

2. Проектирование статически определимой фермы минимального веса / [В.М. Кобринец, А.Н. Орлов, Н.Н. Сорока, Ю.А. Школа]. — Одесса : ОГАСА, 2006. — 148 с.
3. Лихтарников Я.М. Вариантное проектирование и оптимизация стальных конструкций / Я.М. Лихтарников. — М. : Стройиздат, 1979. — 320 с.
4. Овчинников И.Г. Математическое моделирование процесса взаимодействия элементов конструкций с агрессивными средами / И.Г. Овчинников, В.В. Петров // Деформирование материалов и элементов конструкций в агрессивных средах. — Саратов : СПИ, 1983. — С. 3–11.
5. Алексеенко Б.Г. Расчет и оптимальное проектирование рамных систем, подверженных коррозионному износу / Б.Г. Алексеенко // Математические методы в задачах расчета и проектирования сложных механических систем : сб. научн. трудов. — К., 1992. — С. 4–10.
6. Почтман Ю.М. Динамическое программирование в задачах строительной механики / Ю.М. Почтман, В.А. Бараненко. — М. : Стройиздат, 1975. — 110 с.
7. Алексейцев А.В. Эволюционная оптимизация стальных ферм с учетом узловых соединений стержней / А.В. Алексейцев // Инженерно-строительный журнал. — СПб, 2013. — №5. — С. 28–37.
8. Юрьев А.Г. Эволюционные и генетические алгоритмы оптимизации строительных конструкций / А.Г. Юрьев, С.В. Ключев. — Белгород : Изд-во БГТУ им. В.Г. Шухова, 2006. — 134 с.
9. Кирсанов М.Н. Генетический алгоритм оптимизации стержневых систем / М.Н. Кирсанов // Строительная механика и расчет сооружений. — М., 2010. — № 2. — С. 60–63.

Рецензія/Peer review : 19.3.2017 р.

Надрукована/Printed : 19.4.2017 р.

Стаття прорецензована редакційною колегією

УДК 681.327.12

М.М. БИКОВ, А.Д., ГАФУРОВА, В.В. КОВТУН
Вінницький національний технічний університет

ДОСЛІДЖЕННЯ КОМІТЕТУ НЕЙРОМЕРЕЖ У АВТОМАТИЗОВАНІЙ СИСТЕМІ РОЗПІЗНАВАННЯ МОВЦІВ КРИТИЧНОГО ЗАСТОСУВАННЯ

Стаття присвячена дослідженню процесу прийняття рішень у автоматизованій системі розпізнавання мовців критичного застосування у випадку необхідності узагальнення результатів розпізнавання від кількох класифікаторів. В результаті досліджень сформульовано та математично обґрунтовано метод і алгоритм синтезу комітету нейромережових класифікаторів для оптимізації процесу прийняття рішень автоматизованою системою розпізнавання мовців критичного застосування, який адаптовано до специфіки застосування критичних систем і дозволяє об'єднувати у комітет класифікатори різних типів. Для створення окремих нейромереж і комітетів автори застосували можливість крос-платформної бібліотеки чисельного аналізу ALGLIB, яка підтримує основні мови програмування. Результати тестування автоматизованої системи розпізнавання мовців критичного застосування із блоком прийняття рішень, який містив комітет нейромереж, за даними із бази еталонних записів NOISEUS емпірично доводять адекватність запропонованих теоретичних результатів.

Ключові слова: автоматизована система розпізнавання мовців критичного застосування, інформативні ознаки, комітет штучних нейронних мереж, розпізнавання мовців.

M.M. BYKOV, A.D. HAFUROVA, V.V. KOVTUN
Vinnitsia National Technical University

RESEARCH OF NEURAL NETWORK COMMITTEE IN THE AUTOMATED SPEAKER RECOGNITION SYSTEM OF CRITICAL USE

The article is devoted to the research of the decision-making process of automated speaker recognition system of critical use in a case of necessity of generalization of recognition results from several classifiers. In the result of research a method and a algorithm of neural network committee classifiers synthesis were formulated and mathematically proved to optimize decision making process for automated speaker recognition system of critical use that is customized to the specific application of critical systems and allows you to combine the classifiers committee of different types of neural networks. . To create both a separate neural networks and their committees, the authors used the possibility of cross-platform library of numerical analysis ALGLIB that supports basic programming languages. The results of testing of automated speaker recognition system of critical use with the decision making block that included the neural network committee based on data from the reference database NOISEUS empirically have proved the adequacy of the proposed theoretical results.

Keywords: automated speaker recognition system of critical use, information features, neural networks committee, speaker recognition.

Вступ

Структура автоматизованої системи розпізнавання мовця критичного застосування у першому наближенні містить блок аналого-цифрового перетворення мовного сигналу, блок попереднього його оброблення, блок виділення інформативних для розпізнавання мовця ознак і блок класифікації. Кожен із цих блоків має значну кількість параметрів, вплив яких визначає якість розпізнавання мовця автоматизованою системою, яку традиційно оцінюють імовірністю правильного розпізнавання та імовірністю виникнення

помилки першого та другого роду тестуючи новостворену систему на загальновідомих базах даних мовних сигналів. Втім, остаточне рішення при функціонуванні автоматизованої системи приймає класифікатор, вид, архітектура, параметри та методи навчання якого потребують дослідження і оптимізації.

В сучасних інформаційних системах застосовують різні комплексні методи прийняття рішень, які у інформаційних джерелах [1–7] називають сумішами експертів (mixture of experts), комітетами машин (committee machines), ансамблями класифікаторів (classifier ensembles), сумішами класифікаторів (classifier fusion) і т. ін.

З найбільш вживаних методів комплексної класифікації виділяють бегінг (bagging) [1, 2], бустінг (boosting) [3, 4, 6, 7] та стекінг (stacked generalization, stacking) [5].

Бегінг передбачає формування підмножин у навчальній вибірці довільним чином та їх використання для навчання різноманітних класифікаторів, узагальнюючи потім результати їх роботи із застосуванням обраної агрегувальної функції. Метод дозволяє підвищити якість функціонування системи розпізнавання яка включає різні типи класифікаторів та у випадку обмеженого розміру навчальної вибірки.

Бустінг являє собою жадібний алгоритм послідовного формування композиції алгоритмів класифікації, при якому кожен наступний алгоритм вибирається з метою компенсації недоліків його попередників. Цей принцип користується зараз великою популярністю з причини його універсальності та високої узагальнюючої властивості. Бустінг є найпопулярнішим способом оптимізації дерев класифікації. Втім, його ефективність супроводжується значними вимогами до якості та розміру навчальної вибірки та схильністю до перенавчання.

Стекінг – спосіб об'єднання класифікаторів із формуванням мета-алгоритму навчання, для синтезу якого навчальну вибірку розбивають на дві підмножини, елементи яких не перетинаються, далі навчають базові класифікатори на даних із першої підмножини та тестують на даних із другої підмножини і використовуючи результати тестування як вхідні дані, а реальні дані – як вихідні синтезують мета-алгоритм. Недоліком цього підходу є складність формалізації етапу синтезу мета-алгоритму та значна обчислювальна складність методу.

Узагальнюючи огляд існуючих методів побудови комплексних систем класифікації у задачі розпізнавання образів можна констатувати, що жоден з них не може бути безпосередньо використаним для оптимізації процедури класифікації у автоматизованій системі розпізнавання мовців критичного застосування, т.я. інформативні ознаки, що виділяються із мовного сигналу для подальшого розпізнавання мовця мають різне представлення – значення, множина, матриця і т.д., що вимагає використання класифікаторів різного типу, наприклад, загортальні нейромережі глибокого навчання і перцептронні, які за класичними методами у комітети не об'єднуються. Також специфіка критичних систем полягає у максимізації достовірності роботи системи на протязі життєвого циклу її функціонування, що також слід врахувати при генеруванні комітету класифікаторів.

Постановка завдання

Отже, метою дослідження є формулювання методу і алгоритму синтезу комітету нейромережних класифікаторів для оптимізації процесу прийняття рішень автоматизованою системою розпізнавання мовців критичного застосування, адаптованого до специфіки застосування критичних систем і дозволяє об'єднувати у комітет класифікатори різних типів.

Математичне обґрунтування використання комітетів нейромереж у автоматизованих систем розпізнавання мовців критичного застосування

Сформулюємо теоретичні засади формування комітету штучних нейромереж для класифікації у автоматизованій системі розпізнавання мовців критичного застосування. У найбільш загальному вигляді задачу автоматизованого розпізнавання мовців можна звести до знаходження математичної моделі $f(\cdot)$, яка найкращим чином описуватиме залежність імовірності правильного розпізнавання i -го мовця y_i від значень вектора інформативних ознак $\vec{x}_i = (x_{i1}, x_{i2}, \dots, x_{in})$ із врахуванням того, що мовець буде розпізнано ($m = 1$) або не розпізнано ($m = 0$):

$$y_i(\vec{x}_i) \equiv P(y_i = m) = f(\vec{b}^T, \vec{x}), m = 1; 0. \quad (1)$$

Відношення (1) описує модель бінарного вибору, описуючи яку припустимо, що $f(\cdot)$ – інтегральна нормалізована функція розподілу Лапласа, тоді (1) можна перетворити до виду (2):

$$y_i(\vec{x}_i) \equiv P(y_i = 1) = f(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{z^2}{2}} dz, \quad (2)$$

Із (2) видно, що значення аргументу u функціонально визначається значеннями вектору інформативних ознак $\vec{x} = (x_1, x_2, \dots, x_n)$, тобто $u = f(\vec{x})$. Ідентифікація цієї залежності є нетривіальною задачею, масштаб якої зростає із збільшенням розмірності вектора інформативних ознак \vec{x} та ускладнюється очевидною нелінійністю $u = f(\vec{x})$. Отже, ефективним видається використати математичний апарат штучних нейромереж для ідентифікації функціональної залежності $u = f(\vec{x})$. Структура вхідного та вихідного шарів нейромережі очевидна – на нейрони вхідного шару подаватимуться компоненти вектора

інформативних ознак \bar{x} , кількість елементів якого визначатиме кількість нейронів шару, а на вихідному шарі нейромережі буде один нейрон, який повертатиме значення аргументу u із (2). В процесі навчання штучної нейромережі маніпулюючи значеннями матриці синаптичних ваг W можна в неявному вигляді ідентифікувати залежність $u = f(\bar{x})$. При формуванні навчальної вибірки значення аргументу u вибератимемо з таблиці значень функції Лапласа:

$$u = \begin{cases} 3.9, & \text{якщо } y_i = 1, P(y_i = 1) = \Phi(u) = 1, \\ -3.9, & \text{якщо } y_i = 0, P(y_i = 0) = \Phi(u) = 0. \end{cases} \quad (3)$$

Використовуючи значення інформативних ознак із тестуючої вибірки $\bar{x}_{i_t} = (x_{i_t,1}, x_{i_t,2}, \dots, x_{i_t,n})$ на виході навченої нейромережі отримаємо ідентифіковані значення аргументу u_{i_t} , використовуючи які можна оцінити імовірність того, що мовця i розпізнано вірно:

$$P(y_{i_t} = 1) = f(u_{i_t}) \geq 0.5. \quad (4)$$

Слід враховувати, що мовний сигнал, і відповідно його індивідуальні особливості, мають комплексну динамічну природу, що призводить до присутності у навчальній вибірці екземплярів векторів \bar{x} , компоненти яких суттєво відрізняються від середніх показників для відповідного мовця. Видалення таких екземплярів із навчальної вибірки покращить її однорідність, що пришвидшить навчання нейромережі, проте може зменшити імовірність правильного розпізнавання мовця, що у системах критичного застосування є неприпустимим. Отже, виникає потреба у фільтрації екземплярів навчальної вибірки, яку здійснюватимемо так: вважатимемо $\langle \bar{x}_i, y_i \rangle$ екземпляр навчальної вибірки аномальним, якщо хоча б один компонент вектора \bar{x}_i виходить за межі інтервалу

$$x_{ij} \in [\bar{x}_j - 3\sigma_{x_j}; \bar{x}_j + 3\sigma_{x_j}], \quad (5)$$

де \bar{x}_j – середнє по вибірці значення інформативної ознаки x_j ; σ_{x_j} – середньоквадратичне по вибірці відхилення інформативної ознаки x_j відносно значення \bar{x}_j . Екземпляри навчальної вибірки, які не увійшли до інтервалу (5) переносяться із навчальної у тестуючу вибірку.

Узагальнюючи сказане, сформулюємо алгоритм синтезу класифікатора із застосуванням штучних нейромереж для автоматизованої системи розпізнавання мовців критичного застосування:

1. Формуємо альтернативні факторні простори, які відрізняються включеними інформативними для розпізнавання мовця ознаками: $I_k = \langle \bar{X}_j, Y \rangle$, $j = \overline{1, N_{I_k}}$, $k = \overline{1, L}$, де N_{I_k} – кількість інформативних ознак X_j у факторному просторі I_k , L – кількість альтернативних факторних просторів.

2. Будуємо штучні нейромережі, які навчаємо на даних, сформованих для визначених у п.1 факторних просторів. Проводимо тестування створених класифікаторів на тестових даних, використовуючи для оцінювання імовірності правильного розпізнавання правило (4) із фіксацією випадків правильного розпізнавання та виникнення помилок першого та другого роду. Ці дані узагальнюємо у вигляді критерію відбору найінформативнішого факторного простору в вигляді

$$K_{I_k} = N_{I_k}^+ r_1 - N_{I_k}^1 r_2 - N_{I_k}^2 r_3, \quad (6)$$

де $N_{I_k}^+$ – кількість правильно розпізнаних мовців для факторного простору I_k , $N_{I_k}^1$ та $N_{I_k}^2$ – кількість помилок першого та другого роду відповідно для факторного простору I_k , r_1, r_2, r_3 – вагові коефіцієнти відповідних якісних показників роботи автоматизованої системи критичного застосування, значення яких вибератимемо за правилом Фішберна [8]:

$$r_i = \frac{2(n-i+1)}{(n+1)n}, \quad (7)$$

де n – кількість інформативних ознак факторного простору I_k .

3. Створюємо комітет нейромережових класифікаторів із нейромереж одного типу, які можуть мати різну архітектуру шарів та їх кількість, а також різні функції активації нейронів. При навчанні нейромережового комітета використовуємо значення критерію (6). Для підвищення ефективності нейромережового комітета після його створення та навчання відбувається фільтрація, за результатами якої із комітету виключаються неефективні нейромережі. Фільтрацію пропонується проводити за правилом

$$K_{I_k} \geq \Delta, \quad (8)$$

де Δ – поріг фільтрації, значення якого встановлюється на основі критерію (6).

Проведемо емпіричне дослідження ефективності запропонованого алгоритму синтезу оптимального нейромережового комітету для автоматизованої системи розпізнавання мовців критичного застосування, взявши в якості бази еталонних записів безкоштовну базу даних NOIZEUS [9]. Ця спеціалізована база записів мовних сигналів сформована Школою інжинірингу та комп'ютерних наук Еріка Джонсона при

Університеті Техасу в Далласі (США) для дослідження алгоритмів покращення звуку і складається із 30 речень англійської розмовної мови, вимовлених трьома чоловіками та трьома жінками (по 5 на кожного мовця, частота дискретизації записів складає 25 кГц, але задля додавання шуму була зменшена до 8 кГц) та записів типових побутових та техногенних шумів, які можна додавати до чистого сигналу, щоб отримати записи мовних сигналів із рівнями шум/сигнал 0 дБ, 5 дБ, 10 дБ, 15 дБ відповідно. Для навчання нейромережових класифікаторів використовувалися 10% бази еталонних записів рівномірно розподілених за класами мовців та за співвідношенням шум/сигнал. Решта 90% еталонних мовних записів розділялися навпіл між контрольною та тестуючою вибірками.

При виборі множини інформативних для розпізнавання мовця ознак автори врахували специфіку процесу мовотворення (частота/період основного тону на вокалізованих фрагментах мовного сигналу, коефіцієнти лінійного прогнозу, спектр Фур'є), особливості сприйняття мови людиною (кепстральні коефіцієнти) та динамічними характеристиками мови. Кожна із цих ознак має як переваги так і недоліки при її використанні для розпізнавання мовця, залежно від способу проведення процедури запису мовних сигналів, емоційного стану мовця, параметрів навколишнього середовища, наявними обчислювальними ресурсами і т. ін., але в комплексі вони утворюють цілісне представлення особи мовця у факторному просторі.

Отже, для формування альтернативних факторних просторів із мовного сигналу виділяємо множину барк-кепстральних коефіцієнтів, які отримуємо аналогічно відомій методиці розрахунку кепстральних коефіцієнтів мовного сигналу [10, 11], виконуючи відображення дискретного спектру Фур'є за школою Барк-частот, перші 13 коефіцієнтів якої охоплюють спектральну область до 8 кГц, якої достатньо для розпізнавання мовця. Для врахування динамічних характеристик мовного сигналу екстрагуємо похідні від розрахованих барк-кепстральних коефіцієнтів, аналізуючи сусідні фрейми мовного сигналу збільшуючи кількість елементів вектору інформативних ознак до 39. Також до вектору інформативних ознак включено значення періоду основного тону для базового та сусідніх фреймів мовного сигналу, які визначили за авторським методом [12], що використовує автокореляційну функцію із постфільтрацією значень та аналіз концентрації фактів перетину амплітудних значенням мовного сигналу нульового рівня, та піки 2-4 формант, які отримуватимемо аналізуючи згладжений спектр мовного сигналу з розрахунком 10 коефіцієнтів лінійного прогнозу для кожного із фреймів. Загалом кількість елементів факторного простору дорівнює 72, які відповідно до п.1 алгоритму авторами було рівномірно розподілено по чотирьом альтернативним факторним просторам (АФП) АФП1 - АФП4 відповідно.

Для виконання п.2 та 3 алгоритму, які вимагали створення як окремих нейромереж так і їх комітетів, автори застосовували можливості крос-платформної бібліотеки чисельного аналізу ALGLIB [13], яка підтримує основні мови програмування (C++, C#, Pascal, VBA) і кілька операційних систем (Windows, Linux, Solaris). Програмний пакет ALGLIB дозволяє створювати нейронні мережі прямого поширення без прихованих шарів, з одним або двома прихованими шарами, які можуть мати одну зі стандартних стискаючих функцій активації. Вихідний шар може бути, зокрема, лінійним із SOFTMAX-нормалізацією виходів, який зручно застосовувати для задач класифікації, коли виходи мережі повинні бути невід'ємні і їх сума має дорівнювати одиниці, що дозволяє безпосередньо використовувати дані з вихідного шару для оцінювання результатів розпізнавання в імовірнісному виді.

У випадку реалізації комітетів нейромереж у ALGLIB є два обмеження – кількість мереж у комітеті визначається на етапі формування і не може змінюватися динамічно і комітет має бути однорідним, тобто нейромережі, що входять до його складу, мають бути архітектурно аналогічні.

Для навчання нейромереж у ALGLIB можна використати дві підпрограми – MLPEBaggingLM і MLPEBaggingLBFGS. Перша навчає індивідуальні нейромережі із допомогою алгоритму із сімейства Лівенберга-Марквардта, друга – із допомогою L-BFGS алгоритму, суттєвою перевагою останнього у випадку навчання комітетів нейромереж є те, що в результаті його роботи окрім безпосередньо навченого комітету також відбувається і внутрішнє оцінювання похибки узагальнення.

Відповідно до п.2 для обчислення критерію (6) було синтезовано базову нейромережу (трьохшаровий перцептрон), яка по чергово формувалася під класифікацію у кожному із чотирьох сформульованих вище факторних просторах.

Перед навчанням базової нейромережі навчальну вибірку досліджували на виявлення аномальних екземплярів інформативних ознак за правилом (5), які примусово включалися у тестуючу вибірку. Результати роботи базової нейромережі представлено у таблиці 1.

Аналіз даних з таблиці 1, в якій узагальнено результати оцінювання роботи базових нейромереж із чотирма факторними просторами із застосуванням критерію (6), дозволяє зробити висновок про більшу інформативність другого факторного простору у порівнянні з іншими, а також недоцільність включення аномальних екземплярів значень інформативних ознак у навчальну вибірку.

Таблиця 1

Аналіз результатів тестування штучних неймереж, сформованих для виконання п.2 алгоритму

Альтернативний простір	Кількість випадків правильного розпізнавання мовців	Частка правильно розпізнаних мовців, %	Всього помилок	Помилки 1-го роду	Помилки 2-го роду	К		
Без аналізу даних навчальної вибірки згідно (5)								
АФП1	171	89,06	1	2	15	6	9,53	7
АФП2	184	95,83	4	2	5	3	9,84	8
АФП3	178	92,71	4	1	9	5	5,18	8
АФП4	161	83,85	1	3	17	4	2,51	7
Після аналізу даних навчальної вибірки згідно з (5)								
АФП1	174	90,63	8	1	14	4	1,70	8
АФП2	185	96,35	4	2	4	3	0,67	9
АФП3	180	93,75	4	1	9	5	6,84	8
АФП4	164	85,42	1	1	18	0	4,36	7

На основі сформованої системи інформативних ознак створено комітет з десяти неймереж (перцептронів), дослідження ефективності яких для розпізнавання мовців із мовної бази NOIZEUS наведено у таблиці 2.

Оптимізація комітету неймереж проводилася за правилом (8) із застосуванням порогу фільтрації $\Delta = 0.98$, після чого у комітеті залишилося три неймережі (ШНМ2, ШНМ7, ШНМ9 відповідно).

Для кожного із восьми мовців із мовної бази NOIZEUS розраховано емпіричне середнє значення атрибуту \bar{u} , описаного у (3), за формулою

$$\bar{u} = \frac{1}{N_{NM}} \sum_{i=1}^{N_{NM}} u_i, \tag{9}$$

де N_{NM} – кількість неймереж, які пройшли процедуру фільтрації. Як видно із даних таблиці 2, узагальнений із допомогою (8) результат роботи комітету неймереж показує імовірність правильного розпізнавання 98,27%.

Таблиця 2

Результати тестування комітету неймереж

Штучні неймережі	Кількість випадків правильного розпізнавання мовців	Частка правильно розпізнаних мовців, %	Всього помилок	Помилки 1-го роду	Помилки 2-го роду
ШНМ1	168	88,64	24	15	9
ШНМ2	186	98,02	6	2	4
ШНМ3	181	95,07	11	7	4
ШНМ4	186	97,64	6	3	3
ШНМ5	183	96,32	9	5	4
ШНМ6	186	98,09	6	4	2
ШНМ7	187	98,30	5	2	3
ШНМ8	176	92,70	16	9	7
ШНМ9	187	98,48	5	3	2
ШНМ10	185	97,50	7	5	2
Після фільтрації згідно з п.3	187	98,27	5	2	3

Висновки

Отже, сформульовану вище мету дослідження авторами виконано, а отримані результати можна узагальнити так:

- Авторами створено та математично обґрунтовано алгоритм синтезу комітету нейромереж для поняття рішень в автоматизованій системі розпізнавання мовця критичного застосування. Проведені емпіричні дослідження довели адекватність отриманих теоретичних результатів.
- Проведені експериментальні дослідження із розпізнавання мовців доводять, що комітет нейромереж, сформований за авторським алгоритмом, показує меншу помилку узагальнення, ніж самостійна нейромережа оптимальної архітектури у виконанні задачі розпізнавання мовців для автоматизованої системи критичного застосування.
- Комітет нейромереж, навчений за допомогою L-BFGS алгоритму, показує меншу схильний до перенавчання, ніж традиційна нейронна мережа.
- Звичайні нейромережі потребують втручання дослідника для оптимізації архітектури (перебору параметрів нейромереж із одним/двома прихованими шарами і різним ступенем регуляризації), тоді як комітет нейромереж виконує оптимізацію автоматично – достатньо створити необхідну кількість нейромереж із одним прихованим шаром і надмірною кількістю нейронів в ньому (30–100).
- Процес навчання комітету нейромереж в кілька разів більш трудомісткій, ніж навчання звичайної нейромережі. Наприклад, комітет з десяти нейромереж навчається в 2-3 рази повільніше, ніж одна однотипна нейромережа.
- Потребує дослідження питання репрезентативності навчальної вибірки для комітету нейромереж – інколи результати розпізнавання мовців комітетом були гірші порівняно із результатами звичайних нейромереж з причини недонавченості комітету.

Література

1. Izenman A.J. Modern Multivariate Statistical Techniques // Springer Texts in Statistics. Springer Science&Business Media New York, 2013. – 632 p.
2. Breiman L. Bagging predictors // Machine Learning, 1996. – 24. – P. 123–140.
3. Freund Y. A Short Introduction to Boosting / Freund Y., Schapire R. E. // Journal of Japanese Society for Artificial Intelligence, 1999. – 14 (5). – P. 771–780.
4. Breiman L. Random Forests // Machine Learning, 2001. – 45. – P. 5–32.
5. Wolpert D.H. Stacked Generalization // Neural Networks, 1992. – 5. – P. 241–259.
6. Valiant L.G. A theory of the learnable // Communications of the ACM, 1984. – 27(11). – P. 1134–1142.
7. Schapire R. The strength of weak learnability // Machine Learning, 1990. – 5(2). – P. 197–227.
8. Фишберн П. Теория полезности для принятия решений / Фишберн П. – М. : Наука, 1978. – 352 с.
9. NOIZEUS: A noisy speech corpus for evaluation of speech enhancement algorithms. URL: <http://ecs.utdallas.edu/loizou/speech/noizeus/>.
10. Martinez J. Speaker recognition using Mel frequency Cepstral Coefficients (MFCC) and Vector quantization (VQ) techniques / J. Martinez, H. Perez, E. Escamilla, M.M. Suzuki // Electrical Communications and Computers (CONIELECOMP) 22nd International Conference on 27-29 Feb. 2012. – IEEE, 2012. – P. 248–251.
11. Khalil A.A. Efficient speaker identification from speech transmitted over bluetooth based system / Khalil A.A., Mostafa Saad E.S., El-Nabi M.A., Abd El-Samie F.E. // Computer Engineering & Systems (ICCES), 2013 8th International Conference on 26–28 Nov. 2013. – IEEE, 2013. – P. 190–193.
12. Биков М.М. Метод виділення основного тону на основі модифікованої математичної моделі слухової системи людини / М.М. Биков, В.В. Ковтун, А. Раїмі // Вісник Вінницького політехнічного інституту. – 2011. – № 5. – С. 130–135.
13. ALGLIB: a cross-platform numerical analysis and data mining library. URL: <http://www.alglib.net/>.

Рецензія/Peer review : 5.3.2017 р. Надрукована/Printed :19.4.2017 р.
Рецензент: д.т.н., проф. Бісікало О.В.