www.opentrainer.ru/articles/ostillyator-chaykina-chaikin-oscillator/> (дата обращения: 01.09.2019).

Показатель P_6 служит для дополнительного анализа RSI , показывая сколько дней подряд RSI подает сигнал одного типа. При этом данный индикатор может быть как положительным, так и отрицательным, в зависимости от типа сигнала (то есть если сигнал меньше нуля, то количество дней берется со знаком «минус»).

Индикатор тренда был выбран в качестве индикатора P_7 . Классический тренд определяется по наклону прямой, проведенной через локальные максимумы цен. Однако невозможно провести идеальную линию через все локальные максимумы в выбранном интервале исследования, поэтому, как правило, линия проводится эмпирически, чтобы пройти через наибольшее число максимумов цены с минимальным отклонением от них. Тренд имеет три основных состояния: повышающийся, падающий и отсутствие тренда (флэт). Стоит отметить, что из-за отсутствия единой методики тренд может быть определен по-разному. В рамках данного исследования тренд определяется по изменению 22-х дневной ЕМА от максимума дневных цен. Если ЕМА растет второй день подряд, то данный индикатор равен единице до тех пор, пока ЕМА не начнет падать. Соответственно, наоборот: если ЕМА падает второй день подряд, то данный индикатор равен -1 до тех пор, пока ЕМА не начнет расти. Если ЕМА рос или падал в течение только одного дня, то индикатор равен нулю.

Параметры P_8 и P_9 основаны на использовании индикатора Чайкина, который рассчитывается следующим образом:

$$CHI = EMA_3(A/D) - EMA_5(A/D),$$

$$A/D = \frac{C_c - C_o}{Max - Min} \cdot V, \quad (7)$$

где A/D – индикатор накопления/распределения (accumulation/distribution);

C_c – цена закрытия в день расчета индикатора;

C_o – цена открытия в день расчета индикатора;

Max – максимум цены в день расчета индикатора;

Min – минимум цены в день расчета индикатора;

V – объем торгов в день расчета индикатора;

EMA_3, EMA_5 – операции вычисления экспоненциальной скользящей средней с окнами усреднения в 3 и 5 дней соответственно.

Показатель P_8 показывает, в течение скольких дней подряд эта разность находится выше или ниже нуля. При этом если разность меньше нуля, то количество дней берется со знаком «минус».

Показатель P_9 показывает, в течение скольких дней подряд данная разность растет или падает. При этом если значение разности падает, то количество дней берется со знаком «минус».

Последний индикатор P_{10} показывает, был ли максимум цены выше, чем в предыдущий день (значение равно 1), или минимум ниже, чем в предыдущий день (значение равно -1). Если цены не вышли за рамки ценового диапазона предыдущего дня, то значение данного параметра равно нулю. Случаи, когда в течение дня одновременно максимум был выше, а минимум – ниже, чем в предыдущий день были исключены из выборки в связи с их редкостью.

Примеры индикаторов из выборки для акций ПАО «Сбербанк» представлены в таблице 2.

3. Торговая стратегия

Для построения модели комитета первоначально нужно сформировать торговую стратегию. Базовыми условиями торговой стратегии являются:

Таблица 2.

Примеры индикаторов из выборки для акций ПАО «Сбербанк»

Дата	Индикаторы									
	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}
05.02.2010	0,2	1	-0,001	0	-0,36	1	-1	-2	-2	-1
08.02.2010	0,3	2	-0,108	0	-0,342	2	-1	-3	-3	-1
09.02.2010	0	3	-0,291	1	-0,14	3	-1	-4	-4	-1
10.02.2010	-0,1	4	-0,375	2	-0,075	4	-1	-5	0	1
11.02.2010	-0,2	5	-0,166	3	-0,144	5	-1	-6	1	1

1. Стартовый капитал, то есть установленная сумма в рамках, которой можно открыть позицию⁶;

2. Стоп-лосс (заранее установленная величина максимального убытка, при достижении которого текущая открытая позиция будет закрыта)⁷.

В первую очередь, определим, что торговый сигнал поступает от P_3 и P_5 . Если оба параметра меньше 0,25, то это сигнал к покупке (открытию длинной позиции), а если они больше 0,75, то подается сигнал к продаже (открытию короткой позиции). Согласно стратегии, сигнал формируется по окончании торгов дня, но будем считать, что позиция при соответствующем сигнале открывается в последние минуты торгов. При этом позиция закрывается, если появляется сигнал к обратному направлению движения рынка, либо если убыток по позиции превысил величину стоп-лосса, выбранного для этой позиции.

Для выбора величины стоп-лосса существуют различные методы. В рамках данного исследования будем использовать формулу расчета с привязкой к средней за 7 дней дневной волатильности цены⁷, которая будет ограничивать потери для каждого наблюдения. Соответственно стоп-лосс будет рассчитываться по следующей формуле:

$$SL = \bar{V}_7 \cdot s, \quad (8)$$

где \bar{V}_7 – средняя дневная волатильность цены за 7 дней;

s – коэффициент, который показывает, какой долей от средней дневной волатильности позволено рисковать в рамках торговой стратегии.

Описанную стратегию схематично можно представить следующим образом (рисунк 6).

Торговая стратегия, представленная на рисунке 6, построена с использованием двух индикаторов технического анализа – SSO и RSI . Сделки по торговой стратегии заключаются только при поступлении од-

нонаправленного сигнала к покупке или продаже от обоих индикаторов. Однако при этом необходимо понимать, что даже получая сильные сигналы от технических индикаторов, возможно получение как прибыли, так и убытков от торговой деятельности.

4. Задача машинного обучения

Для постановки задачи уменьшения убытков торговую стратегию необходимо представить как задачу дискриминантного анализа, в которой будут распознаваться следующие классы объектов:

1. Класс 1 – множество позиций по торговой стратегии, которые закрывались с прибылью, и прибыль была больше максимального убытка⁸ за период удержания позиции⁹;

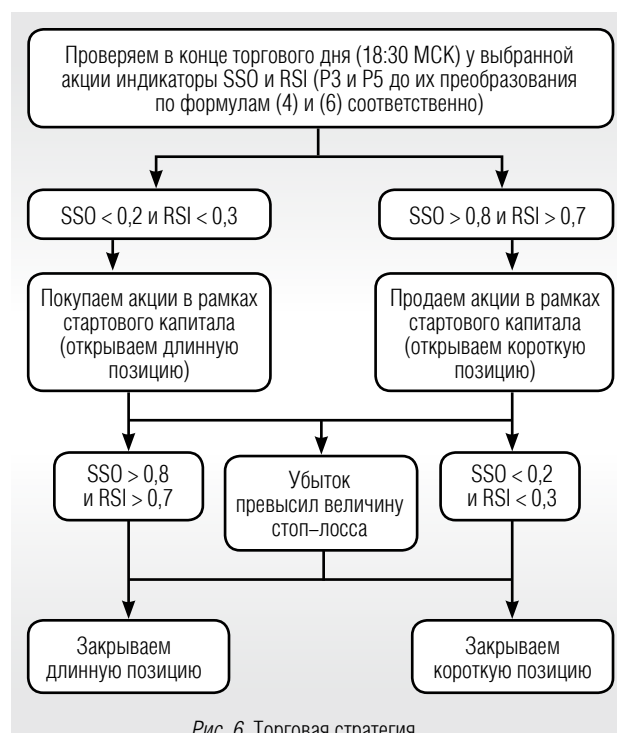


Рис. 6. Торговая стратегия

⁵ Открытие позиции – это первичная покупка или продажа определенного объема финансового инструмента.

⁶ Закрывать торговую позицию – значит совершить обратную по отношению к открытию позиции торговую операцию. Например, если открытие позиции произошло в рамках сделки покупки определенного объема финансового инструмента, то обратная сделка будет выражена последующей продажей приобретенного объема финансового инструмента либо в полном объеме (полное закрытие позиции), либо какой-то его части (частичное закрытие позиции).

⁷ Дневная волатильность – это величина, показывающая, на сколько процентов дневной максимум цены выше дневного минимума цены.

⁸ При открытии позиции на бирже постоянно осуществляется ее переоценка по текущим ценам. Соответственно, максимальный убыток по позиции – это максимальная сумма снижения стоимости позиции относительно ее стоимости при открытии

⁹ Период удержания позиции – это время с момента первичной покупки или продажи определенного объема финансового инструмента до момента торговой операции, обратной по отношению к первой. Подробнее с понятиями открытия и закрытия позиций можно ознакомиться по ссылке: https://www.metatrader5.com/ru/mobile-trading/android/help/trade/positions_manage/open_positions (дата обращения: 01.09.2019)

2. Класс 2 – множество позиций по торговой стратегии, которые закрывались с убытком или прибыль была меньше максимального убытка за период удержания позиции.

В рамках данного исследования в качестве финансового инструмента для анализа были выбраны акции Сбербанка, как наиболее ликвидного актива на рынке акций Московской биржи. Для анализа были использованы дневные котировки по акциям Сбербанка с сайта компании Финам¹⁰ за период с 2010 г. по декабрь 2018 г.

Все наблюдения были разделены на обучающую (выборка, на которой модель обучается) и контрольную (выборка, на которой модель не обучалась) выборки, как показано в *таблице 3*. Такое разделение необходимо, чтобы была возможность проверять согласованность работы модели на этих двух множествах, что позволяет снижать риск переобучения модели [25].

Таблица 3.
Основные параметры обучающей и контрольной выборок

Выборка	Период	J_1	J_2	J
Обучающая	2010–2017 гг.	332	116	448
Контрольная	2018 г.	37	15	52
Итого	2010–2018 гг.	369	131	500

Таблица 3 дает краткое описание обучающей и контрольной выборок для торговой стратегии представленной на *рисунке 6*¹¹. В соответствии с ней количество наблюдений в J_1 примерно в 2,5–3 раза больше, чем в J_2 , и контрольная выборка составляет 11,6% от обучающей выборки. Определившись с задачей классификации и изучаемыми выборками, начнем расчет математической модели.

5. Результаты классификации

Для сравнения по модели (2) были построены различные комитеты, включающие от трех до семи членов. Качество разделения оценивалось при помощи F-меры:

$$SL = \overline{V}_7 \cdot s, \quad (9)$$

где A – точность (доля объектов, отнесенных комитетом к искомому классу и при этом действительно являющимися объектами этого класса);

C – полнота (доля объектов искомого класса, которые комитет смог распознать).

Изучая результаты среди КЕ, решающее правило для случая из пяти членов с результатом 86,4% для J_1 и 68,9% для J_2 показало наилучшую разделяющую способность на обучающей выборке. Для КС наилучшим стал случай из шести членов с результатом 89,1% для J_1 и 73,6% для J_2 . Здесь не представлены результаты для КБ, поскольку математическая модель не смогла найти удовлетворительное по качеству решение для логики большинства¹².

С точки зрения качества модели наибольшее значение имеет результат, полученный на контрольной выборке. Это связано с тем, что этот результат не имеет риска переобучения и поэтому при наличии достаточно большого числа наблюдений объективно оценивает прогностическую способность модели. Соответствующие результаты для F-меры, полученные на контрольной выборке, представлены на *рисунке 7*.

Как видно из *рисунка 7*, КЕ из пяти членов и КС из шести членов также показывают наилучшие результаты. Таким образом, можно отметить, что обучающая и контрольная выборки имеют сильные взаимные связи. Дальнейшее сравнение данных двух комитетных решений позволяет увидеть, что результаты у КЕ из пяти членов выше, чем у КС из шести членов.

Также была рассмотрена модель минимизации количества ошибок при прогнозировании верных или неверных сигналов, поступающих от технических индикаторов. Однако, в рамках этой модели не была учтена величина потенциальных прибылей или убытков в зависимости от принятых решений. Поскольку конечной целью моделирования является получение прибыли, введение данной величины в модель может улучшить конечный результат. Для этого в модели для каждого наблюдения обучающей и контрольной выборок рассчитаем величину потенциальных прибылей или убытков:

¹⁰ Финам – МосБиржа акции – Сбербанк: <https://www.finam.ru/profile/moex-akcii/sberbank/export/> (дата обращения: 01.09.2019).

¹¹ Выборки были построены для торговой стратегии со стоп-лоссом, выбранным по формуле (8) с коэффициентом $s = 2,3$.

¹² При поиске решения в рамках модели (2) при $V=1$ решение с наименьшей величиной φ было найдено при $m=0$, что соответствует КЕ.

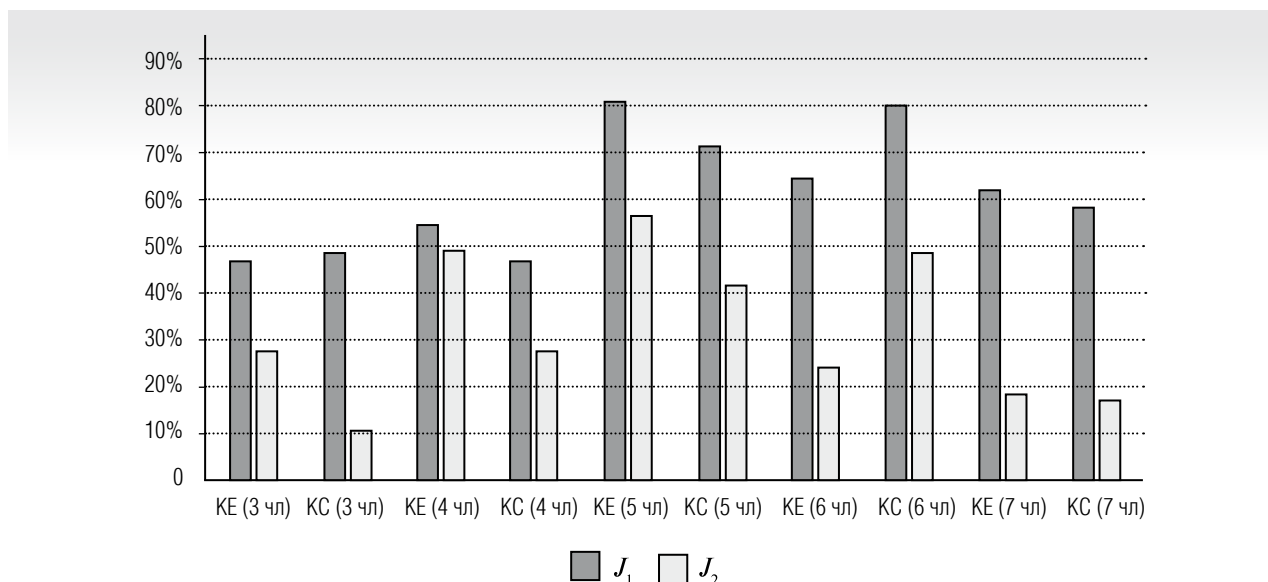


Рис. 7. Результаты для F-меры на контрольной выборке

$$PL = \left(\frac{C_o}{C_c} - 1 \right) \cdot D, \quad (10)$$

где C_o – цена открытия позиции (цена первичной покупки/продажи финансового инструмента, в результате которой сформировалась торговая позиция);

C_c – цена закрытия позиции (цена, по которой ранее открытая позиция была закрыта на бирже, в результате чего по позиции были зафиксированы прибыль или убыток);

D – направление позиции (1 – если финансовый актив куплен; –1 – если финансовый инструмент продан).

Если величина, полученная по формуле (10), больше нуля, то по позиции была получена прибыль, в противном случае имеет место убыток. Для учета данной величины в модели (2) необходимо присвоить веса каждому наблюдению с добавлением соответствующих изменений в ограничениях

$$\sum_{j \in J_1} d_j \text{ и } \sum_{j \in J_2} d_j; \quad (11)$$

$$\sum_{j \in J_1} d_j w_j \leq \sum_{j \in J_1} w_j \varphi,$$

$$\sum_{j \in J_2} d_j w_j \leq \sum_{j \in J_2} w_j \varphi,$$

где w_j – вес j -го наблюдения.

Также дополнительно уменьшим величину коэффициента s в формуле расчета стоп-лосса (8) до 1,5, чтобы снизить уровень убытков по каждой позиции, сформированной по торговой стратегии. Ввиду введения величины стоп-лосса произойдут небольшие

изменения с разделением множества на классы J_1 и J_2 , поскольку 16 наблюдений из J_2 были закрыты по стоп-лоссу (иначе говоря, они были распознаны как J_1). На новом обучающем множестве были также построены различные комитеты с учетом условия (11). Наилучший результат был получен при КБ из семи членов с результатами по F-мере на обучающей выборке для J_1 – 92,1%, для J_2 – 76%, а на контрольной выборке – для J_1 – 83,1%, для J_2 – 51,9%. Данное решающее правило для КБ из семи членов представлено в таблице 4.

Попробуем сравнить по потенциальной норме дохода предыдущее лучшее решающее правило с решающим правилом из таблицы 4. Для этого предположим, что по торговой стратегии каждый раз открывается позиция с одинаковым стартовым капиталом. В таблице 5 представлены результаты годовых сравнений суммарного роста/падения стоимости позиций, выбранных по КБ из семи членов и КЕ из пяти членов.

Результаты, представленные в таблице 5, показывают, что для КБ из семи членов прибыли по верно распознанным наблюдениям из J_2 и убытки по ошибочно распознанным наблюдениям из J_1 имеют значительно меньший разброс значений, чем для КЕ из пяти членов. При этом убытки по ошибочно распознанным наблюдениям из J_1 для КБ из семи членов в два раза меньше, чем у КЕ из пяти членов, в то время как прибыли по верно распознанным наблюдениям из J_2 лишь на 9,6% меньше. Более того, для КБ из семи членов нет такого годового периода, чтобы

Таблица 4

Решающее правило для КБ из семи членов ($m = 2$)

Члены	Кoeffициенты										
	β'_0	β'_1	β'_2	β'_3	β'_4	β'_5	β'_6	β'_7	β'_8	β'_9	β'_{10}
$t = 1$	-1,095	-0,041	0,088	0,006	-0,297	-0,021	0,471	1	0,296	-0,001	-0,298
$t = 2$	-0,352	0,64	-0,006	-0,531	-0,419	0,222	0,168	1	-0,355	0,158	0,224
$t = 3$	-123,18	-153,46	-4,664	6,719	89,83	-39,54	-33,635	-1	25,408	4,748	-75,751
$t = 4$	-1,859	-0,437	-0,193	0,112	0,004	-0,067	0,042	1	-0,062	0,157	0,045
$t = 5$	0,469	-1,666	-0,626	-0,369	-0,787	-0,1	-0,435	1	-1,118	-1,843	-0,142
$t = 6$	-1,573	0,164	0,03	0,058	-0,548	0,05	0,149	-1	-0,155	0,55	-0,723
$t = 7$	-0,906	0,211	0,022	-0,151	0,104	0,124	-0,141	-1	0,073	-0,011	0,263

Таблица 5.

Оценка прибылей и убытков от применения решающих правил

Тип комитета	Класс	% прибыли и убытка по годам									
		2010	2011	2012	2013	2014	2015	2016	2017	2018	Всего
КБ из семи членов	J_1	-13,6	+0,6	-7,4	-32,0	-26,9	-24,6	-8,5	-5,8	-17,3	-136
	J_2	+87,1	+134,3	+37	+46,4	+78,6	+34,8	+67,9	+63,3	+62,8	+612
КЕ из пяти членов	J_1	-53,3	-8,7	-4,5	-14,3	-66,7	+3,8	-43,7	-55,6	-28,4	-271
	J_2	+83,5	+137,8	+35	+65,2	+89,7	+28,3	+101,7	+52,7	+83,7	+677

убытки были больше прибыли, тогда как для КЕ из пяти членов в 2017 году наблюдалось превышение убытков над прибылями. Таким образом, минимизация убытков является более предпочтительным решением для оптимизации торговой стратегии, чем просто минимизация числа нарушений.

Заключение

В рамках исследования показано, что из себя представляет метод комитетов и как в рамках одной математической модели частично-целочисленного программирования могут рассчитываться различные логики комитетов. На практическом примере рассмотрено применение метода комитетов для увеличения точности прогноза биржевых рыночных цен при техническом анализе фондового рынка. Применение комитетных конструкций позволило вывести объективные решающие правила, определяющие, когда следует следовать рекомендациям на основе рыночных индикаторов, а

когда стоит воздержаться от активной торговли на фондовом рынке. Более того, в рамках комитетных конструкций, как модели частично-целочисленного программирования, могут быть сформированы сложные критерии оптимизации с учетом задачи максимизации прибылей и минимизации убытков. Модель с такой формулировкой задачи позволяет получить более стабильные результаты с потенциально более высокой итоговой прибылью, по сравнению с моделями на основе простой минимизации числа ошибок классификации. Таким образом, комитетные конструкции могут быть использованы на финансовых рынках как для формирования новых торговых стратегий, так и для увеличения доходности существующих стратегий. ■

Благодарности

Статья подготовлена в рамках государственного задания для Института экономики Уральского отделения Российской академии наук на 2019 год.

Литература

1. Малышенко К.А., Малышенко В.А., Квятковская Е.О. Теоретические основы анализа фондового рынка: система показателей и классификация методов // Научный журнал КубГАУ. 2017. Том 129. № 5. С. 1292–1303.
2. Найман Э. Малая энциклопедия трейдера. М.: Альпина Паблишер. 2015.
3. Ablow С.М., Kaylor D.J. Inconsistent homogeneous linear inequalities // Bulletin of the American Mathematical Society. 1965. vol. 71. № 5. P. 724.
4. Osborne M.L. The seniority logic: A logic for a committee machine // IEEE Transactions on Computers. 1977. vol. 26. № 12. P. 1302–1306. DOI: 10.1109/TC.1977.1674798.
5. Takiyama R.A. General method for training the committee machine // Pattern Recognition. 1978. Vol. 10. № 4. P. 255–259. DOI: 10.1016/0031-3203(78)90034-1.
6. Мазуров Вл.Д. Метод комитетов в задачах оптимизации и классификации. М.: Наука, 1990.
7. Мазуров Вл.Д., Хачай М.Ю. Комитеты систем линейных неравенств // Автоматика и телемеханика. 2004. № 2. С. 43–54. DOI: 10.1023/B:AURC.0000014716.77510.61
8. Мазуров Вл.Д., Хачай М.Ю. Бустинг и полиномиальная аппроксимируемость задачи о минимальном аффинном разделяющем комитете // Труды Института математики и механики УрО РАН. 2013. Т. 19. № 2. С. 231–236.
9. Мазуров Вл.Д., Хачай М.Ю., Поберий М.И. Задачи комбинаторной оптимизации, связанные с полиэдральной комитетной отделимостью конечных множеств // Труды Института математики и механики УрО РАН. 2008. Т. 14. № 2. С. 89–102.
10. Мазуров Вл.Д., Хачай М.Ю., Рыбин А.И. Комитетные конструкции для решения задач выбора, диагностики и прогнозирования // Труды Института математики и механики УрО РАН. 2002. Т. 8. № 1. С. 66–102.
11. Мазуров Вл.Д., Хачай М.Ю. Комитетные конструкции как обобщение решений противоречивых задач исследования операций // Дискретный анализ и исследование операций. 2003. Сер. 2. Т. 10. № 2. С. 56–66.
12. Мазуров Вл.Д., Хачай М.Ю. Параллельные вычисления и комитетные конструкции // Автоматика и телемеханика. 2007. № 5. С. 182–192. DOI: 10.1134/S0005117907050165.
13. Мазуров Вл.Д., Смирнов А.И. Интерпретация противоречивых изображений на основе систем линейных неравенств // Труды Института математики и механики УрО РАН. 2012. Т. 18. № 3. С. 144–154.
14. Хачай М.Ю. Вопросы вычислительной сложности процедур обучения распознаванию в классе комитетных кусочно-линейных решающих правил // Автоматика и телемеханика. 2010. № 3. С. 178–189. DOI: 10.1134/S0005117910030136.
15. Хачай М.Ю. О существовании комитета большинства // Дискретная математика. 1997. Т. 9. № 3. С. 82–95.
16. Хачай М.Ю. Об оценке числа членов минимального комитета системы линейных неравенств // Журнал вычислительной математики и математической физики. 1997. Т. 37. № 11. С. 1399–1404.
17. Кувшинов Б.М., Ширияев О.В. Метод комитетов в задачах распознавания образов в условиях неопределенности априорной информации // Вестник Южно-Уральского государственного университета. Серия: Математика. Механика. Физика. 2002. Т. 2. № 3. С. 34–43.
18. Система классификации многопараметрических объектов для задач распознавания образов с неточной априорной информацией / Б.М. Кувшинов и [др.] // Информационные технологии. 2001. № 11. С. 37–43.
19. Никонов О.И., Чернавин Ф.П. Построение рейтинговых групп заемщиков физических лиц с применением метода комитетов // Деньги и кредит. 2014. № 11. С. 52–54.
20. Никонов О.И., Чернавин Ф.П., Медведева М.А. Проблемы классификации: метод комитетов // Сборник материалов XII Международная научно-практическая конференция по проблемам экономического развития в современном мире «Устойчивое развитие российских регионов: экономическая политика в условиях внешних и внутренних шоков», Екатеринбург, 17–18 апреля 2015 г. С. 867–874.
21. Nikonov O.I., Medvedeva M.A., Chernavin F.P. Using the committee machine method to forecasting on the FOREX // IEEE 2015 Second International Conference on Mathematics and Computers in Sciences and in Industry (MCSI 2015). Sliema, Malta, 17–19 August 2015. P. 240–243.
22. Чернавин Ф.П. Применение метода комитетов для решения задач классификации // XII Международная конференция «Российские регионы в фокусе перемен». Екатеринбург, 16–18 ноября 2017 г. Екатеринбург: УПИ, 2018. Ч. 1. С. 437–447.
23. Ким Дж.О., Мьюллер Ч.У., Клекка У.Р. Факторный, дискриминантный и кластерный анализ. М.: Финансы и статистика, 1989.
24. Элдер А. Как играть и выигрывать на бирже. Психология. Технический анализ. Контроль над капиталом. М.: Альпина Паблишер. 2016.
25. Ширияев В.И. Нейросетевые методы в анализе финансовых рынков. М.: КомКнига. 2007.

Об авторе

Чернавин Николай Павлович

младший научный сотрудник, Институт экономики, Уральское отделение Российской академии наук, 620014, г. Екатеринбург, ул. Московская, д. 29;

E-mail: ch_k@mail.ru

ORCID: 0000-0002-2093-9715

Application of the committee machine method to analysis of stock market technical indicators

Nikolay P. Chernavin

E-mail: ch_k@mail.ru

Institute of Economics, the Ural Branch of Russian Academy of Sciences
Address: 29, Moskovskaya Street, Ekaterinburg 620014, Russia

Abstract

In this article we study problems of the committee machine method when applied to decision-making when there are many signals from different technical indicators of a stock exchange market. The committee machine method is a data classification method which can find non-linear data dependencies by construction of several linear classifiers. In the framework of this research, the basis for committee machine construction is a unified partially integer programming model, within which various logics of committee structures can be implemented. The subject of the research is the interrelation of indicators of technical indicators of a stock exchange market with pricing for financial instruments of stock exchange trading. Accordingly, the goal of the research is to show the efficiency of committee structures for solving the problems of forecasting the future value of financial instruments listed on stock exchange markets. To accomplish this goal, basic stock exchange data on Sberbank shares were collected from the Moscow Stock Exchange for the period from 2010 to 2019. On the basis of this, the technical indicators and interrelated parameters were calculated. They were used as data for the committee machine models with different numbers of committee members and voting logics. The result of the calculation was to obtain definitive rules, which when applied in speculative trading on the stock exchange market can generate stable profits. For comparison, we show the solutions of a similar problem by classical classification methods. The comparison shows that methods which work with the non-linear data dependencies provide results in terms of classification quality similar to committee machine results. This research may be interesting to the professional traders, investment analysts, specialists in data science and students with a mathematical and/or financial specialization.

Key words: committee machine method; stock exchange market; technical analysis; machine learning.

Citation: Chernavin N.P. (2019) Application of the committee machine method to analysis of stock market technical indicators. *Business Informatics*, vol. 13, no 4, pp. 73–86. DOI: 10.17323/1998-0663.2019.4.73.86

References

1. Malyshenko K.A., Malyshenko V.A., Kvyatkovskaya E.O. (2017) Theoretical bases of the analysis of the stock market: System of indicators and classification of methods. *Scientific Journal of KubSAU*, vol. 129, no 5, pp. 1292–1303 (in Russian).
2. Neiman E. (2015) *The small encyclopedia of trader*. Moscow: Alpina Publisher (in Russian).
3. Ablow C.M., Kaylor D.J. (1965) Inconsistent homogeneous linear inequalities. *Bulletin of the American Mathematical Society*, vol. 71, no 5, p. 724.
4. Osborne M.L. (1977) The seniority logic: A logic for a committee machine. *IEEE Transactions on Computers*, vol. 26, no 12, pp. 1302–1306. DOI: 10.1109/TC.1977.1674798.
5. Takiyama R.A. (1978) General method for training the committee machine. *Pattern Recognition*, vol. 10, no 4, pp. 255–259. DOI: 10.1016/0031-3203(78)90034-1.
6. Mazurov V.I.D. (1990) *The method of committees in optimization and classification problems*. Moscow: Nauka (in Russian).
7. Mazurov V.I.D., Khachai M.Yu. (2004) Committees of systems of linear inequalities. *Automation and Remote Control*, vol. 65, no 2, pp. 193–203. DOI: 10.1023/B:AURC.0000014716.77510.61.
8. Mazurov V.I.D., Khachai M.Yu. (2013) Boosting and the polynomial approximability of the problem on a minimum affine separating committee. *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, vol. 19, no 2, pp. 231–236 (in Russian).
9. Mazurov V.I.D., Khachai M.Yu., Poberii M.I. (2008) Combinatorial optimization problems related to the committee polyhedral separability of finite sets. *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, vol. 14, no 2, pp. 89–102 (in Russian).
10. Mazurov V.I.D., Khachai M.Yu., Rybin A.I. (2002) Committee constructions for solving problems of selection, diagnostics, and prediction. *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, vol. 8, no 1, pp. 66–102 (in Russian).
11. Mazurov V.I.D., Khachai M.Yu. (2003) Committee constructions as generalization of solutions to controversial problems of operations research. *Discrete Analysis and Operations Research*, ser. 2, vol. 10, no 2, pp. 56–66 (in Russian).

12. Mazurov V.D., Khachai M.Yu. (2007) Parallel computations and committee constructions. *Automation and Remote Control*, vol. 65, no 2, pp. 193–203. DOI: 10.1134/S0005117907050165.
13. Mazurov V.D., Smirnov A.I. (2012) Interpretation of contradictory images by means of systems of linear inequalities. *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, vol. 18, no 3, pp. 144–154 (in Russian).
14. Khachai M.Yu. (2010) Computational complexity of recognition learning procedures in the class of piecewise-linear committee decision rules. *Automation and Remote Control*, vol. 71, pp. 528–539. DOI: 10.1134/S0005117910030136.
15. Khachai M.Yu. (1997) The existence of the majority committee. *Discrete Mathematics*, vol. 9, no 3, pp. 82–95 (in Russian).
16. Khachai M.Yu. (1997) Estimate of the number of members in the minimal committee of a system of linear inequalities. *Computational Mathematics and Mathematical Physics*, vol. 37, no 11, pp. 1399–1404 (in Russian).
17. Kuvshinov B.M., Shiryayev O.V. (2002) The method of committees in the problems of pattern recognition in the conditions of uncertainty of a priori information. *Bulletin of the South Ural State University, Series "Mathematics. Mechanics. Physics"*, vol. 2, no 3, pp. 34–43 (in Russian).
18. Kuvshinov B.M., Shiryayev O.V., Bogdanov D.V., Shaposhnik I.I., Shiryayev V.I. (2001) The system for classification of multiparametric objects for pattern recognition problems with inaccurate a priori information. *Information Technologies*, no 11, pp. 37–43 (in Russian).
19. Nikonov O.I., Chernavin F.P. (2014) Construction of rating groups of individual borrowers using the method of committees. *Money and Finance*, no 11, pp. 52–54 (in Russian).
20. Nikonov O.I., Chernavin F.P., Medvedeva M.A. (2015) Classification problems: the method of committees. Proceedings of the *XII International Scientific and Practical Conference on the Problems of Economic Development in the Modern World "Sustainable Development of Russian Regions: Economic Policy in the Conditions of External and Internal Challenges"*, Ekaterinburg, 17–18 April 2015, pp. 867–874 (in Russian).
21. Nikonov O.I., Medvedeva M.A., Chernavin F.P. (2015) Using the committee machine method to forecasting on the FOREX. Proceedings of the *IEEE 2015 Second International Conference on Mathematics and Computers in Sciences and in Industry (MCSI 2015)*, Sliema, Malta, 17–19 August 2015, pp. 240–243.
22. Chernavin F.P. (2018) Application of the method of committees for classification problems. Proceedings of the *XII International Conference "Russian regions in focus of change"*, Ekaterinburg, 16–18 November 2017. Ekaterinburg, UPI, part 1, pp. 437–447 (in Russian).
23. Kim J.O., Mueller C.W., Klecka W.R. (1989) *Factor, discriminant and cluster analysis*. Moscow: Finance and Statistics (in Russian).
24. Elder A. (2016) *Trading for a living: Psychology, trading tactics, money management*. Moscow: Alpina Publisher (in Russian).
25. Shiryayev V.I. (2007) *Neural network methods in the analysis of financial markets*. Moscow: KomKniga (in Russian).

About the author

Nikolay P. Chernavin

Junior Researcher, Institute of Economics, the Ural Branch of Russian Academy of Sciences,
29, Moskovskaya Street, Ekaterinburg 620014, Russia;

E-mail: ch_k@mail.ru

ORCID: 0000-0002-2093-9715