# Committees: History and Applications in Machine Learning

Vladimir D. Mazurov[1,2](✉) and Ekaterina Yu. Polyakova[2]

[1] Krasovsky Institute of Mathematics and Mechanics, Ekaterinburg, Russia
vldmazurov@gmail.com
[2] Ural Federal University, Ekaterinburg, Russia
ekaterina.y.polyakova@gmail.com

**Abstract.** The article outlines a brief history and applications of the committee theory. The use of committees in the problems of recognition and optimization is discussed. The application of the committee structures, ambiguous interpretation of non-formalized and contradictory data are given. The ways of rational regard on environmental factors in the context of a lack of resources are considered. The question of the numerical finding of committee structures is discussed, and these results are directly related to the theory of voting. The class of non-classical logics also contains MK-logic (Mazurov, Khachay). This section of non-classical logic includes the works by N. A. Vasiliev, L. Wittgenstein, J. Lukashevich, and Latin American mathematicians having a wrong term in their titles parainconsistent logic. One of the important results achieved by M. Yu. Khachay: For arbitrary positive integers $q$ and $k$, $k < q$, the minimum estimate of the subsystem power is given that is resolvable by a committee of $k$-elements for the inconsistent system having a committee of $q$-elements. Further the history of this field will be mentioned.

**Keywords:** Committee · Existence · Linear inequalities · Affine case

## 1 Introduction

In 1965, S.B. Stechkin and I.I. Eremin set a problem of substantiating the necessary and sufficient condition for the existence of a linear inequalities system committee. We solved this problem in 1966, and further ways were opened to continue the research and applications. First, we obtained the results on the conditions for the existence of various modifications of the committees and their applications in economics, engineering, medicine and biology (see. e.g. [17–19]). Further, some of them were extended by Khachay [11,13], Rybin [9,20,24], Kobylkin [15], Gainanov [5] who obtained a number of valuable results in the field of the algorithmic analysis of the committee constructions.

### 1.1 Some Applications

As the examples of applications the following ones can be referred to:


I. Bykadorov et al. (Eds.): MOTOR 2019, CCIS 1090, pp. 3–16, 2019.
https://doi.org/10.1007/978-3-030-33394-2_1

- Issues of substantiation and application of numerous specific problems of mathematical programming and recognition solved by us with post-optimal solutions analysis;
- tasks of researching operations with the Ural theme (geology, mining, metallurgy);
- algebraic factor analysis;
- factors and their names;
- design issues;
- mathematical structuralism;
- morphology and structural analysis;
- applied structuralism.

### 1.2 Non-classical Logics

MK-logic (logic of Mazurov-Khachay);
Vasiliev logic;
Lukashevich logic;
parainconsistent logic;
post-optimization analysis and MK-logic.

By its very nature, recognition is associated with mathematical epistemology, mathematical theory of neural networks. Within these disciplines, we voluntarily or involuntarily approach to the questions of the essence of human intelligence, human mentality, formal or informal logic. What is absolutely certain is the fact that we at least imitate human mentality as much as we can. In this case, we use the following scheme of operation of neural networks:

$$? \to S \to A \to R,$$

where "?" denotes the reality that is unknown for us. This reality affects the $S$-layer—the network sensor unit, and we obtain the pixel array $x$. It arrives at the input of the block $A$—the block of associative neurons. The result of the work of these neurons is received on the $R$-block – the block of resulting neurons. If the experimental data diverges from the ones constructed according to the $S \to A \to R \to x$ scheme, then a network correction is done. A. Novikov investigated the linear correction method, having proved that if there is a solution $x$, then the result is obtained through a finite number of corrections. Researchers of neural networks unintentionally seek to interpret the work of the network as a learning process of artificial intelligence. It is characteristic that Rosenblatt entitled his book Principles of Neurodynamics [23], Nilson—Learning Machines [21], and Vapnik in his book "Statistical Learning Theory" [25] entitled the philosophical section "Some General Remarks".

Vapnik poses a question why the pattern recognition problem arouses such a great interest among scientists of various specialties. It seems that the answer to this question was obtained both in the works of Zhuravlev [27,28] and in the seminal works of Vapnik (see, e.g. [26]).

The third direction, close to the theory of Yu.I. Zhuravlev, is the analysis of collective decisions—the method of committees.

## 2 Committee Solutions: Basic Concepts

Let we be given by a ground set $X$ and a family of its non-empty subsets $D_1, D_2, \ldots, D_m$. Let us consider the system of abstract inclusions

$$x \in D_j \quad (j \in [m]), \tag{1}$$

where $[m]$ is the integer segment $\{1, 2, \ldots, m\}$. If $\bigcap_{j=1}^{m} D_j = \varnothing$, system (1) is known as *infeasible*. Any non-empty subset $L \subset [m]$ induces a *subsystem* of system (1). Without loss of generality, we do not distinguish a subset $L$ and the appropriate subsystem. If $D(L) = \bigcap_{j \in L} D_j \neq \varnothing$, the subsystem is known as feasible. Any subsystem maximal by inclusion of system (1) is known as its *maximal feasible subsystem* or m.f.s.

**Definition 1.** *A finite sequence $Q = (x^1, \ldots, x^q)$, $x^i \in X$, such that, for any $j \in [m]$,*

$$|\{i \colon x^i \in D_j\}| > q/2$$

*is known as a committee generalized solution of system* (1)*.*

The number $q$ is called *a length* of the committee $Q$, and we state that system (1) has a committee solution (or is solvable by a committee) of length $q$. A committee of minimal length $q$ (for a given system (1)) is known as its *minimum* committee solution or just a minimum committee.

We can easily represent the set of all committee solutions of the system (1) as follows. Introduce the vector-function

$$\varphi \colon X \to \{-1, 1\}^m, \quad \text{where } \varphi_j(x) = \begin{cases} 1, \text{ if } x \in D_j, \\ -1, \text{ otherwise.} \end{cases}$$

By its construction, the image $\varphi(X)$ is a finite set. Let $\varphi(X) = \{\varphi^1, \ldots, \varphi^s\}$. Without loss of generality, we can suppose that, the vectors $\varphi^i$ are incomparable, i.e., for any $i_1$ and $i_2$, the equation $\varphi^{i_1} \geq \varphi^{i_2}$ implies $i_1 = i_2$. This implies that any $\varphi^i$ is a characteristic vector of some m.f.s. of system (1).

According to Definition 1, a finite sequence $Q$ is a committee solution of (1) if and only if, by some permutation, $Q$ can be represented in the form

$$(\underbrace{y^{1,1}, \ldots, y^{1,z_1}}_{z_1}, \ldots, \underbrace{y^{s,1}, \ldots, y^{s,z_s}}_{z_s}), \tag{2}$$

where $\varphi(y^{i,l}) = \varphi^i$ and $z_1, \ldots, z_s$ are nonnegative integers, such that

$$\sum_{i=1}^{s} z_i \varphi^i \geq e = [1, 1, \ldots, 1]^T.$$

It can be easily verified that a sequence $Q$ is a minimum committee of system (1) if and only if the vector $z$ is an optimal solution in the following integer linear program

$$\min \left\{ \sum_{i=1}^{s} z_i \colon \sum_{i=1}^{s} z_i \varphi^i \geq e, \ z \in \mathbb{Z}_+^s \right\}.$$

Evidently, if a system (1) is feasible, its minimum committees are of length 1 and coincide with regular solutions. In other circumstances, any minimum committee has more than 1 entry.

Existence conditions for committee solutions of system (1) can be easily represented in terms of graph and hypergraph of its m.f.s.

**Definition 2.** *A finite graph $G = (V, E)$ is known as the m.f.s. graph of system (1) if its nodeset $V$ consists of index sets $J_1, J_2, \ldots, J_s$ of the system, and*

$$\{J_i, J_j\} \in E \iff J_i \cup J_j = [m].$$

**Assertion 1.** *Let $s$ be a natural number, $J_1, \ldots, J_{2s-1}$ be a cycle in the m.f.s. graph of system (1), and $x^i \in D(J_i)$. Then, the sequence $Q = (x^1, \ldots, x^{2s-1})$ is a committee solution of system (1).*

To obtain necessary and sufficient conditions we need to introduce a more general notion (see, e.g. [10]).

**Definition 3.** *A finite hypergraph $G = (V, E)$, whose nodeset $V$ coincides with the family $\{J_1, \ldots, J_s\}$ of index sets of the m.f.s. of system (1) such that*

$$\{J_{i_1}, \ldots, J_{i_t}\} \in E \iff \bigcup_{k=1}^{t} = [m],$$

*is known as a m.f.s. hypergraph of system (1).*

**Theorem 1.** *Let $\Gamma = (V\Gamma, E\Gamma)$ be a non-empty*[1] *finite hypergraph without multiple edges. $\Gamma$ is isomorphic to a m.f.c. hypergraph $G = (V, E)$ of an inclusions system*

$$x \in D_j(\Gamma) \quad (j \in [m]) \tag{3}$$

*for some $m = m(\Gamma)$ if and only if $\Gamma$ satisfies the following conditions*

$$\textit{if} \ \ |V\Gamma| > 1, \ \textit{then } E\Gamma \textit{ has no loops} \tag{4}$$

$$(u \in E\Gamma, u \subset w) \Rightarrow w \in E\Gamma. \tag{5}$$

**Definition 4.** *For a m.f.s. hypergraph $\Gamma = (V\Gamma, E\Gamma)$ and natural numbers $\sigma$ and $\tau$, a finite sequence of nodes $S = (v_{i_1}, \ldots, v_{i_{\sigma+1}})$ is known as $(\sigma, \tau)$-simplex in the hypergraph $\Gamma$, if the following inclusion*

$$\{v_{i_j} : j \in L\} \in E\Gamma$$

*is valid for any $L \subset [\sigma + 1]$, $|L| = \tau + 1$.*

The concept of $(\sigma, \tau)$-simplex takes its origin from the geometric reasonings. It can be easily verified that $(2, 1)$-simplex induces a triangle (a cycle of length 3) in the hypergraph $\Gamma$.

**Theorem 2.** *System (1) has a committee generalized solution of length $q$ if and only if its m.f.s. hypergraph has a $(\sigma, \tau)$-simplex for $\sigma = q-1$ and $\tau = \lfloor (q-1)/2 \rfloor$.*

[1] i.e. $E\Gamma \neq \varnothing$.

## 3 Game Theoretic Conditions of Committee Existence

In this section, we set forth a series of necessary existence conditions for committee generalized solutions in terms of optimal strategies for the appropriate antagonistic games between *the Nature* and *the Researcher*. For any natural numbers $q$, $k < q$ and any system of abstract inclusions which has a generalized committee solution of length $q$, we set forth an attainable relative lower cardinality bound for the maximum subsystem having a committee solution of length $k$. To obtain the result, we calculate an upper value of the corresponding zero-sum two-player game.

For any setting of the game under discussion, the first player, we will call him or her *Researcher*, tries to choose a finite sequence of length $k$ attempting to enlarge the subsystem resolved by the sequence as a generalized committee solution. The second player, *Nature*, tries to stop these attempts proposing infeasible systems, which can hardly be solved with committees of length less than $q$.

We demonstrate that, almost always, this game does not have any value when being set in pure strategies, but has a value in mixed ones. Further, we research the asymptotic behavior of our bound in the case of $k = q - h$ for any fixed $h$ and $q \to \infty$. Here, we mainly follow the papers [11,12].

### 3.1 Problem Statement

By some non-empty ground let we be given a set $X$ and a family of its subsets $D_1, D_2, \ldots, D_m$ defining the following abstract system of inclusions

$$x \in D_j \quad (j \in [m]), \tag{6}$$

and natural numbers $q$ and $k < q$. We suppose that system (6) has a generalized committee solution of length $q$ and is meant to provide a lower cardinality bound for a maximum subsystem (of system (6)) satisfied by a committee of length $k$.

To obtain the bound, we consider an arbitrary committee solution $Q = (x^1, \ldots, x^q)$ of system (6). In the same manner as in Sect. 2, we assign to the committee $Q$ an incidence $m \times q$-matrix $A = A(Q)$ with entries

$$a_{ji} = \begin{cases} 1, \text{ if } x^i \in D_j \\ -1, \text{ otherwise.} \end{cases}$$

Denote by $a_j$ the $j$-th row of the matrix $A$. By construction, for any $a_j$,

$$\sum_{i=1}^{q} a_{ji} \geq 1. \tag{7}$$

Subsequently, we denote the set of $m \times q$-matrices $A$ satisfying equation (7) by $M(q)$.

To any subset $I \subset [q]$ of cardinality $k$, we assign

- the characteristic vector $\tau = \tau(I) = \sum_{e \in I} e_i^q$, where $e_i^q$ is the $i$-th orth of the $q$-dimensional Euclidean space $E_q$;
- the subcommittee $Q(I) = (x^i \colon i \in I)$ of length $k$ that is defined by the subset $I$;
- maximal subsystem of system (6) that is resolved by the subcommittee $Q(ISuffi)$; without loss of generality, subsequently, we do not distinguish by construction this subsystem and its index set $J(I)$, which satisfies the equation $J(I) = \{j \colon (a_j, \tau(I)) \geq 1\}$;
- the relative cardinality of the subsystem $J(I)$

$$\delta_{q,k}(I, A) = \frac{|J(I)|}{m}.$$

To any matrix $A \in M(q)$, we assign the number

$$\delta_{q,k}(A) = \max\{\delta_{q,k}(I, A) \colon I \subset [q], |I| = k\}. \tag{8}$$

Let us consider the following zero-sum two-player game $\Gamma = (X, Y, K)$, where the strategy sets of the first and second players are

$$X = \{I \subset [q] \colon |I| = k\} \text{ and } Y = M(q)$$

respectively and the payoff function is $K(I, A) = \delta_{q.k}(I, A)$. To answer the principal question of this Section, we need to calculate an upper value

$$\delta_{q,k} = \min_{A \in M(q)} \delta_{q,k}(A) = \min_{A \in M(q)} \max_{J \subset [q], |I| = k} \delta_{q,k}(I, A)$$

of this game. For the ensuing constructions, we need the following standard notation:

| | |
|---|---|
| $\binom{n}{i} = \frac{n!}{i!(n-i)!}$ | the binomial coefficient |
| $b(i; n, p) = \binom{n}{i} p^i (1-p)^{n-i}$ | the binomial distribution mass function, i.e. probability of $i$ successes in $n$ trials |
| $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ | the standard Gaussian density |
| $\lfloor x \rfloor$ and $\lceil x \rceil$ | the floor and ceiling functions of real argument. |

Also, we call real-valued sequences $\{\xi_n\}$ and $\eta_n$ *asymptotically equivalent* and use the notation $\xi_n \sim \eta_n$, if

$$\lim_{n \to \infty} \frac{\xi_n}{\eta_n} = 1.$$

Let $s = \lceil \frac{q+1}{2} \rceil$ and $t = \lceil \frac{k+1}{2} \rceil$. Similarly to $\tau(I)$, to any subset $S \subset [q]$ of size $s$, we assign the characteristic vector $\sigma(S) = \sum_{i \in S} e_i^q$. Without loss of generality, we suppose that the vector sets $\Sigma = \{\sigma(S)\}$ and $\Theta = \{\tau(I)\}$ are ordered lexicographically by descending and their elements $\sigma^i$ and $\tau^j$ are labelled by natural numbers $i = 1, 2, \ldots, \binom{q}{s}$ and $j = 1, 2, \ldots, \binom{q}{k}$, respectively.

### 3.2 Upper Value

In this subsection, we prove the exact formulas for calculating $\delta_{q,k}$. Then, we find a mixed strategy equilibrium for the game in question.

**Theorem 3.** *For any natural numbers $k < q$, the following equations*

$$\delta_{q,k} = \frac{s}{q} \sum_{l=t-1}^{k-1} \frac{\binom{l}{t-1}\binom{(q-1)-l}{(s-1)-(t-1)}}{\binom{q-1}{s-1}} = \frac{k}{q} \sum_{l=t-1}^{s-1} \frac{\binom{l}{t-1}\binom{(q-1)-l}{(k-1)-(t-1)}}{\binom{q-1}{k-1}} \quad (9)$$

*are valid.*

Theorem 3 generalizes several known results. For example, when substituting in (9) $k = t = 1$, we easily obtain the following.

**Corollary 1.** *Any system (6) that has a committee solution of length $q$ contains a feasible subsystem of relative size at least $\lceil \frac{q+1}{2} \rceil / q$.*

### 3.3 Mixed Strategy Equilibrium

To prove the existence of a mixed strategy equilibrium of the game $\Gamma$, we demonstrate that its mixed extension coincides with a mixed extension of a corresponding matrix game.

By construction the pure strategy set of *Researcher* is finite. His or her set $\bar{X}$ of mixed strategies is

$$\bar{X} = \left\{ x \in \mathbb{R}^{\binom{q}{k}} : \sum_{j=1}^{\binom{q}{k}} x_j = 1, x \geq 0 \right\}.$$

Let us consider the pure strategy set $M(q)$ of *the Nature*. We begin with exclusion from $M(q)$ the matrices containing a row with more than $\lceil \frac{q+1}{2} \rceil$ ones, since these matrices are dominated by some other pure strategies of the second player. Then, because of the evident invariance of the payoff function to any permutation of rows in a matrix and simultaneous cloning of them, we proceed with exclusion of one matrix from any couple of equivalent strategies.

### 3.4 Asymptotic Bounds

In this section we introduce approximate formulas to calculate $\delta_{q,k}$ for large values of $q$ assuming that $k = q - n$ for some fixed integer $n$.

**Theorem 4.** *Let $\lambda$ be an arbitrary natural number.*

*(i) If $n = 2\lambda$, then*

$$\lim_{q\to\infty} \delta_{q,q-n} = \frac{1}{2}\left(1 + b(\lambda; 2\lambda - 1, 0.5)\right). \tag{10}$$

*(ii) If $n = 2\lambda - 1$, then the limit $\lim_{q\to\infty} \delta_{q,q-n}$ does not exist, since*

$$\lim_{s\to\infty} \delta_{2s,2(s-\lambda)+1} = \frac{1}{2} + b(\lambda; 2\lambda - 1, 0.5) \tag{11}$$

$$\lim_{s\to\infty} \delta_{2s-1,2(s-\lambda)} = \frac{1}{2}. \tag{12}$$

**Corollary 2.** *For any fixed natural $k$ and $q > k$ the following equations*

$$\delta_{q,k} \geq 1/2 \quad \text{and} \quad \lim_{q\to\infty} \delta_{q,k} = \frac{1}{2}$$

*are valid.*

**Corollary 3.** *Limit Eqs.* (10) *and* (11) *depend asymptotically on $\lambda$ as follows*

$$\lim_{q\to\infty} \delta_{q,q-2\lambda} \sim \frac{1}{2} + \frac{1}{\sqrt{2\lambda - 1}}\,\varphi\left(\frac{1}{\sqrt{2\lambda - 1}}\right)$$

*and*

$$\lim_{s\to\infty} \delta_{2s,2(s-\lambda)+1} \sim \frac{1}{2} + \frac{2}{\sqrt{2\lambda - 1}}\,\varphi\left(\frac{1}{\sqrt{2\lambda - 1}}\right).$$

*Proof.* The given asymptotic equations follow straight-forwardly from the famous De Moivre – Laplace theorem. According to it, the equation

$$b(i; n, p) \sim \frac{1}{\sqrt{np(1-p)}}\,\varphi\left(\frac{i - np}{\sqrt{np(1-p)}}\right).$$

holds uniformly by $i$. In our case,

$$n = 2\lambda - 1,\ i = \lambda,\ p = \frac{1}{2},$$

which completes the proof.

It is remarkable that the Gaussian approximation for the biggest value of the binomial distribution probability mass function performs well even for rather small values of the parameter $\lambda$. In particular, this is fortified by the following numeric data (Tables 1 and 2).

**Table 1.** Numeric evaluation, part 1

| $\lambda$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\lim_{q\to\infty} \delta_{q,q-2\lambda}$ | 0.750 | 0.688 | 0.656 | 0.637 | 0.623 | 0.613 | 0.605 | 0.598 | 0.593 | 0.588 |
| Gaussian appr. | 0.742 | 0.695 | 0.661 | 0.640 | 0.626 | 0.615 | 0.606 | 0.600 | 0.594 | 0.589 |
| Relative error | 0.011 | 0.010 | 0.008 | 0.005 | 0.005 | 0.003 | 0.003 | 0.003 | 0.002 | 0.002 |

**Table 2.** Numeric evaluation, part 2

| $\lambda$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\lim_{s\to\infty} \delta_{2s,2(s-\lambda)+1}$ | 1.000 | 0.875 | 0.812 | 0.773 | 0.746 | 0.726 | 0.709 | 0.696 | 0.685 | 0.676 |
| Gaussian appr. | 0.984 | 0.891 | 0.823 | 0.781 | 0.752 | 0.730 | 0.713 | 0.699 | 0.688 | 0.678 |
| Relative error | 0.016 | 0.018 | 0.014 | 0.010 | 0.008 | 0.006 | 0.006 | 0.004 | 0.004 | 0.003 |

## 4 Affine Separating Committees and Ensembles of Linear Classifiers

In this section we study the properties of committee solutions of infeasible system of linear inequalities, which is a special kind of an abstract system of constraints (1). In this case, committee solutions are closely connected with the special type of learning algorithms known in literature as *ensemble learning techniques*.

We begin with the common setting of the two-pattern classification problem (see, e.g. [3]). Suppose, we are given a probabilistic triple $(X \times Y, \mathcal{A}, P)$. Here the feature space $X$ and the set $Y = \{-1, 1\}$ of class labels. In many cases, we can suppose that $X$ is a subset of the $n$-dimensional Euclidean space $E_n$. It is required, in the preliminary given family of classifiers $\mathcal{H} \subset [X \to Y]$, to find "the most accurate" $\bar{h}$. Numerous formalizations are admitted due to an accuracy criterion. For simplicity, we focus on the following one

$$\bar{h} = \arg\min\{P(f(x) \neq y) \colon h \in \mathcal{H}\},$$

i.e. on finding a classifier that minimizes the misclassification probability.

If the probabilistic measure is known, this problem has the well-known closed form solution—the Bayes classifier. In the general case studied in this section, when all information about the unknown measure $P$ is exhausted by the finite i.i.d. *training* sample

$$(x_1, y_1), \ldots, (x_m, y_m), \tag{13}$$

the goal is to propose an efficient *learning* algorithm that could find a good approximation to the desired optimal classifier. Within the famous Vapnik-Chervonenkis structural risk minimization learning approach, it is important to design learning algorithms minimizing frequency of misclassification on sample (13) regularized by a capacity of the family $\mathcal{H}$ in terms of its VC-dimension.

We consider the setting of such a learning problem, where it is needed to fit a piecewise linear classifier

$$h(x) = \operatorname{sign} \sum_{j=1}^{k} \alpha_j \operatorname{sign}(c_j^T x - d_j) \tag{14}$$

for some non-negative weights $\alpha_j$, which without loss of generality can be assumed as integers, vectors $c_j$ and real biases $d_j$. In literature (see, e.g. [13], classifier (14) is called an *affine separating committee*. The motivation to study such classifiers arises from the following points:

(i) for any non-contradictory[2] sample (13), there exists a *perfect* affine committee classifier that makes no classification errors on this sample [17]
(ii) the family of affine committees (14) defined over the $n$-dimensional feature space $E_n$ and sharing the property $\sum_{j=1}^{k} \alpha_j = q$ has bounded VC-dimension [14].

We continue with the following notation. Let subsets $A$ and $B$ be defined (by sample (13)) as follows

$$A = \{x_i \colon y_i = 1\}, \quad B = \{x_i \colon y_i = -1\}. \tag{15}$$

Any classifier $h$ determined by Eq. (14) can be equivalently represented by the following finite sequence $K = K(h) = (f_1, \ldots, f_q)$, such that $q = \sum_{j=1}^{k} \alpha_j$ and

$$\begin{aligned} f_1(x) \equiv \ & \ldots \equiv f_{\alpha_1}(x) \equiv c_1^T x - d_1, \\ f_{\alpha_1+1}(x) \equiv \ & \ldots \equiv f_{\alpha_1+\alpha_2}(x) \equiv c_2^T x - d_2, \\ & \ldots \\ f_{q-\alpha_k+1}(x) \equiv \ & \ldots \equiv f_q(x) \equiv c_k^T x - d_k. \end{aligned}$$

It can be easily seen that an affine separating committee is a natural generalization of the concept of a separating hyperplane in Euclidean spaces. By means of the famous Hyperplane Separation Theorem (see, e.g. [4]), for any finite sets $A$ and $B$, the equation

$$\operatorname{conv}(A) \cap \operatorname{conv}(B) = \varnothing$$

presumes the existence of a linear function $f(x) = c^T x - d$ such that the hyperplane $H = \{x \in E_n \colon c^T x - d = 0\}$ separates these sets, i.e., $f(a) > 0$ and $f(b) < 0$ for any $a \in A$ and $b \in B$, respectively. Therefore, if the sets $A$ and $B$ are separable in the regular case, then there exists an affine committee of length 1 that separates them. For the general case, the following criterion is valid.

**Theorem 5 ([17]).** *Finite subsets $A, B \subset E_n$ can be separated by an affine committee if an only if $A \cap B = \varnothing$.*

[2] For which the condition $x_{i_1} = x_{i_2}$ implies $y_{i_1} = y_{i_2}$.

*Problem 1 (Minimum Affine Separating Committee (MASC)).* For the given sets $A, B \subset E_n$ it is necessary to find an affine separating committee of the minimum length.

In the conclusion of this section, we give a brief outline of the recent results concerning algorithmic analysis of the MASC problem following the paper [13].

**Theorem 6 ([8]).** *The Minimum Affine Separating Committee problem is strongly NP-hard and remains intractable even in the case, when*

$$A \cup B \subset \{x \in \{0,1,2\}^n \colon \|x\|_2 \leq 2\}.$$

*The MASC problem does not refer to the APX approximability class, unless $P \neq NP$.*

According to Theorem 6, the MASC problem is hard to solve not only in the class of exact algorithms but even with any constant approximation ratio. The following theorem extends this result to the spaces of any fixed dimension.

**Theorem 7 ([14]).** *The MASC problem is polynomially solvable in the real line and strongly NP-hard in $n$-dimensional Euclidean space for any fixed dimension $n > 1$.*

It is noteworthy that the claim of Theorem 7 remains valid even in the case, when the set $A \cup B$ is in the *general position*. Usually, a finite set $D \subset E_n$ of size $|D| > n$ is said to be *in general position*, if, for any $D' \subset D$, $|D'| = n + 1$, dimension of the affine hull $\text{aff}(D')$ is equal to $n$. The special setting of the MASC problem given in the $n$-dimensional Euclidean space with additional condition on general position of $A \cup B$ is known as MASC-GP($n$).

Nearly all known results in the scope of efficient algorithm construction for the MASC problem are based on the following theorem, which can be considered as a specification of Theorem 1 to the case, when training sets are in general position.

**Theorem 8 ([9]).** *For any finite subsets $A, B \subset E_n$ being in general position, for which $A \cap B = \varnothing$ and $|A \cup B| = m$, there exists an affine separation committee of length*

$$q \leq 2 \left\lceil \frac{\lfloor (m-n) \rfloor}{n} \right\rceil + 1. \quad (16)$$

Two subsequent geometric properties of finite dimensional Euclidean spaces lead to the proof of Theorem 8 mainly.

*Property 1.* Let $Z$ be a finite subset of $E_n$ and $\varnothing \neq Z' \subset Z$ such that $|Z'| \leq n$ and be in general position. Then, there exist open half-spaces $L_1 = \{x \colon c_1^T - d_1 < 0\}$ and $L_2 = \{x \colon c_2^T - d_2 < 0\}$ such that $Z \subset L_1 \cup L_2$ and $Z' \subset L_1 \cap L_2$.

*Property 2.* Let $A$ and $B$ be non-empty finite subsets of $E_n$, where $A \cup B$ is in general position and of size $m > n$. Then, for any subsets $A' \subset A$ and $B' \subset B$ of common size $A' \cup B' = n$ there exist $A'' \supseteq A'$ and $B'' \supseteq B'$ and a function $f(x) = c^T x - d$, such that $f(a) > 0$ and $f(b) < 0$ for any $a \in A''$ and $b \in B''$ respectively, and $|A'' \cup B''| \geq \lceil (m+n)/2 \rceil$.

Indeed, the proof of Theorem 8 extends the proof of Theorem 5 and propose a polynomial time approximation algorithm for the problem MASC-GP($n$) with time complexity bound $O(m/n \times T_n)$ and the approximation ratio $O(m/n)$. Here, $T_n$ signifies the difficulty of solving a Kramer system of linear equations over $n$ variables.

*Remark 1.* Bound (16) is tight. In particular, it is attained on sets mentioned in [6] and called *uniformly distributed sets*. In [7], the MASC problem is shown to be polynomially solvable over such sets. The formal definition is as follows

**Definition 5.** *A finite set $Z = A \cup B \subset E_n$ is known as uniformly distributed (by Gale), if $A \cap B = \varnothing$, $|A \cup B| = n + 2k$ for some natural $k$ and, for any non-trivial hyperplane $H = \{x \in E_n \colon f(x) \equiv c^T x - d = 0\}$, there exist $A' \subset A$ and $B' \subset B$, $|A' \cup B'| \geq k$, such that $f(a) > 0$ and $f(b) < 0$ for any $a \in A'$ and $b \in B'$, respectively.*

It is known that, for any natural numbers $n$ and $k$, there exists a uniformly distributed subset $Z = A \cup B \subset E_n$ of size $2k + n$. Thus, in terms of machine learning, it can be stated that, any time, when a training sample is defined by a uniformly distributed subset, the algorithm proposed in the proof of Theorem 5, in the family of the smallest VC dimension, in linear time with respect to the sample length, will obtain a committee classifier (14) without making any classification errors.

Nowadays, the advanced approximation algorithms for the MASC problem are based on the synthesis of the mentioned above approach and the famous Multiple Weights Update technique (see, e.g. [2]). Characteristics of the *BoostedGreedyCommittee* algorithm [13] which has the best known approximation factor are shown in the following theorem.

**Theorem 9.** *BoostedGreedyCommittee finds an $O(((m \ln m)/n)^{1/2})$-approximate solution for the MASC problem in time $m^{O(n)}$. If, for the given sets $A$ and $B$, there is a minimum committee $(f_0, f_1, \ldots, f_{q-1})$ such that, for any $t = 1, \ldots, (q-1)/2$ and any $a \in A$, $b \in B$ the following equation*

$$(f_{2t-1}(a) > 0 \vee f_{2t}(a) > 0) \wedge (f_{2t-1}(b) < 0 \vee f_{2t}(b) < 0)$$

*is valid, then the approximation factor of this algorithm is $O(\ln m)$.*

## 5 Conclusion

This survey does not pretend to be called exhaustive. We intentionally restrict ourselves to some theoretic results concerning the committees, leaving without considering their applications to numerous practical decision making problem in economy, industry, and medicine forwarding the interested reader to recent papers presenting interesting results in procatice, e.g. [1,5,16,22].

## References


1. Akberdina, V., Chernavin, N., Chernavin, F.: Application of the committee machine method to forecast the movement of exchange rates and oil prices. Digest Finan. **23**(1), 108–120 (2018). https://doi.org/10.24891/df.23.1.108
2. Arora, S., Hazan, E., Kale, S.: The multiplicative weights update method: a meta-algorithm and applications. Theory Comput. **8**(1), 121–164 (2012)
3. Bishop, C.M.: Pattern Recognition and Machine Learning. Information Science and Statistics. Springer, New York (2007)
4. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2009)
5. Gainanov, D., Berenov, D.: Algorithm for predicting the quality of the product of metallurgical production. In: Evtushenko, Y., Khachay, M., Khamisov, O., Kochetov, Y., Malkova, V., Posypkin, M. (eds.) Proceedings of the VIII International Conference on Optimization and Applications (OPTIMA-2017), Petrovac, Montenegro, 2–7 October 2017, pp. 194–200. No. 1987 in CEUR Workshop Proceedings, Aachen (2017). http://ceur-ws.org/Vol-1987/paper29.pdf
6. Gale, D.: Neighboring vertices on a convex polyhedron. Linear Inequalities Relat. Syst. **38**, 255–263 (1956)
7. Khachai, M.: Computational complexity of the minimum committee problem and related problems. Dokl. Math. **73**, 138–141 (2006). https://doi.org/10.1134/S1064562406010376
8. Khachai, M.: Computational and approximational complexity of combinatorial problems related to the committee polyhedral separability of finite sets. Pattern Recogn. Image Anal. **18**(2), 236–242 (2008)
9. Khachai, M., Rybin, A.: A new estimate of the number of members in a minimum committee of a system of linear inequalities. Pattern Recogn. Image Anal. **8**, 491–496 (1998)
10. Khachai, M.: On the existence of majority committee. Discrete Math. Appl. **7**(4), 383–397 (1997). https://doi.org/10.1515/dma.1997.7.4.383
11. Khachai, M.: A relation connected with a decision making procedure based on mojority vote. Dokl. Math. **64**(3), 456–459 (2001)
12. Khachai, M.: A game against nature related to majority vote decision making. Comput. Math. Math. Phys. **42**(10), 1547–1555 (2002)
13. Khachay, M.: Committee polyhedral separability: complexity and polynomialapproximation. Mach. Learn. **101**(1–3), 231–251 (2015). https://doi.org/10.1007/s10994-015-5505-0
14. Khachay, M., Poberii, M.: Complexity and approximability of committee polyhedral separability of sets in general position. Informatica **20**(2), 217–234 (2009)
15. Kobylkin, K.: Constraint elimination method for the committee problem. Autom. Remote Control **73**, 355–368 (2012). https://doi.org/10.1134/s0005117912020130
16. Lebedeva, E., Kobzeva, N., Gilev, D., Kislyak, N., Olesen, J.: Psychosocial factors associated with migraine and tension-type headache in medical students. Cephalalgia **37**(13), 1264–1271 (2017). https://doi.org/10.1177/0333102416678389

17. Mazurov, V.: Committees of inequalities systems and the pattern recognition problem. Kibernetika **3**, 140–146 (1971)
18. Mazurov, V.: Committee Method in Problems of Optimization and Classification. Nauka, Moscow (1990)
19. Mazurov, V., Khachai, M.: Committees of systems of linear inequalities. Autom. Remote Control **65**, 193–203 (2004). https://doi.org/10.1023/b:aurc.0000014716.77510.61
20. Mazurov, V., Khachai, M., Rybin, A.: Committee constructions for solving problems of selection, diagnostics, and prediction. In: Proceedings of the Steklov Institute of Mathematics (suppl. 1), pp. S67–S101 (2002)
21. Nilsson, N.: Learning Machines: Foundations of Trainable Pattern Classifying Systems. McGraw-Hill, New York (1965)
22. Pandey, T.N., Jagadev, A.K., Dehuri, S., Cho, S.B.: A novel committee machineand reviews of neural network and statistical models for currency exchangerate prediction: an experimental analysis. J. King Saud Univ. Comput. Inf. Sci. (2018). https://doi.org/10.1016/j.jksuci.2018.02.010. http://www.sciencedirect.com/science/article/pii/S1319157817303816
23. Rosenblatt, F.: Principles of neurodynamics: perceptrons and the theory of brain mechanisms (1986). https://doi.org/10.1007/978-3-642-70911-1_20
24. Rybin, A.: On some sufficient conditions of existence of a majority committee. Pattern Recorgn. Image Anal. **10**(3), 297–302 (2000)
25. Vapnik, V.N.: Statistical Learning Theory. Adaptive and Learning Systems for Signal Processing, Communications, and Control. Wiley (1998)
26. Vapnik, V.N.: The Nature of Statistical Learning Theory. Statistics for Engineering and Information Science, 2nd edn. Springer, New York (2000). https://doi.org/10.1007/978-1-4757-3264-1
27. Zhuravlev, Y.I.: Correct algebras over sets of incorrect (heuristic) algorithms. I. Cybernetics **13**(4), 489–497 (1977). https://doi.org/10.1007/BF01069539
28. Zhuravlev, Y.I.: Correct algebras over sets of incorrect (heuristic) algorithms. II. Cybernetics **13**(6), 814–821 (1977). https://doi.org/10.1007/BF01068848