

[DOI: 10.17323/1998-0663.2019.4.73.86](https://doi.org/10.17323/1998-0663.2019.4.73.86)

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

Nikolay P. Chernavin 

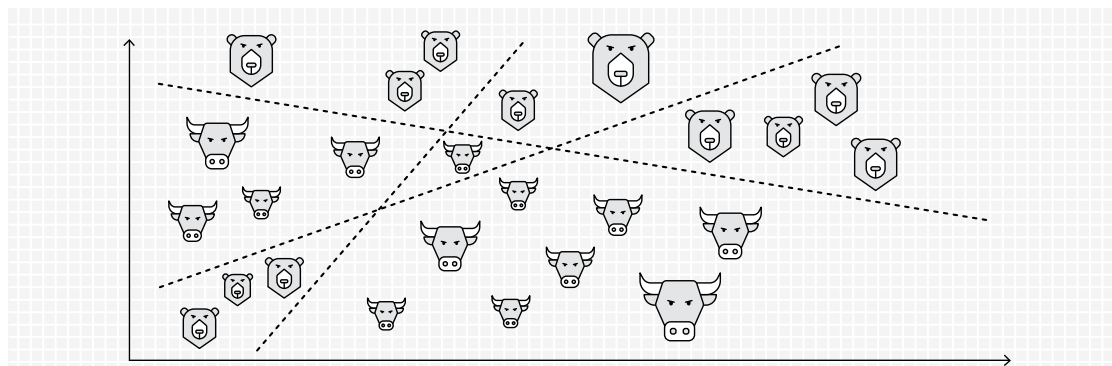
E-mail: ch_k@mail.ru

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.

Graphical abstract



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

Introduction

At present, methods of technical analysis can rightly be called the most popular analytical tools in the financial markets industry. Technical analysis is a set of methods for predicting the dynamics of stock quotes through study of internal market factors, such as changes in prices, trading volumes, open interest, etc. The popularity of technical analysis methods is largely explained by two points. First, simple analytical methods do not require a strong professional knowledge base. Second, the possibility of obtaining economic benefits from the results of technical analysis is available to almost any person who has the minimum capital to open a stock exchange account.

A general picture of the technical analysis methods including their classification by data processing complexity is presented in *Figure 1* (compiled by the author using [1, 2]).

In accordance with *Figure 1*, technical analysis can be distinguished by formal-analytical and visual-graphic methods. The formal-analytical methods study stock exchange market

data over the past period in the form of different indicators. On this basis, these methods predict growth or fall in stock prices. The visual-graphic methods are based on analysis through the study of graphical models built on stock exchange market data. This article explores the use of the committee method to evaluate the results of the technical analysis of stock prices.

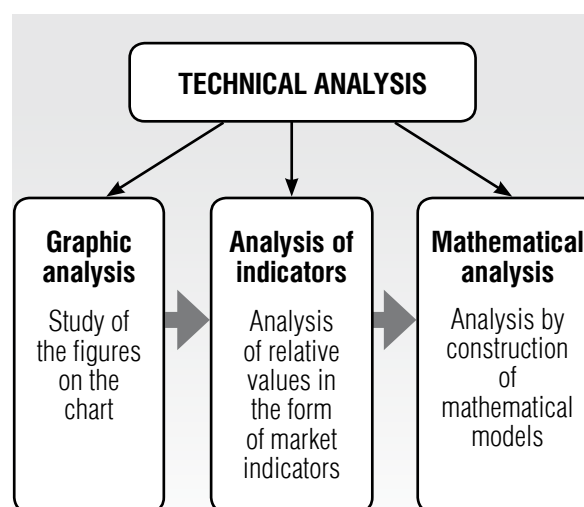


Fig. 1. Classification of technical analysis methods based on data processing complexity

The committee machine method is one of the machine learning algorithms of binary classification. The term “committee machine” appeared for the first time in 1965 in an article by Ablow and Kaylor [3]. Further significant theoretical and practical research on the committee machine method was conducted by Osborne [4], Takiyama [5], as well as in proceedings of the N.N. Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of Russian Academy of Sciences (IMM UB RAS), mainly by Mazurov and Khachai [6–16].

There is also interesting research into committee machine optimization by its confidence in the decision chosen that was conducted by Kuvshinov and Shiryaev [17–18]. Considerable reduction of the problem of committee machine construction to the task of linear partial-integer programming was studied by Chernavin and Nikonov [19–22].

1. Methodology of the research

Committee constructions are used to solve the problems of nonlinear discriminant analysis. Discriminant analysis is a group of machine learning methods that allow you to identify differences between groups and make it possible to classify objects according to the principle of maximum similarity. A necessary condition for applying discriminant analysis is the availability of a set of features (variables) [23], including a classifying (dependent) variable representing the class of the object, and discriminatory (independent) variables that reveal differences between objects of different classes.

The result of discriminant analysis is a discriminant function. The canonical discriminant function of the linear form has the following mathematical representation [23]:

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

where d – classifying variable;

I – a set of features;

J – a set of all the considered objects in the space of the i -th number of features;

β_0 – free coefficient ensuring the fulfillment of the required conditions;

β_i – a discriminant function coefficient for the i -th feature;

x_{ij} – a discriminatory variable for the i -th feature of the j -th object.

Committee machine constructions use several discriminatory functions called the committee machine members. Depending on a value of the classification variable, it is said that a member of the committee votes “for” or “against” a particular decision. The final decision is made relying on the decisions of all the committee members and their processing using the committee machine logic. There are three main committee machine logics: unanimity (CU), majority (CM) and seniority (CS) logics.

To clarify the difference between different logics of the committee machines, we consider graphical examples for the class of objects: “circles” and “stars”. These objects are characterized by two parameters along the axes X_1 and X_2 . Accordingly, the necessary conditions for discriminant analysis are fulfilled by the presence of a classifying variable (binary division into the classes: for example, “circles” = 0 and “stars” = 1), as well as the presence of discriminatory variables, expressed by the values of each object along the X_1 and X_2 axes. The results of discriminatory analysis are displayed in the form of lines representing individual members of the committee machine (linear discriminant functions), and the arrows for each line which indicate the direction of voting of the corresponding committee member.

In a case of the CU, an object belongs to a certain class only if all committee machine members vote for this class, otherwise it will belong to another class. A graphical example of the CU of three members is presented in *Figure 2*.

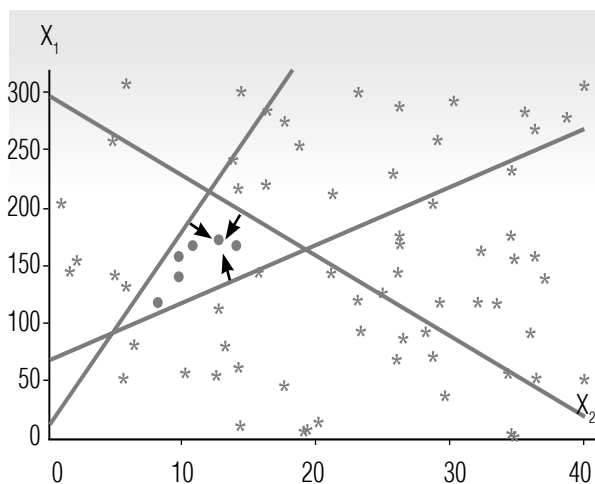


Fig. 2. An example of CU in a two-dimensional space

Figure 2 shows that the “circles” are in the area of the graphic for which all committee machine members voted unanimously. All observations in the remaining areas will contain “stars.”

In accordance with the names, CM is a committee machine where an object belongs to a certain class if a majority of committee machine members vote for this class; otherwise it belongs to another class. A graphical example of CM of three members is presented in Figure 3.

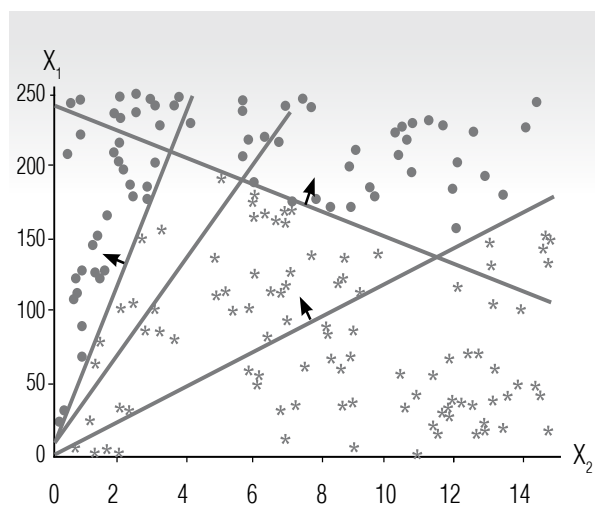


Fig. 3. An example of CM in the space of two features

Figure 3 shows that the “circles” are in the area of the graphic, for which two or more committee machine members voted. All observations in the remaining areas will contain “stars.”

The last logic requires different committee members to have certain weights reflecting their power in the voting process. As a result, an object belongs to a certain class if there are enough weights from the votes of the committee machine members for this class; otherwise it belongs to another class. A graphical example of CS of three members is presented in Figure 4 (in addition to the previous designations, the lines correspond to the weights of the committee members).

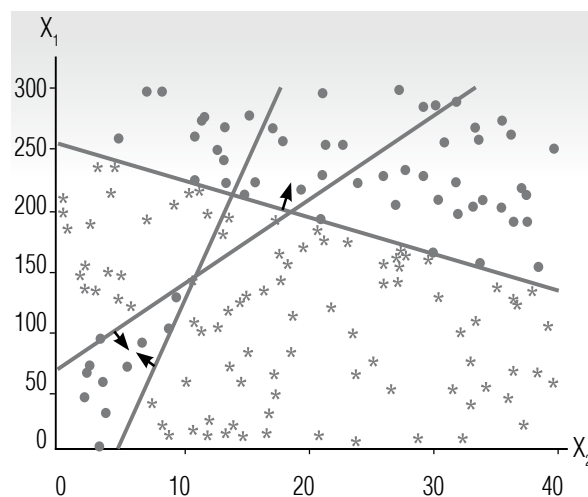


Fig. 4. An example of CS in the space of two features

Figure 4 shows that the “circles” are in the area of the graphic, which have a sum of vote weights more than three. All observations in the remaining areas will contain “stars.”

The examples presented reflect cases where the committee machine allows us to classify objects into classes precisely. However, in real life, as a rule, not all objects can be accurately attributed to a certain class. Therefore, the committee machine should classify the objects

in such a way that the quality of the division into classes would be best.

To illustrate this idea, let's suppose that there is some initial division of objects into classes of "circles" and "crosses." In Figure 5, a graphical example of such classification is presented.

Figure 5 shows the case of the CU committee machine of three members in the space of two features, where there are examples when objects are classified to the wrong class. For clarity, such objects are circled on the graph. The presence of the classification errors is a natural phenomenon which can be explained by many factors, such as incorrect initial markup of objects into classes, lack of initial information for more accurate classification, insufficient number of committee members, incorrect committee machine logic, etc.

Mathematically, a committee machine can be described as a linear partially integer programming model. Below an example of such a model is presented:

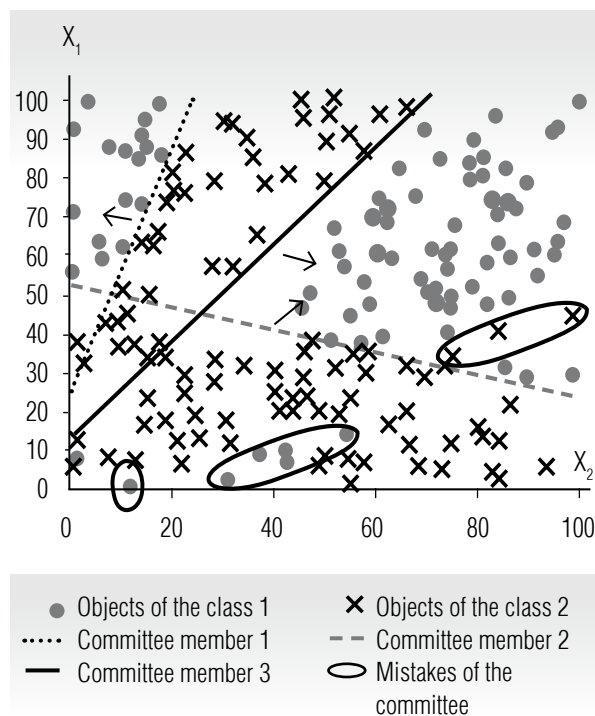


Fig. 5. An example of CM in the space of two features in a case where there are classification mistakes

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

where J_1 – a set for classification in the space of the i -th number of features belonging to class 1;

J_2 – a set for classification in the space of the i -th number of features belonging to class 2;

I – a set of observation parameters;

T – a set of the committee members;

i, j, t – indices of the corresponding sets;

x_{ij} – i -th parameter of j -th observation (discrimination variables);

β_i^t – a coefficient for the i -th parameter of t -th committee machine (decision variables);

β_0^t – an absolute term of the t -th committee machine (decision variable);

z_j^t – a Boolean variable to fixate vote direction of the t -th hyperplane for the j -th observation;

d_j – a Boolean variable to fixate errors of a committee machine classification for the j -th observation;

L – a number much greater by value than parameters of a model (introduction of such constant is necessary to make it possible for $L \cdot z_j^t$ and $L \cdot d_j$ to be used as the residuals);

ε – a small number close to 0 (introduction of such constant makes not strict inequalities equivalent to the strict inequalities ($>$ or $<$), so this way we exclude as a decision a case when all coefficients are equal to zero);

V^t – a weight of the t -th committee machine

member (constants which are calculated as a power of two);

φ – a Chebyshev approximation (minimax) variable. Using this value for minimization allows us to search for the optimal solution in such a way that the share of classification errors for each class of objects will be minimized at the same time;

sv – a sum of all committee machine weights ($\sum_{i \in T} V^i$);

m – a minority of a committee machine (a variable in the range $0 \leq m \leq sv - 1$).

This model (2) can be used for each of the committee machine logics. If $V = 1$, then depending on m there may be constructed CU ($m = 0$ or $m = sv - 1$) or CM ($0 < m < sv - 1$). If $V = 2'$, then depending on m there may be constructed CU ($m = 0$ or $m = sv - 1$) or CS ($0 < m < sv - 1$).

2. Description of assigned task

The main problem of the technical analysis methods is the presence of a multitude of false signals, the screening of which requires experience and intuition. Accordingly, we will try to conduct such screening with implementation of the committee machines. To accomplish this task, 10 market indicators were selected (Table 1).

The basis for the selection of indicators is a set of indicators described by Elder in his Triple Choice system [24]. At the same time, there is no preservation of the philosophy of the Triple Choice system associated with triple sifting of a deal before making a decision, since it is assumed that the committee method based on the composition of the indicators is able to reveal a decisive rule that has greater potential profitability.

Let's consider in more details the calculation of the indicators presented in Table 1.

To calculate indicators P_1 and P_2 , it is necessary to calculate MACD histogram in the following way:

1. Calculate 12-day EMA of daily close prices;

2. Calculate 26-day EMA of daily close prices;

3. Calculate a “fast” line, also known as MACD line (26-day EMA of daily close prices calculated on 12-day EMA of daily close prices);

4. Calculate a “slow” line, also known as signal line (9-day EMA calculated on values of a “fast” line);

5. Calculate an MACD histogram as the difference between “fast” MACD line (step 3) and “slow” signal line (step 4).

Indicator P_1 shows how many days the MACD histogram grows or falls. Moreover, if the indicator falls, then the number of days is taken with the “minus” sign.

Indicator P_2 shows how many days the MACD histogram is more or less than zero. In this case, if the indicator is less than zero, then the number of days is taken with a “minus” sign.

The SSO in the indicators P_3 and P_4 is used to define the overbought and oversold market states. It is calculated as follows:

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

where r – time period of SSO (in this research $r = 5$);

s – a time interval of the averaging (in this research $s = 3$);

Max_r – a maximum in a time interval r ;

Min_r – a minimum in a time interval r ;

C_c – a closing price on the moment of SSO calculation;

$Mean_n$ – function to calculate a mean value in a period of n days (in this research $n = 5$).

There are different opinions regarding the overbought and oversold market states. Usually, it is considered that the market is over-

Table 1.

Analyzed indicators

| Indicator | Description | Value range |
|-----------|---|--------------|
| P_1 | How many days in a row does a moving average convergence divergence histogram (MACD) decrease or increase | Integer |
| P_2 | How many days in a row is the MACD ¹ histogram more or less than zero | Integer |
| P_3 | Signal from a slow stochastic oscillator (SSO ²) | from 0 to +1 |
| P_4 | How many days in a row does SSO give a strong signal | Integer |
| P_5 | Signal from a relative strength index (RSI ³) | from 0 to +1 |
| P_6 | How many days in a row does RSI give the same signal | Integer |
| P_7 | Trend calculated by dynamics of an exponential moving average (EMA) from the price maximums | -1, 0, +1 |
| P_8 | How many days in a row does the Chaikin indicator equal more or less than zero | Integer |
| P_9 | How many days in a row does the Chaikin indicator ⁴ decrease or increase | Integer |
| P_{10} | Is the price maximum higher or is the minimum lower compared to the previous day. Otherwise it equals zero. | -1, 0, +1 |

sold when $SSO \leq 0.2$, and is overbought, if $SSO \geq 0.8$. Consequently, there is no meaning to use this indicator directly in a model, because it would not fully reflect the logic of oversold and overbought. Thus, as a model parameter this indicator needs to be transformed by the following formula:

$$P_3 = (2 \cdot (SSO - 0.5))^3 \quad (4)$$

Indicator P_4 is used for a more detailed description of P_3 . It indicates how many days SSO is in the overbought area (taken with a “minus” sign) or oversold (taken with a “plus” sign). This indicator is used for approximate digitization of an estimate of an SSO chart by the eyes of a trader, when not only current values are estimated, but also previous dynamics.

Indicator P_5 is based on the *RSI* indicator, which is calculated as follows:

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

where *RS* – a ratio of the average value of increasing closing prices for seven days to the average value of lowering closing prices for seven days.

In this research the market is considered oversold when $RSI \leq 0.3$, and is overbought, if $RSI \geq 0.7$. This indicator is transformed the same way as P_3 by the following formula:

$$P_5 = (2 \cdot (RSI - 0.5))^3 \quad (6)$$

Indicator P_6 additionally analyzes *RSI* by showing how many days in a row the *RSI* sends

¹ For more details on the calculation and application, see [24]

² For more details on the calculation and application, see [24]

³ For more details on the calculation and application, see [24]

⁴ The Chaikin indicator shows what is happening with the trend based on the ratio of the difference between closing and opening prices to the difference between the maximum and minimum prices multiplied by the volume of trades. For more details on the calculation and application, see <https://www.opentrainer.ru/articles/ostsillyator-chaykina-chaikin-oscillator/> (access date: 01 September 2019)

a signal of the same type, and this indicator can be either positive or negative depending on the type of signal (that is, if the signal is less than zero, then the number of days is taken with a “minus” sign).

The daily trend indicator was chosen as indicator P_7 . Classically, the trend is determined by the slope of the line drawn through the local price highs. However, it is impossible to draw an ideal line through all local highs in the selected study interval; therefore, as a rule, the line is drawn empirically, so as to pass through the greatest number of price maxima with the minimum deviation from them. A trend has three main states: rising, falling and no trend (flat). In this study, the trend is determined by the change in the 22-day exponential moving average (EMA) from a maximum daily price. If the EMA is growing for the second day in a row, then this indicator equals one until it begins to fall. Accordingly, and vice versa, if the EMA falls for the second day in a row, then this indicator equals -1 until it starts to grow. If the EMA grew or fell within just 1 day, then the indicator is zero.

Indicators P_8 and P_9 reflect the signals given by the Chaikin indicator that is calculated as follows:

$$CHI = EMA_3(A/D) - EMA_5(A/D), \quad (7)$$

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

where A/D – accumulation/distribution indicator;

C_c – closing price in a day of calculation;

C_o – opening price in a day of calculation;

Max – maximum price in a day of calculation;

Min – minimum price in a day of calculation;

V – trade volume in a day of calculation;

EMA_3, EMA_5 – operations of EMA calcula-

tion with a time period of 3 and 5 days respectively.

Indicator P_8 shows how many days in a row the Chaikin indicator is more or less than zero. Moreover, if the indicator is less than zero, then the number of days is taken with a “minus” sign.

Indicator P_9 shows how many days in a row the Chaikin indicator is rising or falling. Moreover, if the indicator falls, then the number of days is taken with the “minus” sign.

The last indicator P_{10} shows whether the maximum price was higher than the previous day (equals 1) or the minimum lower than the previous day (equals -1). Otherwise, if prices have not gone beyond the price range of the previous day, then the parameter equals zero. The market also has a case when during the day the maximum was simultaneously higher than the previous day, and the minimum was lower than the previous day. Mostly they are linked with unpredictable events like news and are quite rare. Considering this fact, they were excluded from a list of the samples to reduce information noise from the intersection of the samples differing by the trading environment.

Examples of indicators for PJSC Sberbank shares are presented in *Table 2*.

3. Trading strategy

To build a committee machine model, one needs to form a trading strategy. The basic parts of a trading strategy are:

1. Start-up capital, that is, a fixed amount within which you can open a position⁵;
2. Stop-loss (a preset value of maximum loss at which the current open position will be closed⁶).

Let's determine that a trading signal comes from P_3 and P_5 . If both parameters are less than

⁵ To open a position means to initially purchase or sell a certain volume of a financial instrument

⁶ To close a trading position means to complete a reverse trade operation in relation to opening a position. For example, if a position was opened as part of a purchase transaction of a certain volume of a financial instrument, then the reverse transaction will be expressed by the subsequent sale of the acquired volume of the financial instrument either in full (complete closing of the position) or some part of it (partial closing of the position)

Table 2.

Examples of indicators for PJSC Sberbank shares

| Date | Indicators | | | | | | | | | |
|------------|------------|-------|--------|-------|--------|-------|-------|-------|-------|----------|
| | 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 |

0.25, then this is a signal to buy (open a long position), and if they are more than 0.75, then a signal to sell (open a short position) is given. According to the strategy, the signal is formed at the end of the trading day, but we will assume that the position with the corresponding signal opens in the last minutes. In this case, we determine that the position is closed if a signal appears with a reverse direction, or losses are greater than a chosen stop-loss value.

There are various methods for choosing a stop loss value. In the framework of this study, we used the calculation formula with reference to an average price volatility⁷ of seven days, which will limit losses for each observation. Accordingly, a stop-loss will be calculated using the following formula:

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

where \overline{V}_7 – an average price volatility of seven days;

s – a coefficient that shows how much of the average daily volatility is allowed to risk in the framework of the trading strategy.

The described strategy can be schematically presented as presented in *Figure 6*.

The trading strategy shown in *Figure 6* is built by using two indicators of technical analysis – *SSO* and *RSI*. Trades in a trading strategy are formed only upon receipt of a same signal to buy or sell from both indicators. However, it must be understood that even after receiving strong signals from technical indicators, it is possible to experience losses from the trading activities.

4. A problem of machine learning

To reduce potential losses, a trading strategy must be presented as a discriminant analysis problem, in which the following classes of objects will be recognized:

1. Class 1 is a set of the positions that were closed with a profit and the profit was greater than the maximum loss⁸ during the holding of the position⁹;

2. Class 2 is a set of the positions that were closed with a loss, or a profit that was less than the maximum loss while the position was held.

After determining the strategy, it is necessary to choose a financial instrument for which the decision rule will be built. As part of the research,

⁷ Daily volatility is a value that shows by how many percent the daily maximum is higher than the daily minimum

⁸ When a position is opened on the stock exchange, it is constantly reevaluated at the current prices. Accordingly, the maximum loss on a position is the maximum amount of decrease in the value of a position relative to its value at the opening

⁹ The period of position holding is a period from the time of an initial purchase or sale of a certain volume of a financial instrument to the time of a trade operation that is inverse to the first deal. More information on the concepts of the position opening and closing can be found at: https://www.metatrader5.com/ru/mobile-trading/android/help/trade/positions_manage/open_positions (access date: 01.09.2019)

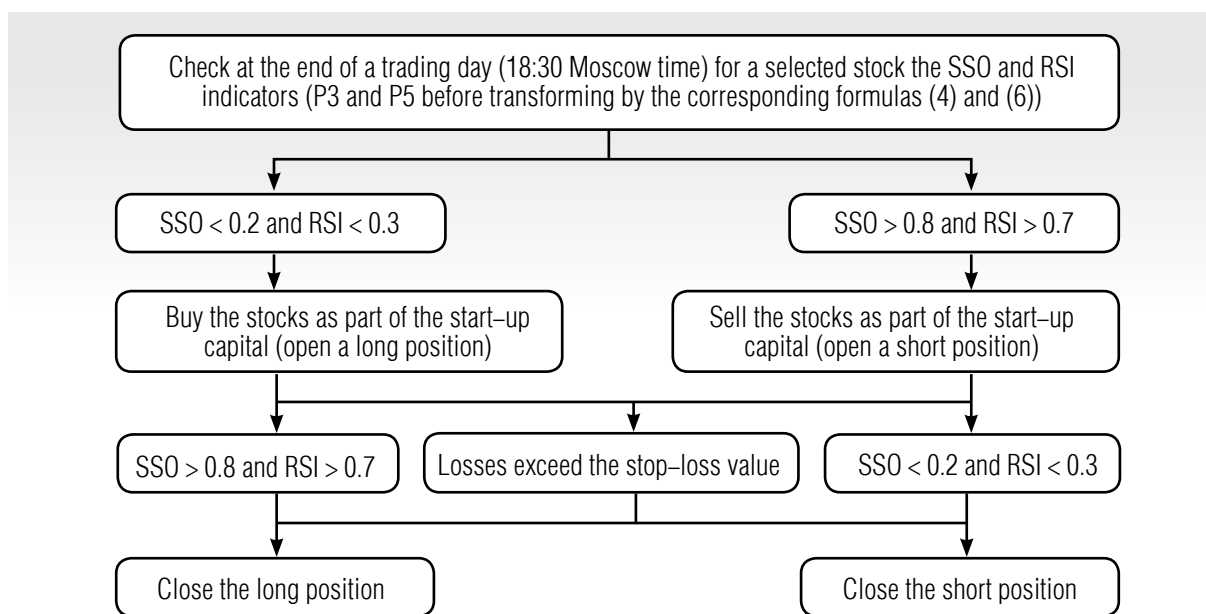


Fig. 6. Trading strategy

Sberbank shares were chosen as such a tool given that they are the most liquid asset on the Moscow Exchange’s share market. For the analysis, daily quotes on Sberbank shares were downloaded from the Finam¹⁰ website for the period from 2010 to December 2018.

All observations were divided into training (a set with observations, on which a model trains) and test (a set with observations, on which a model does not train) sets, as shown in Table 3. This distinction is necessary to be able to check the consistency of the model on these two sets, which reduces the risk of retraining the model [25].

Table 3.

Description of the training and test sets

| Set | Period | J_1 | J_2 | J |
|----------|-----------|-------|-------|-----|
| Training | 2010–2017 | 332 | 116 | 448 |
| Test | 2018 | 37 | 15 | 52 |
| Total | 2010–2018 | 369 | 131 | 500 |

Table 3 provides a short description of the training and test sets for a trading strategy described in Figure 6¹¹. According to it, the number of samples in J_1 is about 2.5–3 times greater than in J_2 , and the test set equals 11.6% size of the training set. With the classification task and sets defined let’s start our calculation of the mathematical model.

5. Classification results

For comparison, we constructed various committees from 3 to 7 committee machine members. Classification quality was evaluated using the F-measure metric, calculated by the following formula:

$$SL = \bar{V}_7 \cdot s, \tag{9}$$

where A – accuracy (the share of objects assigned by a committee machine to the desired class and at the same time really being objects of this class);

¹⁰ Finam – Moscow exchange stocks – Sberbank: <https://www.finam.ru/profile/moex-akcii/sberbank/export/> (access date: 01 September 2019)

¹¹ Samples were built for a trading strategy with a stop-loss selected by formula (8) with a coefficient $s = 2.3$

C – completeness (the share of objects of the desired class that a committee machine was able to recognize).

Studying the results among CU, the decisive rule for the case of five members with the result of 86.4% for J_1 and 68.9% for J_2 has shown the best separating ability on the training set. For the CS, the best case consists of six members with the result 89.1% for J_1 and 73.6% for J_2 . Here are not presented results for the CM, since the mathematical model has not found an appropriate quality solution for the majority logic¹².

To evaluate the quality of the model, the most important thing is its results on the test set. This is linked to the fact that the results obtained on the test set have no risk of over-training and subject to the condition that the size of a test set is big enough it can objectively assess the predictive ability of a model. Accordingly, the F-score metric results of the model on the test set are presented in *Figure 7*.

Studying the results on the test set in *Figure 7* it can be seen that the previously reviewed CU of five members and the CS of six members also have the best result. Thus, we can state that the training and test sets really have strong links

between each other, and that proves the fact that our problem can be studied by construction of committee machines. By further comparing these two committee machine decisions by the F-score metric, it can be seen that CU results of five members are better than for CS of six members.

So in this model we studied minimization of a number of errors in predicting the correct or incorrect signals from technical indicators. However, in the framework of this model, the value of the potential profits or losses, depending on the accepted decisions was not taken into account. Since the ultimate goal of modeling is to make a profit, introducing this evaluation criterion into the model can improve the final result. To do this, in the model for each observation of the training and test samples, we calculate the value of potential gains or losses by the following formula:

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

where C_0 – the opening price of a position (initial purchase/sale price of a financial instrument, as a result of which a trading position was formed);

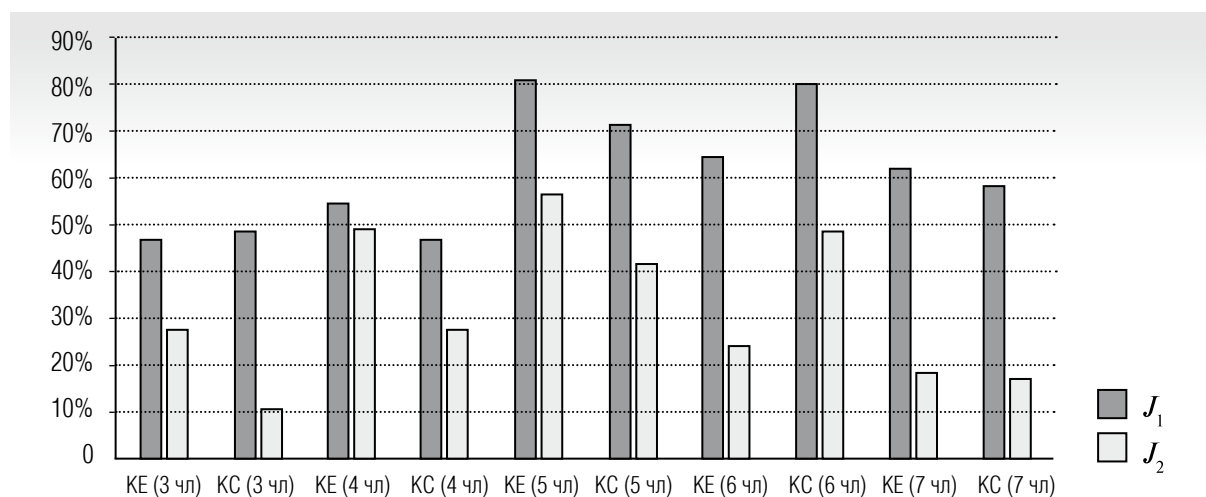


Fig. 7. F-score results on the test set

¹² While searching for a solution in the framework of the model (2) with $V = 1$, the solution with the lowest value of φ was found at $m = 0$, which corresponds to CU

C_c – the closing price of a position (price at which a previously opened position was closed on a stock exchange, as a result of which profits or losses were recorded on the position);

D – position direction (1 – if a financial instrument was bought; –1 – if a financial instrument was sold).

If the value obtained by the formula (10) is greater than zero, then a position gives a profit, otherwise there is a loss by position. Accordingly, the previous model (2) may be changed by addition of weights to each observation with addition of corresponding change in restrictions $\sum_{j \in J_1} d_j$ and $\sum_{j \in J_2} d_j$ in the model (2):

$$\begin{aligned} \sum_{j \in J_1} d_j w_j &\leq \sum_{j \in J_1} w_j \varphi, \\ \sum_{j \in J_2} d_j w_j &\leq \sum_{j \in J_2} w_j \varphi, \end{aligned} \tag{10}$$

where w_j – weight of the j -th observation.

Let’s also additionally reduce the coefficient s in the stop loss calculation formula (8) to 1.5 in order to reduce the level of losses for each position formed by the trading strategy. With addition of this stop-loss parameter we changed the number of observations for each class on the training set, because 16 observations from the J_2 were closed by the stop-loss rule. In other words, they have been classified as the J_1 . Accordingly,

on the new training set we constructed various committees taking into account the conditions (3) and the best results were obtained by CM of seven members with the F-score metric results 92.1% for J_1 and 76% for J_2 on the training set, and 83.1% for J_1 and 51.9% for J_2 on the test set. A corresponding decision rule for CM of seven members is presented in the *Table 4*.

Let’s compare the previous best decision rule with the decision rule from *Table 4* by the value of potential return. To simplify calculations, let’s assume that each deal is formed with the same startint capital. Consequently, *Table 5* presents the results of comparison by years of CM of seven members and CU of five members taking into account the chosen stop-losses.

Based on the profitability results in *Table 5*, we see that for a CM of seven members, income from correctly recognized observations from J_2 and losses from wrongly recognized observations from J_1 have a much smaller variation of values than for a CU of five members. At the same time, losses on wrongly recognized observations from J_1 for a CM of seven members are two times less than for a CU of five members, while incomes from correctly recognized observations from J_2 are only 9.6% less. Moreover, for the considered CM of seven members there is no such annual period when the losses were greater than the income, whereas for the

Table 4.

Decision rule for the CM of seven members ($m = 2$)

| Members | Coefficients | | | | | | | | | | |
|---------|--------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---------------|
| | 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 |

Table 5.

Profit and loss evaluation of decision rules

| Committee type | Class | % of profit or losses by year | | | | | | | | | |
|-----------------|-------|-------------------------------|--------|------|-------|-------|-------|--------|-------|-------|-------|
| | | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | Total |
| CM of 7 members | 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 |
| CU of 5 members | 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 |

case of the CU of five members in 2017 there was an excess of losses over the income. Thus, we can say that optimization with minimization of losses in a case of work with technical indicators provide better results than just minimization by the number of classification errors.

Conclusion

As the result of the study, we have shown what the committee machine method is and how various committee machine logics can be calculated within the framework of a single mathematical model of partially integer programming. Using a practical example, we consider the use of the committee method to increase the accuracy of forecasting prices on the stock exchange market while performing technical analysis. The use of committee machine constructions made it possible to derive objective decision rules that determine when to follow recommendations based on market indicators and when to refrain

from active trading on the stock exchange market. Moreover, within the framework of committee machine constructions, as models of partially integer programming, complex optimization criteria can be formed which take into account the problem of maximizing profits and minimizing losses. A model with this formulation of the problem allows us to get more stable results with a potentially higher total profit, compared with models based on a simple minimization of the number of classification errors. Thus, committee machine constructions can be used on the financial markets both to form new trading strategies and to increase the profitability of existing strategies. ■

Acknowledgements

The article was prepared as part of the state assignment for the Institute of Economics, the Ural Branch of Russian Academy of Sciences for 2019.

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. (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.I.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.I.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., Shiryaev 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., Shiryaev O.V., Bogdanov D.V., Shaposhnik I.I., Shiryaev 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. Shiryaev 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