

**М. В. Куклінський, к.т.н.**  
Національний авіаційний університет,  
просп. Комарова, 1, м. Київ, 03058, Україна  
[maximum\\_inc@ua.fm](mailto:maximum_inc@ua.fm)

## ВЕКТОРНА ОПТИМІЗАЦІЯ НЕЙРОМЕРЕЖЕВОГО КЛАСИФІКАТОРА ТЕКСТІВ

*На теперішній час нейромережеві технології набули широкого застосування в задачах інтелектуального аналізу даних. До одного із спектрів таких задач можна віднести задачі класифікації текстів. Популярність використання нейромережевих алгоритмів саме в цих задачах пов'язана з високою ефективністю їх роботи. Тому в процесі проектування систем, які застосовують нейромережеві алгоритми, особлива увага приділяється питанню їх архітектури. Це, насамперед, обумовлено вимогами, які ставляться до підвищення ефективності та надійності функціонування цих систем. Якщо певним властивостям нейромережевого класифікатора поставити відповідну кількісну характеристику, яка має сенс критерію, то можна говорити про багатокритеріальну оптимізацію цього класифікатора. Такий критеріальний підхід дозволить вирішити один із недоліків використання текстових класифікаторів, який пов'язаний з неефективною побудовою архітектури його нейронної мережі. У статті розглянуто основні підходи до векторної багатокритеріальної оптимізації нейромережевих класифікаторів та запропоновано метод пошуку компромісно-оптимальної архітектури нейромережевого класифікатора текстів.*

**Ключові слова:** нейронна мережа, векторна оптимізація, критерії якості, нейромережевий класифікатор, текст.

**Вступ.** В методах інтелектуального аналізу даних можна виділити цілу групу задач, в яких необхідно визначити, до якого з відомих класів відносяться досліджувані в них об'єкти. Цю групу прийнято називати задачами класифікації. Очевидно, що результати аналізу в цих задачах приймаються на підставі даних про досліджувані об'єкти, причому усі класи, до яких може бути віднесений об'єкт, відомі заздалегідь.

Саме до такої групи задач відносяться задачі класифікації тексту, актуальність яких завжди залишається високою. З одного боку, це викликано стрімким збільшенням текстової інформації, з другого, – різноманітністю способів її використання, зберігання та обробки. Насамперед, це стосується інформації, яка представлена у цифровому вигляді та оброблюється за допомогою комп'ютерних технологій. Системи, які оброблюють текстову інформацію для вирішення задач класифікації, називаються текстовими класифікаторами.

**Аналіз публікацій та постановка мети.** Питанням класифікації тексту присвячено чимало публікацій. Наприклад, у [1–2] зібрано й описано основні підходи та методи щодо класифікації текстових документів. Показано їх переваги та недоліки.

Серед недоліків можна виділити один, який притаманний більшості класифікаторів, – це істотне зниження продуктивності в разі значного зростання обсягу оброблюваної інформації та збільшення кількості класів, за якими необхідно класифікувати документи. У [3–4] цей недолік вирішуються шляхом розподілення процесу класифікації за допомогою використання багатопроцесорних обчислювальних систем і комплексів.

У [5–7] вирішується ще одна проблема задачі класифікації, коли, наприклад, два тексти, що близькі по ключових словах та входять до одного класу, виявляються зовсім різними, з точки зору їх вузької тематичної спрямованості.

Загалом, проведений аналіз більшості методів показав, що вибір підходящого алгоритму класифікації тексту найчастіше визначається експериментально під конкретну задачу. І, незважаючи на пристойні результати більшості методів, найкращі результати, з точки зору ефективності, показували класифікатори з використанням нейронних мереж.

Насамперед, це викликано такими перевагами нейромережевих класифікаторів:

– урахування термінів довільного походження (слів, словосполучень, метаконструкцій, гіперпосилань тощо);

– неявне врахування взаємозв'язків між термінами (наприклад, асоціативних зв'язків, відносин рід/вид тощо);

– потенційно висока ефективність мережі (якщо взяти мережу з великою кількістю нейронів і шарів, витратити обчислювальні ресурси на її навчання, то можна очікувати високих показників ефективності).

Але й цьому класу алгоритмів притаманні деякі недоліки. Навчання та використання великої нейронної мережі вимагає великої кількості пам'яті і обчислень. Також низька стійкість до «шуму» в навчальній вибірці показує низьку ефективність роботи нейронних алгоритмів з великими ієрархіями класів «у чистому вигляді», що, у свою чергу, призводить до помилок мережі [2].

Метою статті є розкриття проблеми пошуку оптимальної архітектури нейронної мережі, при якій її максимальна помилка не перевищувала б гранично допустиме значення.

**Загальні положення.** Зазвичай задачі класифікації вирішуються в два етапи:

– на першому етапі виділяється навчальна вибірка, до якої входять значення як за-

лежних, так і незалежних змінних вже відомих об'єктів. На підставі навчальної вибірки вибудовується модель визначення значення залежної змінної (функції класифікації);

– на другому етапі побудовану модель застосовують до аналізованих об'єктів (до об'єктів з невизначеним значенням залежної змінної).

У свою чергу, для отримання максимально точної функції класифікації до навчальної вибірки ставляться свої вимоги:

– кількість об'єктів, які входять у вибірку, повинна бути достатньо великою. Чим більше об'єктів, тим точніше буде побудована на її основі функція класифікації;

– у вибірку повинні входити об'єкти, які представляють всі можливі класи;

– для кожного класу в задачі класифікації вибірка повинна містити достатню кількість об'єктів [1].

Майже такі ж вимоги висувуються і за таким же принципом працюють нейронні класифікатори, побудовані на нейронних мережах, складовою частиною яких є штучний нейрон (рис. 1).

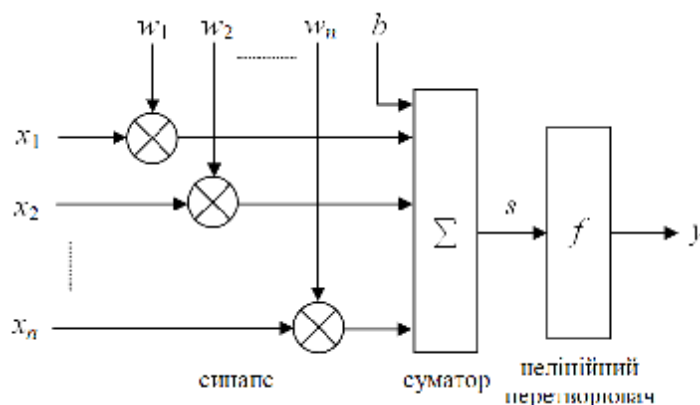


Рис. 1. Структура штучного нейрона

Як видно, він складається з елементів трьох типів: помножувачів (синапсів), суматора і нелінійного перетворювача. Синапси здійснюють зв'язок між нейронами, множать вхідний сигнал на число, яке характеризує силу зв'язку (вагу синапсу). Суматор виконує додавання сигналів, які надходять по синаптичних зв'язках від інших нейронів та зовнішніх вхідних сигналів. Нелінійний перетворювач реалізує нелінійну функцію одного аргументу – виходу суматора. Ця функція називається функцією активації, або передавальною функцією нейрона.

Нейрон, в цілому, реалізує скалярну функцію векторного аргументу. Математична модель нейрона має вигляд:

$$s = \sum_{i=1}^n w_i x_i + b,$$

$$y = f(s),$$

де  $w$  – вага синапсу,  $i=1..n$ ;  $b$  – значення зміщення,  $s$  – результат підсумовування;  $x$  – компонент вхідного вектора (вхідний сигнал),  $i=1..n$ ;  $y$  – вихідний сигнал нейрона;  $n$  – кількість входів нейрона;  $f$  – нелінійне перетворення (функція активації).

У загальному випадку вхідний сигнал, вагові коефіцієнти і зміщення можуть брати дійсні значення, а в багатьох практичних завданнях – лише деякі фіксовані значення. Вихідний сигнал визначається видом функції активації і може бути як дійсним, так і цілим [8].

Нейронна мережа являє собою сукупність нейронів, певним чином з'єднаних один

з одним та з зовнішнім середовищем за допомогою зв'язків, які визначаються ваговими коефіцієнтами. Наприклад, у загальному вигляді структура  $q$ -шарового нейромережевого класифікатора з прямими зв'язками буде наступною (рис. 2).

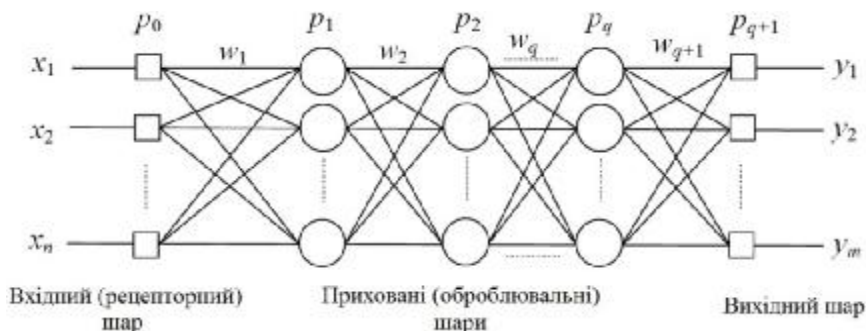


Рис. 2. Загальний вигляд нейромережевого класифікатора з прямими зв'язками

Тут  $x_1, x_2, \dots, x_n$  – ознаки об'єкта класифікації, що становлять вхідний вектор  $x = \{x_i\}_{i=1}^n$ ;  $p_0 = n$  – кількість нейронних елементів у вхідному шарі;  $p_1, p_2, \dots, p_q$  – кількість нейронів в кожному з прихованих (оброблювальних) шарів;  $p_{q+1} = m$  – кількість нейронів у вихідному шарі (кількість класів);  $y = \{y_k\}_{k=1}^m$  – вихідний вектор нейронної мережі, що визначає належність об'єкта класифікації одному з  $m$  класів;  $w_1, w_2, \dots, w_q, w_{q+1}$  – вектори синаптичних вагів нейронної мережі.

Як видно з рис. 2, множина всіх нейронів штучної нейронної мережі розділяється на підмножини, які називаються шарами. Шар – це множина нейронів, на які в кожний такт часу паралельно надходять сигнали від інших нейронів цієї мережі. На виході класифікатора виходить вектор функцій активації  $y = \{y_k\}_{k=1}^m$ . Номер  $j$ , для якого вихід  $y_j$  має максимальну активність, тобто  $\max_{k \in [1, m]} y_k = y_j$ ,

відповідає номеру класу об'єкта класифікації.

Кількість нейронів вхідного шару  $p_0 = n$  визначається розмірністю вхідного вектора ознак і не підлягає змінам. Аналогічно, кількість нейронів вихідного шару  $p_{q+1} = m$  визначається кількістю областей (класів), на які ділиться простір ознак, і теж є

постійною. Кількість же оброблювальних (прихованих) шарів  $q$  і кількість нейронів у кожному з них  $p_1, p_2, \dots, p_q$  становлять поняття архітектури нейронної мережі і можуть служити аргументами при її оптимізації [8].

**Постановка задачі.** Нехай кількість оброблювальних (прихованих) шарів  $q$  є фіксованою і заданою. Тоді аргументами оптимізації архітектури нейронного класифікатора є кількість нейронів у кожному з оброблювальних шарів, а саме складові вектора незалежних змінних  $p = \{p_j\}_{j=1}^q$ . Саме від вибору архітектури  $p$  залежить якість функціонування нейронного класифікатора. Проблема полягає у виборі такої архітектури класифікатора, при якій в заданих умовах функціонування він характеризувався б своїми якнайкращими властивостями.

У загальному вигляді проблема може бути формально представлена задачею

$$p^* = \arg \operatorname{extr}_{p \in P} Y(p), \quad (1)$$

де  $Y(p)$  – цільова функція;  $\operatorname{extr}_{p \in P}$  – оператор екстремізації цільової функції по аргументах  $p$ ;  $P$  – допустима область незалежних змінних.

Якщо кожній властивості нейронного класифікатора поставити у відповідність кількісну характеристику  $f(p)$ , яка має сенс критерію якості його функціонування, то одним

із таких критеріїв буде вірогідність помилки класифікації.

Цей критерій можна визначити експериментально і приблизно представити як кількість помилок класифікації  $e(p)$ , віднесена до загальної, достатньо великої кількості випробувань  $N$

$$f_1(p) = \frac{e(p)}{N}. \quad (2)$$

Передбачається, що зі зростанням у деяких розумних межах кількості нейронів в оброблювальних шарах точність класифікації підвищується, а величина цього критерію зменшується. Гранично допустиме значення помилки мережі має бути відоме з фізичних міркувань і задане як обмеження  $f_1(p) \leq A_1$ .

Другий критерій характеризує час, необхідний для навчання нейронної мережі при даній архітектурі  $p$ . Спостерігається тісна кореляція між таким часом і сумарною кількістю нейронів у прихованих шарах класифікатора. Тому представимо цей критерій у вигляді

$$f_2(p) = \sum_{k=1}^q p_k. \quad (3)$$

Відзначимо, що цим критерієм характеризується і час проходження сигналу через нейронну мережу від входу до виходу. Зі зростанням кількості нейронів значення цього критерію буде збільшуватися. Гранично допустиме значення другого критерію визначається допустимим часом навчання нейронної мережі і задається як обмеження  $f_2(p) \leq A_2$ .

Існують і інші критерії для характеристики різних властивостей нейронного класифікатора. Проте обмежимося тільки наведеними двома основними критеріями, маючи на увазі, що викладена методика допускає включення до розгляду й інших властивостей класифікатора.

Допустима область аргументів оптимізації задається паралелепіпедним обмеженням  $P = \{p | 0 \leq p_k \leq P_u, k \in [1, P_u], u \in [1, q]\}$ , де  $P_u$  – максимальна кількість нейронів в  $u$ -му шарі.

Оскільки обидва критерії, які розглядаються, підлягають мінімізації, чим критерій менший, тим краща відповідна властивість класифікатора, то оператор екстремізації цільової функції набуває вигляду:  $extr = \min_{p \in P}$ .

Таким чином, обидва критерії є суперечливими, невід'ємними, такими, що мінімізу-

ються, і обмеженими. Саме ці передумови дозволяють використати до цільової функції метод скалярної згортки критеріїв по нелінійній схемі компромісів [9].

Така згортка в уніфікованій версії виражається формулою

$$Y(p) = Y[f(p)] = \frac{A_1}{A_1 - f_1(p)} + \frac{A_2}{A_2 - f_2(p)}. \quad (4)$$

де  $f(p) = \{f_r(p)\}_{r=1}^{r=2}$  – двовимірний вектор часткових критеріїв.

Враховуючи (2), (3) і (4), вираз (1) для задачі оптимізації архітектури нейронного класифікатора перетвориться до вигляду

$$p^* = \arg \min_{p \in P} \left[ \frac{A_1}{A_1 - e(p)/N} + \frac{A_2}{A_2 - \sum_{k=1}^q p_k} \right]. \quad (5)$$

Неважко побачити, що у формулі (5) залежність  $e(p)$  апіорі є невідомою. Для її визначення в заданих умовах повинна бути організована спеціальна експериментальна процедура. Тому метод векторної оптимізації архітектури нейронного класифікатора буде теоретико-експериментальним.

**Оптимізація нейромережевого текстового класифікатора.** Як приклад розглянемо у загальних рисах задачу багатокритеріальної оптимізації архітектури нейромережевого класифікатора текстів. Система класифікації тексту [10] складається з двох основних частин: частотний аналізатор з системним словником і власне нейромережевий класифікатор (рис. 3).

На вхід системи надходить текст, на виході виходить номер теми, до якої відноситься цей текст (бізнес, політика, медицина, спорт, просто спам тощо).

Перш ніж приступити до оптимізації архітектури нейромережевого класифікатора текстів, необхідно виконати такі етапи:

– визначаються  $m$  класів, з якими працюватиме система;

– підбираються відповідні учбові тексти  $t_k, k \in [1, m]$  і перевірні (тестові) тексти  $t_l, l \in [1, L], L \geq m$ ;

– з множини учбових текстів спеціальним чином виділяються слова  $v_i, i \in [1, n]$  і формується системний словник  $V$ ;

– частотний аналізатор визначає для кожного слова  $v_i$  з системного словника  $V$  його частоту входження  $x_i$  в даний текст  $t_k$ . Частотна характеристика – це вектор

$x = \{x_i\}_{i=1}^n$  ознак тексту  $t_k$ , розмірність якого дорівнює кількості слів у системному словнику  $v_i \in V$ .



Рис. 3. Загальна схема класифікатора

Отримавши результати частотного аналізу учбових текстів, можна приступати до навчання нейронного класифікатора при деякій архітектурі  $p = \{p_j\}_{j=1}^q$ . Процес навчання нейронної мережі полягає у встановленні таких вагових коефіцієнтів її зв'язків, при яких максимальна помилка мережі на учбових текстах для цієї архітектури не перевищує гранично допустимого значення. Конкретні алгоритми навчання тут не розглядаються.

Після виконання названих етапів можна приступити безпосередньо до теоретико-експериментальної процедури векторної оптимізації.

Для оптимізації архітектури нейромережевого класифікатора можна скористатися одним із пошукових методів, наприклад, методом симплекс-планування. Нехай для визначеності кількість оброблювальних шарів  $q=2$ . Тоді ідею методу в безперервному варіанті можна ілюструвати за допомогою рис. 4.

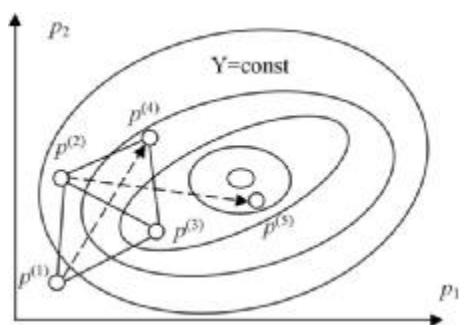


Рис. 4. Ілюстрація пошукового методу

На площині аргументів  $p_1 - p_2$  у деякій стартовій області будуємо початковий регулярний симплекс, який у двовимірному випадку є рівнобедреним трикутником з вершинами

$p^{(1)}, p^{(2)}, p^{(3)}$ . Для кожної архітектури симплексу здійснюємо процес навчання класифікатора і подаємо на вхід серію тестових текстів  $t_l$ . У кожній вершині симплексу експериментально визначаємо кількість помилок класифікації при загальній кількості випробувань  $N = L$ . За формулою (2) одержуємо критерії  $f_1^{(1)}, f_1^{(2)}, f_1^{(3)}$ . За формулою (3) визначаються критерії  $f_2^{(1)}, f_2^{(2)}, f_2^{(3)}$ . Формула (4), яка в цьому випадку виступає не в ролі цільової, а в ролі оцінювальної функції, набуде вигляду

$$Y(p_1, p_2) = \frac{A_1}{A_1 - e(p_1, p_2)/L} + \frac{A_2}{A_2 - p_1 - p_2}. \quad (6)$$

Для архітектури стартового симплексу вона дасть значення скалярних згорток  $Y^{(1)}, Y^{(2)}, Y^{(3)}$ . Порівнюючи між собою ці значення, знаходимо, що одне з них, наприклад,  $Y^{(1)}$ , виявилось більшим (тобто гіршим), ніж інші. З великою вірогідністю можна стверджувати, що архітектура  $p^{(4)}$ , яка буде одержана дзеркальним відображенням гіршої в початковому симплексі точки  $p^{(1)}$  відносно центру протилежної грані, виявиться кращою. Здійснивши всі розрахунки для архітектури  $p^{(4)}$ , утворюємо новий симплекс з вершинами  $p^{(2)}, p^{(3)}$  і  $p^{(4)}$ . Порівнявши значення  $Y^{(2)}, Y^{(3)}, Y^{(4)}$ , виявимо, що одна з точок, наприклад  $p^{(2)}$ , виявилася гіршою за інші в сенсі другого симплексу. Відобразивши цю точку відносно центра протилежної грані другого симплексу, одержимо архітектуру  $p^{(5)}$ , і так

далі доти, поки не одержимо архітектуру  $p^*$ , відповідну мінімуму цільової функції.

Запропонований пошуковий метод передбачає проведення серії експериментів. Одержані при цьому експериментальні дані можуть бути використані для побудови аналітичної регресійної моделі часткового критерію  $f_1(p) = e(p)/L$ . За допомогою такої моделі можна здійснювати не пошукову, а аналітичну векторну оптимізацію архітектури інших нейромережових класифікаторів такого ж вигляду. Якщо це виявиться складним, то проводиться пошукова процедура, але вже із застосуванням не натурального, а обчислювального експерименту, що істотно простіше.

Проте, вирішуючи задачу побудови регресійної моделі, необхідно задавати вигляд апроксимуючої залежності, який буде відомим з точністю до коефіцієнтів регресії. Аналіз задачі дає припущення, що у цьому випадку, з достатньою практичною точністю, можна обмежитися лінійною регресією:

$$f_1(p_1, p_2) \approx (a_1 p_1 + a_2 p_2) / L,$$

де  $a_1, a_2$  – коефіцієнти регресії, які визначаються експериментальними даними за методом найменших квадратів. Лінійна регресійна модель може бути перевірена на адекватність методами математичної статистики та, при необхідності, може ускладнюватися.

**Висновки.** Розглянутий метод передбачає старт пошукової процедури від архітектури, яка може знаходитися достатньо близько до оптимальної точки. Якщо в процесі пошуку має місце зростання кількості нейронів в оброблювальних шарах, то теорія нейронних мереж характеризує цей підхід як конструктивний. При надмірній стартовій кількості нейронів підхід буде деструктивним.

Застосування запропонованого підходу векторної оптимізації дозволить визначити архітектуру нейромережового класифікатора, при якій системно ув'язуються суперечливі критерії ефективності його функціонування, а сама отримана архітектура буде компромісно-оптимальною.

### Список літератури

1. Анализ данных и процессов : учебн. пособие / [А. А. Барсегян, М. С. Куприянов, И. И. Холод и др.]. – [3-е изд., перераб. и

доп.]. – С.Пб. : БХВ–Петербург, 2009. – 512 с.

2. Шабанов В. И. Модели и методы автоматической классификации текстовых документов : дис. ... канд. техн. наук : 05.13.11 / Шабанов Владислав Игоревич. – М., 2003. – 227 с.
3. Жора Д. В. Распараллеливание алгоритмов функционирования классификатора со случайными подпространствами / Д. В. Жора // Проблемы программирования. – 2006. – № 2–3 [спец. вып.]. – С. 124–134.
4. Пескишева Т. А. Параллельный алгоритм обучения текстового классификатора для многопроцессорной системы с иерархической архитектурой / Т. А. Пескишева, Е. В. Котельников, О. А. Пестов // Вопросы современной науки и практики. Университет им. В. И. Вернадского. – 2011. – № 3 (34). – С. 103–110.
5. Захаров В. Н. Автоматическая оценка подобия тематического содержания текстов на основе сравнения их формализованных смысловых описаний / В. Н. Захаров, А. А. Хорошилов // Электронные библиотеки : перспективные методы и технологии, электронные коллекции : труды 14-й всеросс. научн. конф. RCDL–2012., 15–18 окт. 2012 г. – С. 143–149.
6. Шабанов В. И. Метод классификации текстовых документов, основанный на полнотекстовом поиске / В. И. Шабанов, А. М. Андреев // Труды первого российского семинара по оценке методов информационного поиска / под ред. И. С. Некрестьянова. – С.Пб. : НИИ Химии СПбГУ, 2003. – 132 с.
7. Добрынин В. Ю. Оценка тематического подобия текстовых документов / В. Ю. Добрынин, В. В. Клюев, И. С. Некрестьянов // Электронные библиотеки : перспективные методы и технологии, электронные коллекции : труды 2-й всеросс. научн. конф. RCDL–2000, 26–28 сент. 2000 г. – С. 204–210.
8. Круглов В. В. Искусственные нейронные сети. Теория и практика / В. В. Круглов, В. В. Борисов. – [2-е изд., стереотип.]. – М. : Горячая линия – Телеком, 2002. – 382 с.
9. Воронин А. Н. Векторная оптимизация динамических систем / А. Н. Воронин, Ю. К. Зиатдинов, А. И. Козлов. – К. : Техніка, 1999. – 284 с.

10. Борисов В. С. Самообучающийся классификатор текстов на естественном языке / В. С. Борисов // Кибернетика и системный анализ. – 2007. – № 3. – С. 169–176.
- References**
1. Barseghyan, A. A., Kupriyanov, M. S., Holod, I. I. et al. (2009). The analysis of the data and processes. St. Petersburg: BHV–Peterburg, 512 p. [in Russian].
  2. Shabanov V. I. (2003) Models and methods for automatic classification of text documents: dis. ... for Ph.D. in Engineering: 05.13.11. Moscow, 227 p. [in Russian].
  3. Zhora, D. V. (2006). Parallelization of algorithms for functioning of random subspace classifier. *Problemy prohranuvannya*, (2–3), pp. 124–134 [in Russian].
  4. Peskischeva, T. A. Kotelnikov, E. V. and Pestov, O. A. (2011). Parallel algorithm of text classifier learning for multiprocessor system with a hierarchical architecture. *Voprosy sovremennoy nauki i praktiki*, 3 (34), pp. 103–110 [in Russian].
  5. Zakharov, V. N. and Khoroshilov, A. A. (2012) Automatic evaluation of the similarity of texts thematic content based on a comparison of their formal semantic descriptions. *Elektronnyye biblioteki: perspektivnyye metody i tehnologii, elektronnyye kollektzii: trudy 14-j vseross. nauchn. konf. RCDL–2012, October 15–18*), pp. 143–149 [in Russian].
  6. Shabanov, V. I. and Andreev, A. M. (2003) Classification method of text documents based on full-text research. *Trudy pervogo rossiyskogo seminaru po otsenke metodov informatsionnogo poiska*. In: J. S. Nekrestyanov (Ed.). St. Petersburg, NII Himii S.PbGU, 132 p. [in Russian].
  7. Dobrynin, V. Yu., Klyuev, V. V. and Nekrestyanov, I. S. (2000) Evaluation of thematic similarity of text documents. *Elektronnyye biblioteki: perspektivnyye metody i tehnologii, elektronnyye kollektzii: trudy 2-j vseross. nauchn. konf. RCDL–2000 (September 26–28)*, pp. 204–210 [in Russian].
  8. Kruglov, V. V. and Borisov, V. V. (2002) Artificial neural networks. Theory and practice. Moscow: Goryachaya liniya – Telekom, 382 p. [in Russian].
  9. Voronin, A. N., Ziatdinov, Yu. K. and Kozlov, A. I. (1999) Vector optimization of dynamic systems. Kiev: Tehnika, 284 p. [in Russian].
  10. Borisov, V. S. (2007). Self-learning classifier of natural language text. *Kibernetika i sistemnyy analiz*, (3), pp. 169–176 [in Russian].

**M. V. Kuklinskiy, Ph.D.**  
National Aviation University,  
Komarova ave., 1, Kyiv, 03058, Ukraine  
[maximum\\_inc@ua.fm](mailto:maximum_inc@ua.fm)

## VECTOR OPTIMIZATION OF NEURAL NETWORK CLASSIFIER OF TEXTS

*At present time, neural network technologies have found the wide application in problems of intellectual analysis of the data. It is possible to relate such problems of text classification to spectra of such problems. The popularity of neural network algorithms being used in these problems is connected with high efficiency of their work. Therefore in the course of designing of systems in which neural network algorithms are applied, the special attention is given to the issues of their architecture construction. First of all, it is connected with the requirements shown to increase the efficiency and reliability of these systems' functioning.*

*Being possible to put the corresponding quantitative characteristic that has a meaning of criterion to certain properties of neural network classifier, it gives the possibility to speak about multicriteria optimization of this classifier. As a criterion, the probability of classification error and time necessary for a neural network training is chosen in this article at specified architecture. The presence is supposed for other criteria to characterize various properties of neural classifier. However, using only*

*these two basic criteria, it is supposed that the explained technique includes other properties of the qualifier taken into consideration.*

*Both chosen criteria are inconsistent, integral, of such a type that are possible to minimize and are limited. These preconditions allow a method of scalar convolution under nonlinear scheme of compromises to be used in objective function of criteria, and to show the search of optimum architecture on an example of a simplex-planning method.*

*The application of the proposed approach of vector optimization will allow to determine the architecture of neural network classifier where conflicting criteria of functional efficiency are systematically coordinated, and the received architecture will be optimal and have compromise solution. Also, the proposed method will provide the decision of one of the basic limitations of text qualifiers use which is connected with inefficient construction of its neural network architecture.*

**Keywords:** *neural network, vector optimization, quality criteria, neural network classifier, text.*

*Стаття надійшла до редакції 26.12.2015 р.*

*Рецензенти: Ю. К. Зіатдінов, д.т.н., професор,  
А. М. Воронін, д.т.н., професор*