

Ю. С. Здітовецький¹
О. В. Бісікало¹
Ю. Ю. Іванов¹

ІНТЕЛЕКТУАЛЬНА ІНФОРМАЦІЙНА СИСТЕМА РОЗПІЗНАВАННЯ ТА АНАЛІЗУ СКЛАДУ ПРОДУКТІВ ХАРЧУВАННЯ

¹Вінницький національний технічний університет

Враховуючи довгострокові наслідки COVID-19, лікарями запропоновано дотримуватись певного режиму харчування, оскільки якісні продукти важливі для здорового способу життя, а також боротьби з різноманітними хронічними хворобами. Розвиток технологій спонукає виробників до збільшення прибутку та здешевлення виробництва, використовуючи ароматизатори та специфічні харчові домішки (Е-домішки), а також їхні синтетичні комбінації, кожна з яких має низку особливостей. Міжнародний збірник харчових стандартів Codex Alimentarius включає в себе перелік близько 500 видів оригінальних Е-домішок, які можуть бути природними, ідентичними природним або синтетиками. Низка таких домішок має негативний вплив на організм людини, частина з них не повністю вивчена, що потенційно формує ризик генетичних мутацій, а відповідно аутоімунного та канцерогенного ефектів у майбутньому. Саме тому для покупця певного продукту харчування важлива швидка ідентифікація складу продукту, аналіз відповідних адитивів у ньому, їхньої небезпеки онлайн біля вітрини з використанням мобільних пристроїв та мережі Інтернет.

Останнім часом сучасні розробки розглядаються через призму штучного інтелекту, відповідно такі моделі, які будуть орієнтовані на роботу з продуктами харчування, допоможуть підтримати прагнення людини до здорового харчування. У статті стисло описано розроблену інтелектуальну інформаційну систему, яка для досягнення поставленої мети застосовує моделі машинного навчання на наповненій базі даних з продуктами харчування, сканер бар-кодів з додатковою процедурою коригування, якщо він сильно пошкоджений, апарат регулярних виразів, метрик подібності тексту, а також систему рейтингування продуктів. Відповідний програмний додаток працює в трьох режимах на платформах iOS та Android: розпізнавання продукту, бар-кодів, аналіз складу та оцінювання товару. Програма дозволяє отримати інформацію про продукт, склад, перелік адитивів, наукову інформацію по ним, рейтинг «корисності» товару, порівняння його з продуктами-аналогами тощо.

Ключові слова: продукт харчування, склад продукту, Е-домішки, система розпізнавання, нейромережі, програмування, програмний додаток.

Вступ

В останні роки сфера охорони здоров'я набула дуже значного значення. Наприклад, багато людей у всьому світі страждає від довгострокових наслідків COVID-19, і лікарі пропонують відповідально підходити до здорового способу життя, щоб полегшити ці наслідки. Відповідно більша обізнаність про громадське здоров'я спонукала людей змінити свої харчові звички. Багато з них стурбовані довгостроковими харчовими наслідками для їхнього здоров'я, намагаючись додати натуральнішу їжу у свій раціон [1]. Моделі штучного інтелекту, орієнтовані на роботу з продуктами харчування, можуть підтримати прагнення людей до здорового харчування, що може допомогти уникнути або покращити симптоми низки хронічних захворювань [2], [3].

З розвитком технологій спостерігається значне додавання різноманітних харчових домішок або адитивів, Е-домішок (*Europe food additives*) у продукти харчування. Таких домішок у світі існує близько 500 видів, не враховуючи ароматизатори та комбіновані адитиви. Система нумерації з унікальними номерами кожної Е-домішки прийнята для публікації у міжнародному збірнику харчових стандартів Codex Alimentarius [4]. Домішки можуть бути природними, ідентичними природним або штучно синтезованими речовинами, які взаємодіють з іншими складовими, що може за-

грозувати нормальним метаболічним процесам у організмі людини, формувати генетичні мутації, спричиняти розвиток і ріст пухлин. Але, слід зазначити, що деякі харчові домішки мають корисні властивості, наприклад, дозволяють підвищити строк придатності продуктів, покращують смак та вигляд, не завдаючи шкоди [5]—[7].

Отже, необхідно систематично контролювати наявність адитивів у продуктах та особливості їхнього маркування на етикетках товарів. Оскільки запам'ятати перелік корисних та шкідливих речовин досить важко, то актуальною для покупця є задача швидкої ідентифікації *E*-домішок у продуктах харчування, а також розпізнавання їхньої небезпеки онлайн, тобто в магазині біля вітрини з використанням мобільних пристроїв та мережі Інтернет. Загальний вигляд інтелектуальної інформаційної системи для виконання таких дій показано на рис. 1 [8].



Рис. 1. Структура системи

Метою роботи є розробка інтелектуальної інформаційної системи розпізнавання харчових продуктів та *E*-домішок у них за маркуванням на етикетці, яка завдяки гнучкому поєднанню робочих режимів з використанням сформованої бази даних продуктів, неймереж зі сканером баркодів та процедурою їхньої корекції, а також аналізу з використанням метрик подібності тексту та низки спеціальних регулярних виразів для оцінки якості продукту харчування, дозволяє, на відміну від аналогів, значно розширити функціональні можливості, забезпечуючи користувача актуальною інформацією. Порівняння розробки з іншими відомими додатками наведено у табл. 1, в якій визначено кількість переваг (+) та недоліків (–) функціональності програм щодо роботи з продуктами харчування та аналізом інформації, отриманої від них.

Таблиця 1

Порівняння з додатками-аналогами

#	<i>AI in Nutrition</i> [8]	<i>Open Food Facts</i> [9]	<i>FoodVisor</i> [10]	<i>Сканер їжі</i> [11]	<i>Запропонована розробка</i>
Розпізнавання продуктів на вітрині	–	+	–	+	+
Розпізнавання страв	+	–	+	–	–
Формування дієти	–	–	+	–	–
Розпізнавання бар-кодів	–	+	–	+	+
Аналіз складу на <i>E</i> -домішки	–	+	+	–	+
Аналіз складу на калорійність	+	–	+	+	–
Наукова інформація	–	–	–	–	+
Система рейтингування та підбору аналогів	–	–	–	–	+
База продуктів	відсутня (–)	велика (+)	велика (+)	мала (–)	велика (+)
Веб-сайт	–	+	+	–	+
<i>Всього (+)</i>	2	5	6	3	7
<i>Рейтинг</i>	5	3	2	4	1

Результати роботи

Розпізнавання тексту та зображень є одним зі складних напрямків комп'ютерного зору та машинного навчання, як у теоретичному, так і у практичному сенсі. Але глибокі нейронні мережі здійснили революцію, змінивши підхід до розв'язання такого роду задач, дозволивши замінити каскад різноманітних алгоритмів однією нейромережею без значної втрати точності, а в більшості випадків, дозволяючи отримати рекордні результати [2], [3]. У розробці авторів для розв'язання поставленої задачі застосовано технології для взаємодії з нейронними мережами *Tensorflow* (архітектура *VGGNet*) та бібліотеку алгоритмів комп'ютерного зору *OpenCV*, а також сканер бар-кодів, апарат регулярних виразів та метрик подібності тексту з мови програмування *Python*. Слід зазначити, що дуже важливою та трудомісткою задачею з погляду витраченого часу є формування та наповнення бази зразків з описом відповідних об'єктів, в нашому випадку товарів, їхнього складу та відповідних бар-кодів, з яких найпоширенішими є *QR*-коди. Середній рейтинг додатку на основі більше 10 тисяч скачувань та 536 відгуків з використанням 5-бальної шкали оцінювання користувачами складає 4 бали. Оскільки розробка є комерційним проектом, то коротко опишемо математичну модель та основні особливості роботи програмного продукту.

Режим 1 (розпізнавання продуктів). Під навчанням нейромереж розглядається налаштування їхньої архітектури, ваг нейронів та зв'язків між ними. Пошук значення ваг синапсів, які з'єднують нейрони, є задачею багатовимірної глобальної оптимізації нелінійної, багатоекстремальної функції помилки (у нашому випадку середня сума квадратів відхилень) [12]:

$$E(w) = E(X, Y) = \frac{1}{n} \cdot \sum_{i=0}^m (y_i^t - y_i^r)^2 \rightarrow \min, w_i \in \{-\infty; +\infty\}, i = \overline{1, n}, \quad (1)$$

де X, Y — датасет (розміром n) прикладів для навчання та правильних значень; m — кількість виходів мережі (товари); y_i^t та y_i^r — правильне та отримане мережею значення.

Параметри нейромережі коригуються за допомогою детермінованого класу методів, заснованого на аналізі стану поточних параметрів, величинах входів, фактичних та бажаних значеннях сигналу на виходах мережі. Для оптимізації функції втрат $E(w)$ використано нову ефективну модифікацію алгоритму *ADAM*, запропоновану Тато–Нкамбу. Ідея алгоритму полягає у накопиченні значень градієнта, калібруванні моментів (математичного очікування m_t та дисперсії v_t) та врахуванні компоненти корекції (*look ahead + correction component*). Правило оновлення ваг синапсів w_{t+1} на новій ітерації $t+1$ з урахуванням інформації, отриманої на попередніх кроках, задається у такому вигляді [12], [13]:

$$w_{t+1} = w_t - \eta \Delta_t = w_t - \eta \frac{\beta_1^{(t)} m_{t-1} + (1 - \beta_1^{(t)}) g_t}{\left(\sqrt{\frac{\beta_2^{(t)} v_{t-1} + g_t^2 \cdot \text{sign}(v_{t-1} - g_t^2)}{1 - \beta_2^{(t)}} + \xi} \right) (1 - \beta_1^{(t)})} - \eta \psi_t; \quad (2)$$

$$\psi_{t+1} = \Delta_t \cdot \text{sign}(g_t) (1 - \beta_1^{(t)}), \quad (3)$$

де η — швидкість навчання; g_t — градієнт; $m_0 = v_0 = 0$; $\beta_1^{(0)} = 0,9$; $\beta_2^{(0)} = 0,999$; $\xi = 10^{-8}$.

Режим 2 (розпізнавання/коригування *QR*-кодів продуктів). Якщо програма рекомендує користувачу зчитати *QR*-код продукту камерою смартфона, то використовується вбудований додаток-сканер, який за стандартом ISO 18004 [14] використовує для коригування помилок вбудовані недвійкові циклічні коди Ріда–Соломона з алгоритмом декодування Берлекемпа–Мессі, пошуком положення помилок та оцінювання їхніх значень за Ченем та Форні [15], що дозволяє відновити інформацію, якщо пошкоджено до $R \leq 30\%$ матриці даних коду. У запропонованій розробці як додаткову процедуру коригування, коли сканувальному обладнанню не вдається зчитати та розпізнати *QR*-код звичайним чином (високий рівень шуму, пошкодження елементів-детекторів або матриці коду $R > 30\%$), застосовано підхід з використанням нейромережі Хеммінга–Ліппмана та ковзного вікна з роботи [16].

Режим 3 (аналіз складу та оцінювання товару). Для аналізу складу продукту використано апарат регулярних виразів [17] та метрик подібності тексту [18]. Програма дозволяє отримати деталь-

ну інформацію про сам продукт, його склад, перелік *E*-домішок (база синонімів на трьох мовах, планується подальше розширення), наукову інформацію про них, основні відомості, рейтинг «корисності» товару, який розраховується за спеціальною формулою з урахуванням відгуків користувачів, шкідливості/корисності складових, особливостей продуктів-аналогів тощо (рис. 2).

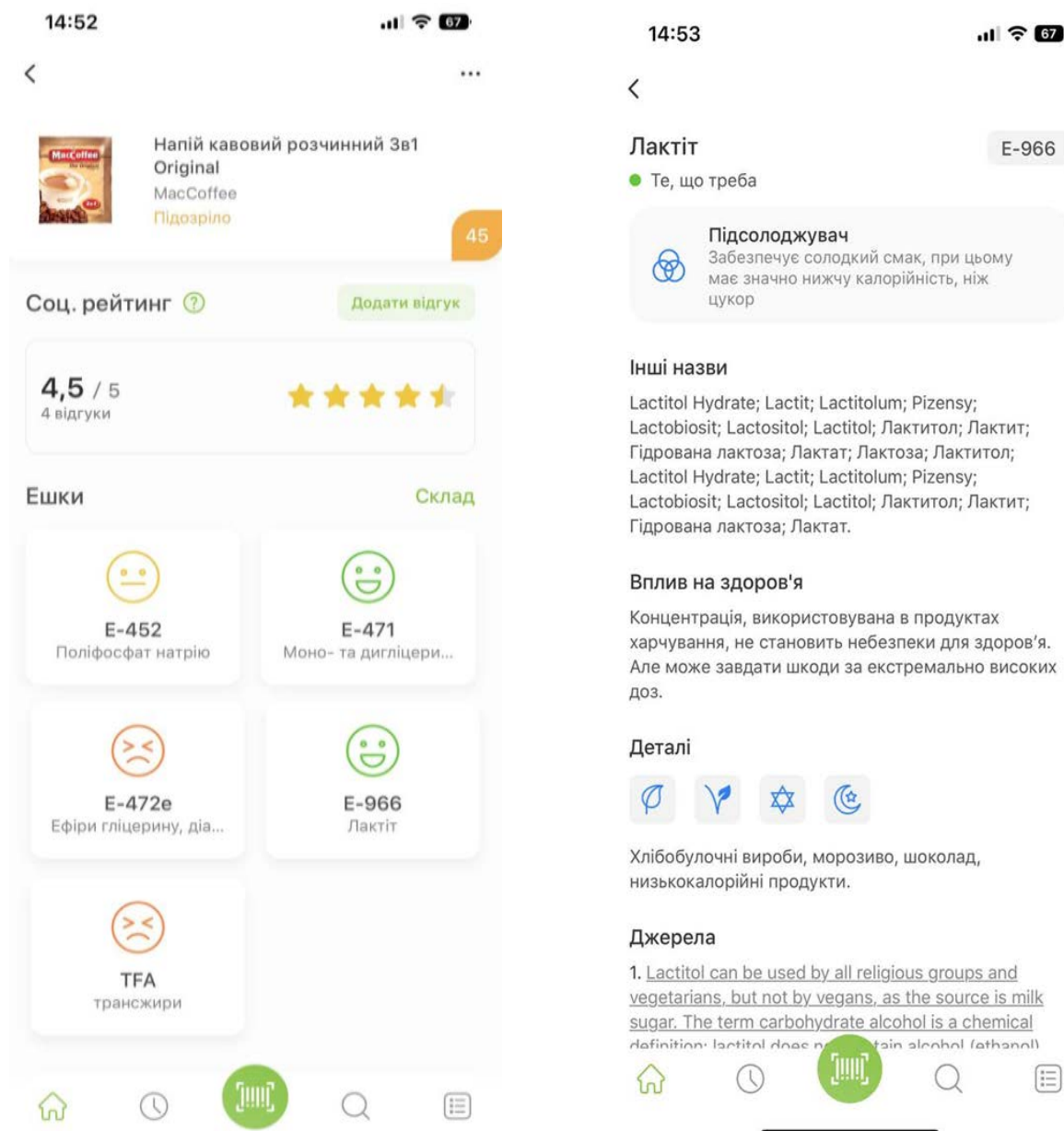


Рис. 2. Приклад роботи мобільного додатку

Кодування внутрішньої частини веб-сайту (*backend*) виконано на базі платформи з відкритим кодом *Node.js* на движку *Chrome V8*, яка дозволяє писати серверний код для веб-додатків та динамічних веб-сторінок [19]. Для розробки презентаційної частини та програмного інтерфейсу користувача (*frontend*) використано платформу *React Native*, яка дозволяє створювати мобільні програми з власним інтерфейсом для операційних систем *IOS* та *Android* [20].

Висновки

У статті стисло подано основні відомості про розроблені інтелектуальну інформаційну систему та відповідний програмний додаток (використовує сучасні технології машинного навчання, апарат регулярних виразів та метрики подібності тексту), який працює в трьох режимах: розпізнавання продуктів, їхніх *QR*-кодів, аналіз складу та розв'язання задачі онлайн-ідентифікації *E*-домішок, а також оцінювання товару. В подальшому планується доповнити базу даних для товарів, збільшити

кількість мов для аналізу речовин у складі продуктів харчування, а також додати можливість головного керування.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] А. В. Бабюк, та ін., *Безпека харчування: сучасні проблеми*. Чернівці, Україна: Книги-XXI, 2005, 454 с.
- [2] С. В. Нагірний, Я. Г. Бондарев, і Л. В. Нечволода, «Використання нейромережових технологій у системах розпізнавання образів для оцінювання безпечності продуктів харчування,» на *Міжнародній науковій конференції Комп'ютерні технології обробки даних*, 2020, с. 40-43.
- [3] T. B. Kumar, et al., “A Novel Model to Detect and Classify Fresh and Damaged Fruits to Reduce Food Waste Using a Deep Learning Technique,” *Journal of Food Quality*, pp. 1-8, 2022. <https://doi.org/10.1155/2022/4661108> .
- [4] ISO 22000 : 2005. *Системи управління безпечністю харчових продуктів – Вимоги до будь-яких організацій харчового ланцюга*. [Електронний ресурс]. Режим доступу: <http://www.codexalimentarius.net> . Дата звернення: 26.01.2022.
- [5] А. А. Дубініна, *Токсичні речовини у харчових продуктах та методи їх визначення*. Київ, Україна: Професіонал, 2007, 375 с.
- [6] В. І. Смоляр, «Сучасні проблеми використання харчових добавок,» *Проблеми харчування*, Київ, Україна: Інститут екології і токсикології імені Л. І. Медведя, с. 5-13, 2009.
- [7] О. Кратко, і М. Янків, «Вивчення небезпечного впливу продуктів харчування на здоров'я людини,» *Грааль науки*, № 1, с. 167-170, 2021.
- [8] K. Shiraly, “AI in Nutrition: How Technology is Transforming what We Eat,” [Electronic resource]. Available: <https://www.width.ai/post/ai-in-nutrition>. Accessed: 20.03.2023.
- [9] *Open Food Facts*. [Electronic resource]. Available: <https://play.google.com/store/apps/details?id=org.openfoodfacts.scanner>. Accessed: 20.03.2023.
- [10] *FoodVisor*. [Electronic resource]. Available: <https://www.foodvisor.io/>. Accessed: 20.03.2023.
- [11] *Сканер їжі*. [Electronic resource]. Available: <https://play.google.com/store/apps/details?id=food.scanner&hl=uk>. Accessed: 20.03.2023.
- [12] Yu.Yu. Ivanov, D. O. Kruts, and H. B. Rakytyanska, “An Algorithm for Training Artificial Neural Network Based on Adaptive Moments Estimation,” на *Міжнародній науково-практичній Інтернет-конференції Електронні інформаційні ресурси: створення, використання, доступ*, Суми / Вінниця, Україна: НІКО / ВНТУ, 2022, с. 117-119.
- [13] A. A. N. Tato, and N. Nkambou, “Improving ADAM Optimizer,” in *International Conference on Learning Representations*, 2018, pp. 1-4.
- [14] ISO/IEC 18004 : 2015. Estonian Centre for Standardization. [Electronic resource]. Available: <https://www.evs.ee/products/iso-iec-18004-2015>. Accessed: 26.01.2023.
- [15] T. K. Moon, *Error Correction Coding: Mathematical Methods & Algorithms*, John Wiley & Sons, 2005, 750 p.
- [16] V. V. Pivoshenko, and Yu.Yu. Ivanov, “Method for Recognition Highly Corrupted Barcodes,” на *VIII Міжнародній конференції з оптико-електронних інформаційних технологій “Photonics-ODS”*, Вінниця, Україна: ВНТУ, 2018, с. 55.
- [17] F. Lopez, and V. Romero, *Mastering Python Regular Expressions*. Packt Publishing, 2014, 110 p.
- [18] Yu. S. Zditovetskiy, O. V. Bisikalo, and Yu. Yu. Ivanov, “Overview and Numerical Examples of the Text Similarity Metrics,” in *Scientific and practical international conference “Science and Technology: Problems, Prospects and Innovations,”* Osaka, Japan: CPN Publishing Group, 2022, pp. 184-188.
- [19] F. Zammetti, *Modern Full-Stack Development: Using TypeScript, React, Node.js, Webpack, and Docker*. APress, 2020, 396 p.
- [20] A. Yudin, *Building Versatile Mobile Apps with Python and REST: RESTful Web Services with Django and React*. APress, 2020, 364 p.

Рекомендована кафедрою автоматизації та інтелектуальних інформаційних технологій ВНТУ

Стаття надійшла до редакції 14.03.2023

Здітовецький Юрій Сергійович — аспірант кафедри автоматизації та інтелектуальних інформаційних технологій, e-mail: yura.zditovetskij@gmail.com ;

Бісікало Олег Володимирович — д-р техн. наук, професор, в. о. завідувача кафедри автоматизації та інтелектуальних інформаційних технологій, e-mail: obisikalo@gmail.com ;

Іванов Юрій Юрійович — канд. техн. наук, доцент, доцент кафедри автоматизації та інтелектуальних інформаційних технологій, e-mail: Yura881990@i.ua .

Вінницький національний технічний університет, Вінниця

Yu. S. Zditovetskyi¹
O. V. Bisikalo¹
Yu. Yu. Ivanov¹

Intellectual Information System for Recognition and Food Product Composition Analysis

¹Vinnitsia National Technical University

Taking into account the long-term COVID-19 effects, doctors have suggested a specific diet, as quality foods are important for a healthy lifestyle, as well as fighting with various chronic diseases. Advances in technology encourage manufacturers to increase profits and reduce production costs by using flavorings and specific food additives (E-additives), as well as their synthetic combinations, each of which has a number of features. The international collection of food standards Codex Alimentarius includes a list of about 500 types of original E-additives, which can be natural, identical to natural or synthetic. A number of such additives have a negative effect on the human organism, some of them have not been fully researched, that potentially creates the risk of genetic mutations and, accordingly, autoimmune and carcinogenic effects in the future. That is why it is important for the buyer of a certain food product to quickly identify the product composition, analyze the relevant additives in it, their dangers “online” at the storefront using mobile devices and the Internet.

Recently, modern developments have been considered through the prism of artificial intelligence, accordingly, such models, which will be oriented to work with food products, will help support people’s desire for healthy food. This article briefly describes the developed intelligent information system, which, to achieve the goal, applies machine learning models on a database with food products, a barcode scanner with an additional correction procedure, if it is highly damaged, a regular expression apparatus, a text similarity metric, and also a product rating system. The corresponding software application works in three modes on the iOS and Android platforms: product recognition, barcodes recognition, composition analysis and product evaluation. The program allows buyer to get information about the product, its composition, an additives list, scientific information on them, the rating of the product “usefulness”, its comparison with similar products, etc.

Keywords: food product, product composition, E-additives, recognition system, neural networks, programming, software application.

Zditovetskyi Yurii S. — Post-Graduate Student of the Chair of Automation and Intellectual Information Technologies, e-mail: yura.zditovetskij@gmail.com ;

Bisikalo Oleh V. — Dr. Sc. (Eng.), Professor, Professor of the Chair of Automation and Intelligent Information Technologies, e-mail: obisikalo@gmail.com ;

Ivanov Yurii Yu. — Cand. Sc. (Eng.), Associate Professor, Associate Professor of the Chair of Automation and Intellectual Information Technologies, e-mail: Yura881990@i.ua