

УДК 637.07:637.068

А. Калініченко, наук. співроб., asya.kalini4enko@gmail.com,  
Л. Арсеньєва, д-р техн. наук,  
В. Пасічний, д-р техн. наук,  
Національний університет харчових технологій, Київ

## ВИКОРИСТАННЯ ЕЛЕКТРОННОГО НОСА ТА ЙМОВІРНІСНОЇ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ ІДЕНТИФІКАЦІЇ КОВБАС

*Запропоновано спосіб ідентифікації ковбасних виробів та визначення масового вмісту в них соєвих продуктів із використанням електронного носа та ймовірнісної нейронної мережі. Оптимізовано архітектуру нейронної мережі для дискримінації зразків за параметрами електронного носа:  $S_i$  (надійність класифікації становить 95,8% за значень параметра згладжування  $\sigma=3,6-54,0$ ) та  $\Delta F_i^{max}$  (100% точність ідентифікації за значень  $\sigma=0,2-1,0$ ).*

*Ключові слова: електронний ніс, ймовірнісна нейронна мережа, ідентифікація, варені ковбаси.*

**Вступ.** Для встановлення ідентичності та відмінностей у складі харчових продуктів широко застосовують мультисенсорну систему з методологією електронний ніс. Методологія системи – інтегральна оцінка складу об'єкта аналізу без попереднього розділення суміші з використанням масиву сенсорів з перехресною чутливістю до окремих сполук або класів речовин.

Через високу вартість якісної м'ясної сировини варені ковбасні вироби зазнають фальсифікації. Виробник замінює м'ясну сировину соєвими продуктами, зазвичай не вказуючи масову частку заміни [1]. До основних летких маркерів соєвих продуктів, які формують специфічний соєвий запах, відносяться: метантиол, диметилтрисульфід, гексаналь, 1-гексанол, 2-пентилфуран, 2-пентилпурин, 2,3-бутандіон, 1-октен-3-ол, транс-2-4-декадіеналь, транс-2-4-нонадіеналь, транс-2-ноненаль, ацетофенон [2, 3]. Більшість із цих сполук є метаболітами мікроорганізмів, притаманних соєвим продуктам [4].

Соєві продукти (ізолят, концентрат), які не містять у своєму складі крохмалю, ідентифікують гістологічним методом [5], ПЛР-аналізом послідовності промотору 35S, присутність якої свідчить про наявність генетичної модифікації геному сої [6], методом електрофорезу [7], імунохроматографічним аналізом. Недоліками цих методів є багатостадійна підготовка проби, складне апаратне оформлення, дорогі реактиви або навіть неможливість визначення масового вмісту соєвих компонентів. На сьогодні відсутні експрес-способи та методики визначення вмісту соєвих продуктів у м'ясних виробках, які необхідні для ефективного моніторингу товарів і вилучення фальсифікованої продукції з обігу.

**Мета роботи** – оцінити можливість дискримінації модельних зразків варених ковбасних виробів з різним вмістом соєвого заміни за рівноважною газовою фазою над ними з використанням електронного носа та ймовірнісної нейронної мережі, надати рекомендації з налаштування ймовірнісної штучної нейронної мережі (Probabilistic neural network, PNN) для класифікації та ідентифікації зразків за вихідними параметрами масиву п'єзосенсорів.

**Матеріали й методика дослідження.** Дослідження проводили на аналізаторі газів з методологією п'єзоелектронний ніс. Матрицю формували із семи сенсорів на базі п'єзокварцових резонаторів РК-169 з основною частотою коливання 10.0 МГц на срібних електродах діаметром 5 мм, які модифікували сорбційними покриттями масою 10–20 мкг. Як сорбційні покриття обрано хроматографічні фази, різні за полярністю, селективні щодо індивідуальних тест-речовин і груп летких сполук соєвих та м'ясних продуктів: поліетиленгліколь себацінат (PEG seb – сенсор 1), поліетиленгліколь адипінат (PEG ad – сенсор 2), дициклогексано-18-краун-6 (DCG18k6 – сенсор 3), Triton X-100 (TX-100 – сенсор 4), поліетиленгліколь 2000 (PEG-2000 – сенсор 5), полідіетиленгліколь сукцинат (PDEG suc – сенсор 6), полівінілпіролідон (PVP – сенсор 7).

Досліджували динаміку зміни якісного та кількісного складу рівноважної газової фази (РГФ) над модельними зразками вареної ковбаси "Лікарської" вищого сорту (в. с.), яка не містить рослинного білкового заміни, і зразками із вмістом соєвого ізоляту 10 (1 с.), 20 (2с.), 30 мас.% (фальсифіковані ковбаси). Зразки зберігали за температури  $20 \pm 1$  °C протягом 4 діб. Відбирали рівноважні пари над пробами масою  $5 \pm 0,1$  г методом дискретної газової екстракції. Індивідуальним шприцом відбирали 3 см<sup>3</sup> РГФ із пробовідбірника об'ємом 60 см<sup>3</sup> і вводили в комірку детектування електронного носа, час вимірювання становив 60 с. У комірку детектування леткі пари розбавляються в 60 разів. За таких умов детектуються мікроконцентрації летких сполук у РГФ, масова чутливість п'єзокварцового мікрозважування для досліджуваних систем залишається постійною в широкому діапазоні розбавлення.

У програмному забезпеченні MAG-Soft реєстрували два вихідні параметри статичних (steady-state) відгуків сенсорів: площі під кривими одиночних сенсорів протягом вимірювання, ( $S_i$ , Гц·с); максимальні зміни частот коливання сенсорів за час вимірювання ( $\Delta F_i^{max}$ , Гц). Параметри перевірені на статичну надійність ( $P = 0,95$ ,  $N = 3$ ).

Для реалізації алгоритмів використовували програмний пакет Matlab R2015b.

**Результати та їх обговорення.** Для визначення вмісту соєвих компонентів у варених ковбасних виробках та класифікації зразків за двома вихідними параметрами електронного носа використовували ймовірнісну штучну нейронну мережу (Probabilistic neural network, PNN).

Досліджуваний багатовимірний масив даних від семи сенсорів містив інформацію про 48 зразків варених ковбасних виробів, поділених на 4 класи (групи) залежно від вмісту соєвого ізоляту (CI):

- 1 клас – зразки вареної ковбаси "Лікарської" в. с.,
- 2 клас – зразки з 10 мас.% CI,
- 3 клас – зразки із 20 мас.% CI,
- 4 клас – зразки із 30 % CI.

Класифікаційний алгоритм має бути робастним, тобто нечутливим до впливу викидів, відхилень та неоднорідностей у вибірці. Отже, оцінювали стійкість алгоритму до варіювання значень, формуючи вибірку із зразків різної якості для врахування зміни газової фази протягом зберігання ковбасних виробів.

Ідентифікацію зразків за відгуками сенсорів електронного носа необхідно проводити за такою методологічною схемою: формування матриці аналітичних сигналів; оцінювання структури даних; визначення, що вважати шумом, а що інформацією; підготовка даних та обробка сигналів; побудова ефективної класифікаційної моделі; перевірка (валідація) моделі; прогнозування надійності класифікаційної моделі на нових даних або тестовому наборі.

Формували дві матриці даних із вихідних параметрів електронного носа для порівняння їхньої ефективності

дискримінації зразків із використанням PNN. Перша матриця включає максимальні зміни частот коливання сенсорів протягом вимірювання ( $\Delta F^{max}$ ). Даний параметр характеризує ефективність сорбції летких сполук на досліджуваному покритті сенсора. Друга матриця складається із площ під кривими сигналів сенсорів.  $S_i$  є аналітичним сигналом, який відображає кінетику сорбції летких сполук на досліджуваному покритті сенсора (дифузю до поверхні плівки та взаємодію з модифікатором) і характеризує весь відгук (період адсорбції, фазу рівноваги (плато) і десорбцію). Відмінності форми статичних відгуків сенсорів обумовлено рівнем чутливості сенсорів до даної сполуки або груп речовин.

Відповідно до методу аналізу даних і поставленого завдання "сирі" експериментальні дані можуть потребувати попереднього оброблення, спрямованого на підвищення ефективності функціонування нейронної мережі за рахунок приведення даних у необхідний формат.

Слід відзначити, що використання необхідної попередньої обробки даних залежить як від виду та структури даних, так і від архітектури нейронної мережі. У деяких випадках для стабілізації алгоритму нейронної мережі процедура нормалізації вхідних векторів включена до алгоритму (наприклад, у нейронній мережі прямого поширення сигналу в Matlab) [8]. Одним із завдань було дослідити необхідність попередньої обробки даних електронного носія для найкращої їх апроксимації із використанням імовірнісної нейронної мережі.

У роботі використовували три види оброблення та трансформування даних: нормалізацію, автошкалювання та оброблення методом головних компонент (МГК).

Для тренування нейронної мережі буває корисною нормалізація вхідних векторів і цілей (класів) таким чином, щоб вони потрапляли до діапазону [-1, 1]. Матриці нормалізуються за мінімальним і максимальним значеннями по кожному рядку в  $[y_{min}, y_{max}]$ .

$$y = \frac{(y_{max} - y_{min})(x - x_{min})}{(x_{max} - x_{min}) + y_{min}} \quad (1)$$

Автошкалювання (z-score) – процес перетворення матриці таким чином, що стандартне відхилення величин за рядками дорівнює 1, а середнє значення 0 ( $y_{mean}$  та  $y_{std}$ ) [8].

$$y = \frac{(x - x_{mean})y_{std}}{x_{std}} + y_{mean} \quad (2)$$

Отже, одні змінні не будуть превалювати над іншими й мати домінуючий вплив на модель через порядок величини.

МГК одночасно розв'язує дві важливі задачі: переведення даних у більш придатну систему координат (початок якої лежить у центрі області даних) і зменшення розмірності (використання тільки необхідної кількості головних компонент (ГК), які відображають структуру даних).

МГК здійснює декомпозиції матриці  $X$  на структурну (добуток двох матриць  $T$  та  $P$ ) та шумову ( $E$ ) [9]:

$$X = TP^t + E = \sum_{k=1}^K t_k p_k^t + E \quad (3)$$

Зазвичай  $X$  – це матриця не вихідних даних, а даних після центрування,  $T$  – матриця рахунків (scores),  $P^t$  – транспонована матриця навантажень (loadings),  $E$  – матриця залишків (помилки). Число стовбців –  $t_k$  у матриці  $T$  та  $p_k$  у матриці  $P$  дорівнює ефективному рангу матриці  $X$ . Ця величина  $K$  називається числом головних компонент (ГК), яких зазвичай менше, ніж кількість стовбців у матриці  $X$ .

Кількість головних компонент, необхідних для моделювання даних, вибирають за відсотком поясненої ними дисперсії, однак не завжди потрібно прагнути 100%, оскільки в такому випадку можна включити шумову складову.

Неефективним є використання лінійних методів, якщо дані є суттєво нелінійними. У таких випадках доцільно використовувати нелінійні функції та методи оброблення даних, такі як нейронні мережі. У даній роботі – це використання матриці рахунків як вхідних даних для нейронної мережі. У такому випадку попереднє оброблення даних методом МГК дає змогу отримати швидко в навчанні нейронної мережі зі зменшеною кількістю вхідних нейронів.

Імовірнісна нейронна мережа (Probabilistic neural networks, PNN) є надійним та швидким алгоритмом класифікації з навчанням. PNN характеризується простою архітектурою і складається з чотирьох шарів: вхідного, двох проміжних (шару зразків та підсумовуючого) і вихідного шару нейронів (рис. 1). У програмному пакеті Matlab проміжні шари називаються радіально-базисний і конкуруючий відповідно [8].

Вхідний шар мережі не виконує розрахунків та слугує для прийому й передачі вхідних векторів наступному проміжному шару нейронів. Кількість нейронів вхідного шару дорівнює кількості вхідних параметрів електронного носія ( $p=7$ ).

Шар зразків містить по одному нейрону на кожний зразок навчальної вибірки. У проміжному шарі PNN, що реалізовано в Matlab, використовується радіально-базисна функція (4) як функція ядра для обчислення імовірного значення тестового вектора  $X$  (нового вхідного зразка) [8, 10]:

$$f_k(X) = \sum_{i=1}^{m_k} e^{-\frac{(x-x_{ki})^T(x-x_{ki})}{2\sigma^2}} \quad (4)$$

де  $X$  – тестовий вектор із сімома параметрами,  $f_k(X)$  – значення імовірності  $X$  для  $k$ -го класу,  $X_{ki}$  –  $i$ -те спостереження  $k$ -го класу з навчального набору даних,  $k$  – [1:4], оскільки досліджуваний масив поділяли на чотири класи залежно від вмісту соєвого ізоляту,  $m_i$  – число спостережень, пов'язаних із класами залежно від вмісту соєвого ізоляту в навчальному наборі даних,  $\sigma$  – значення відхилення Гауссової функції або ступеня згладжування, що визначається експериментально.

Підсумовуючий шар містить чотири нейрони, що еквівалентно кількості класів (категорій). За допомогою функції зваженої суми розраховується імовірність приналежності вхідного вектора до того чи іншого класу.

Вихідний нейрон виконує функції дискримінатора порогового значення. За допомогою функції  $argmax$  визначається нейрон підсумовуючого шару з максимальним вихідним сигналом і виводиться клас, до якого належить представлений вхідний образ.

Отже, PNN дозволяє виразити нелінійні комбінаційні ефекти всіх вхідних параметрів як єдину функцію, яка деталізує зміни складу харчових зразків. Це досягається оптимізацією значення відхилення функції Гаусса, що визначає ширину "дзвону" з центром у кожному кластері та ступінь інтерполяції між точками.

Навчання будь-якої нейронної мережі починається з формування навчальної та тестової вибірок. Досліджуваний масив даних випадковим чином поділяли на навчальну та тестову вибірки різного співвідношення. Під оптимальним об'ємом навчальної вибірки розуміли таке число зразків, яке забезпечувало максимальну надійність класифікації зразків тестової вибірки. Коефіцієнт  $L$ , %, показує, яка частка зразків від їхньої загальної кількості міститься в навчальній вибірці:

$$L = \frac{H}{M} \cdot 100, \quad (5)$$

де  $H$  – кількість зразків навчальної вибірки,  $M$  – загальна кількість зразків.

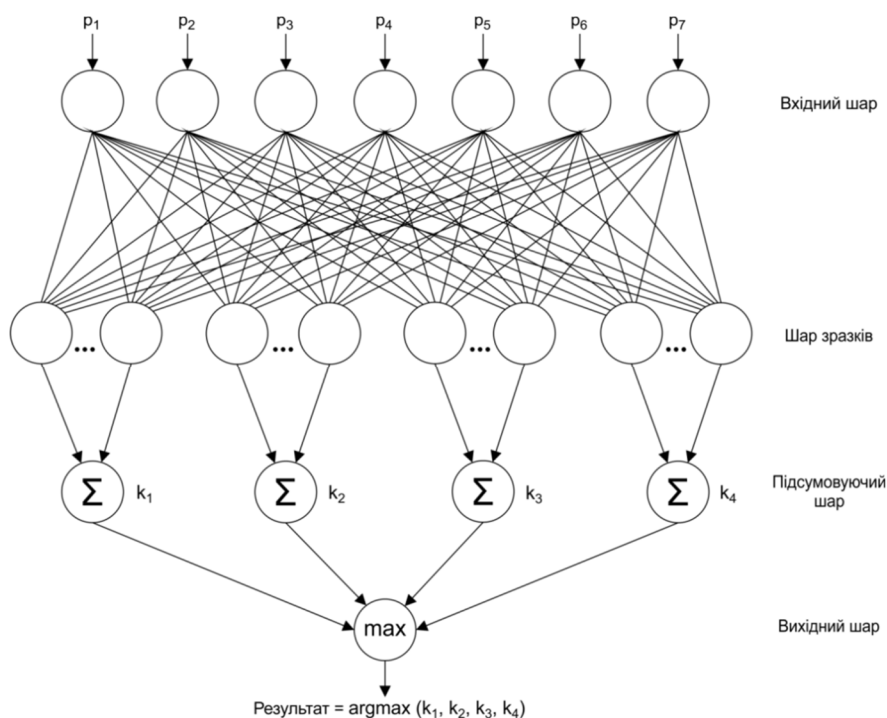


Рис. 1. Архітектура ймовірнісної нейронної мережі для класифікації чотирьох категорій зразків за вихідними параметрами електронного носа

У роботі ненадійність класифікації оцінювали як частку невірно класифікованих зразків тестової вибірки, % [11]:

$$U = \frac{t}{T} \cdot 100 \quad (6)$$

де  $t$  – кількість невірно класифікованих зразків тестової вибірки,  $T$  – загальна кількість зразків тестової вибірки.

Оптимальний і представницький обсяг навчальної вибірки для вихідних параметрів електронного носа ( $S_i$  та  $\Delta F_i^{max}$ ) становив 75 %, тестової вибірки – 25 %.

Надійність класифікації оцінювали як:

$$R = 100 - U \quad (7)$$

Отриману модель перевіряли методом перехресної 25-кратної кросвалідації. Далі під надійністю класифікації розуміли середнє значення надійності класифікації для чотирьох різних пар навчальної та тестової вибірок.

Параметрична оптимізація ймовірнісної нейронної мережі полягає у визначенні параметра згладжування  $\sigma$ , що контролює коефіцієнт масштабування експоненціальної функції активації. Значення  $\sigma$  має бути однаковим для кожного еталонного нейрона.

У лімітуючих випадках межа розділення чотирьох розподілів безперервно змінюється від гіперплощини, що

враховує декілька розташованих поруч векторів (рішення узгодженого фільтра з обчисленим пороговим значенням), коли  $\sigma \rightarrow \infty$ , до дуже нелінійної границі, що являє собою  $k$ -класифікатор найближчих сусідів, коли ступінь згладжування  $\sigma \rightarrow 0$  [12]. Загалом граничні випадки не забезпечують оптимального розділення розподілень.

Однак експериментальним шляхом не складно визначити оптимальне значення  $\sigma$ , оскільки швидкість неправильної класифікації змінюється поступово за невеликих змін  $\sigma$ . Рекомендують обирати значення ступеня згладжування як функцію розмірності вимірювального простору та кількості зразків навчальної вибірки [13].

За налаштуванням в алгоритмі PNN Matlab  $\sigma = 0,1$ .

Установлено, що найвища надійність дискримінації тестових зразків варених ковбас (95,8%) за вихідним параметром електронного носа  $S_i$  без попереднього оброблення може бути отримана за будь-яких значень  $\sigma$  в достатньо широкому інтервалі від 3,6 до 54,0. Крім того, спостерігаємо незначне погіршення точності класифікації зразків за значень  $\sigma$  в інтервалах від 2,4 до 3,5 та від 54,1 до 100,0 (рис. 2).

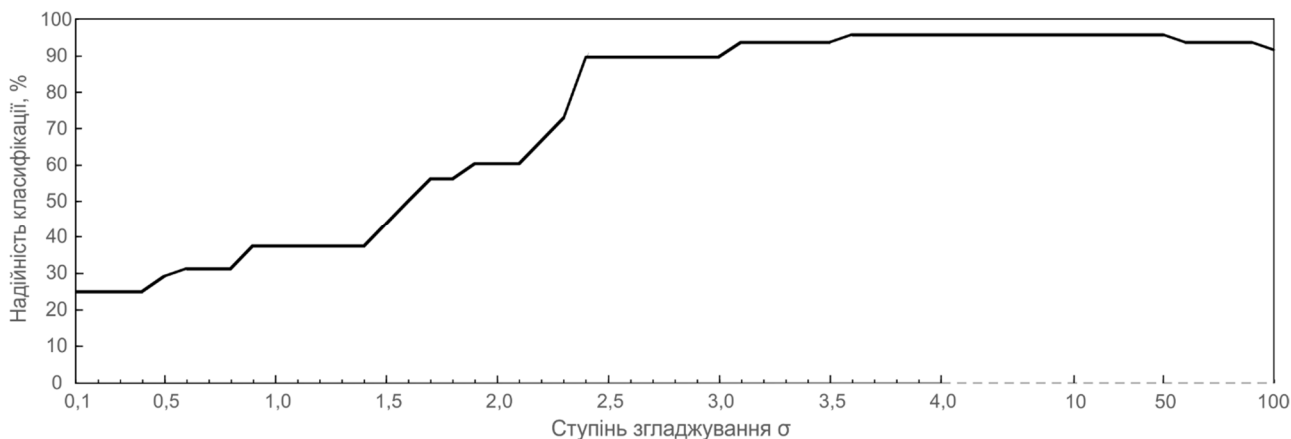


Рис. 2. Залежність надійності класифікації тестових зразків за параметром  $S_i$  від зміни ступеня згладжування  $\sigma$

З наведених у табл. 1 результатів можна зробити висновки, що використання процедур попередньої обробки даних не завжди є виправданим, оскільки зашумлені змінні можуть здійснювати не виправдано великий вплив. У випадку, коли вибіркові значення дисперсії практично однакові для всього набору змінних, необхідність попередньої обробки даних шляхом нормалізації

або автошкалування відсутня. Нормалізація за мінімальним та максимальним сигналами змінює не структуру, а вагу різних частин даних. Наприклад, значення параметра площі під кривою ( $S_i$ ) найбільш полярного сенсора з покриттям PVP під час нормалізації за рядками приймається як максимальне значення інтервалу 1, і його суттєвий вплив на дискримінацію зразків нівелюється, за рахунок чого точність класифікації знижується.

Таблиця 1

Результати ідентифікації зразків варених ковбас з використанням параметра  $S_i$  електронного носа та імовірнісної нейронної мережі

Попередня підготовка даних	Відхилення функції $\sigma$	Ненадійність класифікації, U, %	Надійність класифікації, R, %
Вихідні дані	3,6–54,0	4,2	95,8
Нормалізація	0,1–0,2; 0,9	8,3	91,7
Масштабування	0,1	6,2	93,8
МГК (центрування, 3 ГК), 98 % ПД	3,0–3,1	6,3	93,7
	3,4–3,8	6,2	93,8
МГК (центрування, масштабування, 4 ГК), 92% ПД*	0,1; 0,4–0,7	6,2	93,8

\*ПД – пояснювальна дисперсія

Метод головних компонент передбачає лінійну структуру даних, яка в реальних випадках не спостерігається. У зв'язку з цим, застосування алгоритму завжди пов'язано з втратою дисперсії даних внаслідок їх лінеаризації, що стає причиною обмеженого числа факторних

ознак, що включаються в модель. Однак попереднє оброблення даних МГК дозволяє знизити їх розмірність з семи векторів до трьох (ПД=98–99%) та збільшити швидкість навчання нейронної мережі.

Таблиця 2

Результати ідентифікації зразків варених ковбас з використанням параметра  $\Delta F_i^{\max}$  електронного носа та імовірнісної нейронної мережі

Попередня підготовка даних	Відхилення функції $\sigma$	Ненадійність класифікації, U, %	Надійність класифікації, R, %
Вихідні дані	0,2–1,0	0	100
Нормалізація	0,1–1,0	4,2	95,8
Масштабування	0,7–0,8	2,1	97,9
МГК (центрування, 3ГК), 99 % ПД	0,2–1,0	4,2	95,8
МГК (центрування, масштабування, 5 ГК), 96% ПД	0,1–0,7	4,2	95,8

Максимальну надійність (100%) результатів класифікації зразків варених ковбас за параметром  $\Delta F_i^{\max}$  отримали в інтервалі значень  $\sigma$  від 0,2 до 1,0. Використання попереднього оброблення та трансформації даних незначно погіршило точність ідентифікації зразків, однак збільшилася швидкість навчання нейронної мережі.

**Висновки.** Сформовано та досліджено вибірку модельних зразків варених ковбас без додавання соєвого ізоляту та із вмістом 10, 20, 30 мас. % під час їх зберігання за змін складу газової фази. Оброблення такої неоднорідної вибірки дозволило побудувати робастні, не чутливі до відхилень, класифікаційні моделі.

Проведено параметричну оптимізацію імовірнісної нейронної мережі шляхом визначення інтервалів значень параметра згладжування  $\sigma$  під час апроксимації матриць даних електронного носа. Проведено порівняльне оцінювання надійності класифікаційних моделей, побудованих під час використання різних за фізичним змістом вихідних параметрів електронного носа ( $S_i$  та  $\Delta F_i^{\max}$ ) та імовірнісної нейронної мережі. Встановлено, що найбільш точну модель (100%) отримано із використанням семи параметрів  $\Delta F_i^{\max}$  як вхідних векторів нейронної мережі. Зазначимо, що використання параметра  $S_i$  дозволило побудувати незначно гіршу за точністю класифікаційну модель (95,8%) в широкому інтервалі значення параметра згладжування  $\sigma$ . Класифікаційні моделі із використанням попередньо оброблених даних шляхом нормалізації, автошкалування та МГК виявилися менш надійними для ідентифікації зразків, однак нейронна мережа стала більш стабільною та швидкою в навчанні. Отже, перетворення даних електронного носа до певного формату та виду для вирішення

задачі класифікації ковбас імовірнісною нейронною мережею не є обов'язковим.

Запропоновано експрес-спосіб визначення масової частки соєвих продуктів у варених ковбасних виробках із використанням електронного носа в поєднанні з імовірнісною нейронною мережею, що дає змогу ідентифікувати зразки та виявити фальсифікацію продукції без застосування складного апаратного оформлення, багатоваріантної підготовки проби та коштовних реактивів.

#### Список використаних джерел

1. Ковбаси варені, сосиски, сардельки, хліби м'ясні: ДСТУ 4436-2005. Чинний від 2006.07.01. К.: Держспоживстандарт України, 2006. 32 с. Boiled sausages, frankfurters, sardellas, meat loaves: DSTU 4436-2005. 2006.07.01. Kyiv, Derzhspozhivstandart Ukraini, 2006, 32 p.
2. Russell T. A. Comparison of sensory properties of whey and soy protein concentrates and isolates: a thesis submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Master of Science Department of Food Science. Raleigh, 2004, 132 p.
3. Boatright W.L., Lei Q. J. Food Sci., 2000, 65, P. 819–821.
4. Шпичка А.И., Семенова Е.Ф. Фундаментальные исследования, 2013, 8, С. 1113–1124. Shpichka A.I., Semenova Ye. F. Fundamental'nye issledovaniya, 2013, 8, P. 1113–1124.
5. Мясо и мясные продукты. Метод гистологической идентификации состава: ГОСТ Р 51604-2000. Введен в действие 2000.05.12. М.: Госстандарт России, 2000, 11 с. Meat and meat products. Method of histological identification of composition: GOST R 51604-2000. 2000.05.12. Moscow, Gosstandart Rossii, 2000, 11 p.
6. Прошкин Л. В. Ветеринарно-санитарная экспертиза и методы определения качества и безопасности колбасных изделий. Автореф. дис. на соиск. уч. степени канд. вет. наук, СПб., 2011, 30 с. Proshkin L. V. Veterinary and sanitary expertise and methods of quality and safety evaluation of the sausage products. Abstract of thesis for candidata's degree, SPb., 2011, 30 p.

7. Мясо и мясные продукты. Определение массовой доли растительного (соевого белка) методом электрофореза: ГОСТ Р 53220-2008. Введен в действие 2010.01.01. М.: Стандартинформ, 2009, 12 с.

Meat and meat products. Electrophoretic method of determination of soy proteins mass content: GOST R 53220-2008. 2010.01.01. Moscow, Standartinform, 2009, 12 p.

8. Demuth H., Beale M. Neural Network Toolbox User's Guide. The MathWorks Inc., 2002, 840 p.

9. Esbensen Kim H. Multivariate Data Analysis – in practice. An Introduction to Multivariate Data Analysis and Experimental Design, 5th Edition. Aalborg University, Esbjerg. Oslo: Camo Process AS, 2004, 54–65.

10. Tran D. H., NG A. W. M., Perera B. J. C., Burn S., Davis P. Urban Water J., 2006, 3 (3), 175–184.

11. Rios A., Barcelo D., Buydens L., Cardenas S., Heydorn K., Karlberg B., Klemm K., Lendl B., Milman B., Neidhart B., Stephany R. W., Townshend A., Zschunke A., Valcarcel M. Accred. Qual. Assur., 2003, 8, 68–77.

12. Specht D. F. Neural Netw., 1990, 3, 109–118.

Надійшла до редколегії 23.11.17

А. Калиниченко, науч. сотр., asya.kalini4enko@gmail.com,

Л. Арсеньева, д-р техн. наук,

В. Пасичный, д-р техн. наук,

Национальный университет пищевых технологий, Киев, Украина

### ПРИМЕНЕНИЕ ЭЛЕКТРОННОГО НОСА И ВЕРОЯТНОСТНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ ИДЕНТИФИКАЦИИ КОЛБАС

*Предложен способ идентификации колбасных изделий и определения массового содержания в них соевых продуктов с использованием электронного носа и вероятностной нейронной сети. Оптимизирована архитектура нейронной сети для дискриминации образцов по параметрам электронного носа:  $S_i$  (надёжность классификации составляет 95,8% при значениях параметра сглаживания  $\sigma=3,6-54,0$ ) и  $\Delta F_i^{\max}$  (100% точность идентификации при значениях  $\sigma=0,2-1,0$ ).*

*Ключевые слова: электронный нос, вероятностная нейронная сеть, идентификация, вареные колбасы.*

A. Kalinichenko, Res. Sc., asya.kalini4enko@gmail.com,

L. Arseniyeva, Dr. Sc.,

V. Pasichnyi, Dr. Sc.

National University of Food Technologies, Kyiv, Ukraine

### ELECTRONIC NOSE AND PROBABILISTIC NEURAL NETWORK USE FOR SAUSAGES IDENTIFICATION

*The electronic nose system based on seven quartz-microbalance sensors was used to generate a pattern of the volatile compounds present in sausage samples with different mass content of soy protein isolate (SPI) ranging from 0 to 30% w/w. The original response curve was extracted to two features such as the maximum response value ( $\Delta F_i^{\max}$ , Hz) and area values of sensor response curve and time axis surrounded ( $S_i$ , Hz-s). All parameters subjected to pattern recognition analysis in original, normalised, scaling values and after principal component analysis. In this paper we used probabilistic neural network (PNN) for multiclass discrimination of sausage products. The neural network architecture was optimized for samples discrimination using as input vectors the electronic nose parameters such as  $S_i$  and  $\Delta F_i^{\max}$ . The best classification reliability (95,8%) for model based on dataset of  $S_i$  in original values obtained with the values of the PNN smoothing parameter  $\sigma$  in the range from 3,6 to 54,0. The classification model built by  $\Delta F_i^{\max}$  dataset in original values gave the 100% identification accuracy with the value of the PNN smoothing parameter in the range from 0,2 to 1,0. Furthermore, the different pre-processing techniques applied to datasets led to a slight decrease the prediction performance of the classification models, but the speed of neural network training has increased.*

*This paper presents a novel approach to identification of cooked sausages and determination of soy products mass content using the electronic nose combined with probabilistic neural network. Compared to classical methods, this new approach could represent an alternative and innovative tool for faster and cheaper sausage identification and mass content of soy protein isolate (0, 10, 20, 30% w/w) detection.*

*Keywords: electronic nose, probabilistic neural network, identification, cooked sausages.*