



systems. *Sensors* 2023, 23, 3020.

25. Reegu, F.A.; Abas, H.; Gulzar, Y.; Xin, Q.; Alwan, A.A.; Jabbari, A.; Sonkamble, R.G.; Dziauddin, R.A. Blockchain-based framework for interoperable electronic health records for an improved healthcare system. *Sustainability* 2023, 15, 6337.

26. Saad, G.; Harb, H.; Abouaissa, A.; Idoumghar, L.; Charara, N. A sensing-based patient classification framework for efficient patient–nurse scheduling. *Sustainable Computing: Informatics and Systems* 2023, 38, 100855.

27. Aldousari, A.; Alotaibi, M.; Khajah, F.; Jaafar, A.; Alshebli, M.; Kanj, H. A wearable IoT-based healthcare monitoring system for elderly people. In *Proceedings of the International Conference on Bio-Engineering for Smart Technologies (BioSMART)*, Paris, France, 7–9 June 2023; IEEE: Piscataway, NJ, USA; pp. 1–4.

28. Suliyanti, W.N.; Sari, R.F. Blockchain-based double-layer Byzantine fault tolerance for scalability enhancement in building information modeling information exchange. *Big Data and Cognitive Computing* 2023, 7, 90.

29. Mohan, R.; Ferris, T. The current state of healthcare information exchange (HIE) and proposing a blockchain HIE infrastructure. In *Blockchain in Healthcare: From Disruption to Integration*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 19–36.

30. Parekh, N.; Mangrulkar, R. Enabling blockchain architecture for health information exchanges. In *Unleashing the Potentials of Blockchain Technology for Healthcare Industries*; Elsevier: Amsterdam, The Netherlands, 2023; pp. 77–93.

Отримана в редакції 01.12.2025. Прийнята до друку 12.12.2025. Розміщено в інтернеті 30 березня 2026.

Received 01 December 2025. Approved 12 December 2026. Available in Internet 30 March 2026

УДК 004.896:[658.8:664.6]

## КЛАСТЕРНИЙ АНАЛІЗ ПАРТІЙ БОРОШНА В СИСТЕМІ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ ЕФЕКТИВНОСТІ ЗАМІСУ ТІСТА

## CLUSTER ANALYSIS OF FLOUR BATCHES IN A DECISION SUPPORT SYSTEM FOR THE EFFICIENCY OF DOUGH KNEADING

Рустамов Р.Р.<sup>1</sup>, Жигайло О.М.<sup>2</sup>  
Rustamov R.R.<sup>1</sup>, Zhygailo O.M.<sup>2</sup>

<sup>1,2</sup>Одеський національний технологічний університет, Одеса, Україна

<sup>1,2</sup>Odesa National University of Technology, Odesa, Ukraine

ORCID: <sup>1</sup><https://orcid.org/0009-0000-3667-892X>, <sup>2</sup><https://orcid.org/0000-0001-6986-4673>

Email: <sup>1</sup>[mr.rustamov550@gmail.com](mailto:mr.rustamov550@gmail.com), <sup>2</sup>[dr\\_jam2006@ukr.net](mailto:dr_jam2006@ukr.net)

Copyright © 2026 by author and the journal “Automation of technological and business – processes”.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0>



DOI: [10.15673/atbp.v18i1.3439](https://doi.org/10.15673/atbp.v18i1.3439)

**Анотація.** У статті розглядаються результати оглядово-аналітичного дослідження щодо перспектив застосування методів кластеризації для вирішення завдань автоматизації на етапі аналізу сировини та замісу тіста. Впровадження інструментів аналізу даних спрямоване на розв'язання актуальної проблеми стабілізації якості хлібобулочних виробів в умовах, коли фізико-хімічні та реологічні показники основної сировини (борошна) є суттєво нестабільними та піддаються значним природним коливанням. Існуючі традиційні системи управління хлібопекарським виробництвом мають специфіку, пов'язану з використанням жорстко заданих технологічних програм. Ці програми не здатні адаптуватися до поточної варіабельності характеристик кожної конкретної партії борошна (зокрема, до змін кількості сирової клейковини, індексу її деформації та числа падіння), що неминуче призводить до нестабільності структурно-механічних властивостей напівфабрикатів і



зниження якості готової продукції. Як базовий математичний інструментарій для інтеграції аналітичного та технологічного рівнів управління запропоновано застосування методу кластерного аналізу. У роботі наведено концептуальну структуру автоматизованої системи обробки інформації, яка об'єктивно розділяє масив вхідних даних про сировину на однорідні групи (кластери). Результати показали, що попередня кластеризація показників борошна формує необхідний базис для подальшого автоматизованого визначення оптимальних параметрів програм замісу (співвідношення тривалості та інтенсивності на різних етапах). Впровадження запропонованого підходу в майбутньому буде забезпечувати підвищення якості продукції без необхідності модернізації існуючого на підприємстві технологічного обладнання.

**Abstract.** The article presents the results of a review and analytical study on the prospects of applying clustering methods to solve automation tasks at the stage of raw material analysis and dough mixing. The introduction of data analysis tools is aimed at addressing the urgent problem of stabilizing the quality of bakery products under conditions where the physicochemical and rheological indicators of the main raw material (flour) are significantly unstable and subject to considerable natural fluctuations. Existing traditional bakery production control systems are characterized by the use of rigidly defined technological programs. These programs are unable to adapt to the current variability of characteristics of each specific flour batch (in particular, to changes in wet gluten content, its deformation index, and falling number), which inevitably leads to instability in the structural and mechanical properties of semi-finished products and a decrease in the quality of the finished product. The application of the cluster analysis method is proposed as a fundamental mathematical tool for integrating the analytical and technological levels of control. The paper presents a conceptual structure of an automated information processing system that objectively divides the array of input data on raw materials into homogeneous groups (clusters). The results demonstrated that preliminary clustering of flour indicators forms the necessary basis for the subsequent automated determination of optimal mixing program parameters (the ratio of duration and intensity at different stages). The implementation of the proposed approach will in future ensure improved product quality without the need to upgrade the existing technological equipment at the enterprise.

**Ключові слова:** кластерний аналіз, якість борошна, заміс тіста, автоматизація, режим замісу, хлібопекарське виробництво, інтелектуальний аналіз даних.

**Keywords:** clustering, flour quality, dough kneading, automation, kneading modes, bakery production, data mining.

**Вступ.** Хлібобулочні вироби є однією з базових категорій продуктів харчування, що мають стратегічне значення для продовольчої безпеки та щоденного раціону населення. Сучасні тенденції розвитку харчової промисловості та перехід підприємств до парадигми Індустрії 4.0 вимагають впровадження автоматизованих систем управління на всіх етапах життєвого циклу продукції. Основною проблемою, що перешкоджає стабільному випуску високоякісних виробів, є значна природна варіабельність фізико-хімічних та реологічних властивостей основної сировини - борошна. Показники кількості та якості сировини, індексу деформації клейковини (ІДК) та числа падіння суттєво коливаються залежно від сорту зерна, кліматичних умов вирощування та термінів зберігання. Традиційні системи управління технологічним процесом замісу тіста переважно використовують жорстко задані (статичні) рецептури та час замісу. Вони не здатні оперативної адаптуватися до змінних характеристик кожної нової партії борошна, що неминує призводить до погіршення структурно-механічних властивостей напівфабрикатів, зниження якості готової продукції та економічних втрат підприємства. Відповідно, виникає об'єктивна необхідність удосконалення систем управління шляхом впровадження методів інтелектуального аналізу даних для попередньої об'єктивної оцінки сировини та адаптивного вибору параметрів роботи обладнання.

#### Аналіз літературних даних і постановка проблеми

У дослідженні [1] представлено систематизацію показників якості пшеничного борошна, що використовуються у світовій практиці. Узагальнено фізико-хімічні індикатори, зокрема вміст білка, зольність, вологість, характеристики клейковини та їх вплив на функціональні властивості тіста. Основний акцент зроблено на впорядкуванні критеріїв оцінювання якості сировини. Поглиблений аналіз показав, що жоден окремий показник не може бути використаний як універсальний критерій оцінки, оскільки формування реологічних властивостей тіста є результатом складної взаємодії білкових, ферментативних та водно-фізичних характеристик. Дослідниками також встановлено, що варіація навіть одного параметра, наприклад вмісту білка, може по-різному впливати на властивості тіста залежно від загального складу системи, що свідчить про нелінійний характер залежності між показниками якості. Це ускладнює використання традиційних підходів до оцінювання та вимагає застосування методів багатовимірної аналізу. У джерелі [2] наведено результати дослідження реологічних властивостей борошна із застосуванням експериментальних методів, що дозволяють оцінити поведінку тіста в процесі замісу. Дослідження базується на варіюванні складу сировини шляхом додавання солодового борошна, що дозволило проаналізувати зміну таких параметрів, як водопоглинальна здатність, час розвитку тіста, стабільність та ступінь розрідження. У результаті встановлено, що збільшення частки добавки призводить до зміни ферментативної активності системи, що, у свою чергу, впливає на структурно-механічні властивості тіста. Зокрема, спостерігається зниження стабільності тіста та зміна його еластичності, що прямо відображається на якості кінцевого продукту. Отримані залежності демонструють високу чутливість технологічного процесу до



складу сировини та підтверджують необхідність точного підбору режимів замісу залежно від її характеристик. У дослідженні [3] висвітлено підходи до аналізу багатовимірних даних якості борошна та розглянуто можливості групування зразків за сукупністю показників. Окрім цього, наведено результати експериментального дослідження впливу рослинних добавок на якість хліба. Встановлено, що збільшення концентрації добавки призводить до зростання виходу продукції та втрат при випіканні, а також до зменшення об'єму хліба. Виявлено підвищення кислотності, вологості та зольності, а також зміну кольорових характеристик м'якушки. Водночас текстурні властивості залишаються відносно стабільними при невеликих концентраціях добавки, проте їх погіршення спостерігається при значному збільшенні її вмісту. Показано, що використання добавки суттєво підвищує вміст фенольних сполук та антиоксидантну активність продукту. Результати сенсорного аналізу свідчать, що найбільш прийнятними для споживачів є зразки з невисоким рівнем добавки, тоді як збільшення її концентрації негативно впливає на органолептичні показники. Отримані результати демонструють складний та неоднозначний вплив складу сировини на якісні характеристики готової продукції, що підтверджує необхідність комплексного врахування декількох параметрів при оцінюванні та оптимізації технологічних режимів. У дослідженні [4] розглянуто функціональні характеристики тіста та їх зв'язок із показниками якості борошна. Проаналізовано вплив структурних параметрів на сприйняття якості продукції. Додатково встановлено, що навіть за відсутності впливу вологості, відмінності у структурі та текстурі хліба можуть суттєво впливати на процес споживання, зокрема на характер жування та формування харчового болюсу. Показано, що зразки з меншою пористістю м'якушки потребують більших зусиль під час жування, що свідчить про визначальну роль внутрішньої структури продукту. Разом з тим встановлено, що незначні зміни у рецептурі, які впливають на текстуру, не завжди можуть бути чітко ідентифіковані споживачами під час сенсорного оцінювання. Це свідчить про складний взаємозв'язок між інструментально вимірюваними характеристиками та суб'єктивним сприйняттям якості. Отримані результати підтверджують, що структурно-механічні властивості продукції формуються під впливом сукупності факторів і мають суттєвий вплив на її експлуатаційні характеристики, що потребує комплексного підходу до їх аналізу.

Однак, слід зазначити, що розглянуті дослідження не мають комплексної практичної реалізації для автоматизованого вибору програм замісу тіста, оскільки вони переважно зосереджені на окремих аспектах оцінювання якості борошна або аналізу його властивостей, без інтеграції цих підходів у єдину систему підтримки прийняття рішень.

#### **Мета і завдання дослідження**

Метою дослідження є підвищення результативності управління якістю хлібобулочних виробів шляхом розроблення та впровадження системи підтримки прийняття рішень, що базується на кластеризації показників якості сировини та автоматизованому виборі й коригуванні програм замісу тіста відповідно до поточних технологічних умов.

Для досягнення визначеної мети необхідно:

- провести аналіз існуючих підходів до оцінювання якості борошна та управління процесом замісу тіста.
- проаналізувати показники якості сировини, що можуть бути використані для кластеризації;
- застосувати метод кластерного аналізу для формування кластерів борошна;
- Встановити відповідність між сформованими кластерами та програмами замісу тіста.

#### **Методи і матеріали досліджень**

Головним методологічним підґрунтям дослідження є діалектичний метод пізнання, який дозволив розглядати систему управління якістю хлібобулочних виробів як динамічну систему, у межах якої показники якості сировини безпосередньо впливають на параметри технологічного процесу замісу тіста.

У роботі застосовано загальнонаукові методи дослідження: аналіз і синтез - для дослідження існуючих підходів з управління якістю та вибору програм замісу; індукцію і дедукцію - для формування висновків щодо залежності між характеристиками борошна та режимами замісу; порівняння та спостереження - для оцінювання стабільності технологічних показників. Основним спеціальним методом дослідження є метод кластерного аналізу, який застосовано для групування партій борошна за сукупністю фізико-хімічних показників (вологість, вміст і якість клейковини, число падіння тощо).

#### **Результати досліджень**

Аналіз та підготовка даних. Експериментальною базою для інтелектуального аналізу використовувався масив лабораторних даних щодо показників якості 731 партії борошна. В якості основних вхідних змінних було обрано три ключові фізико-хімічні параметри, що найбільше впливають на реологію тіста: вміст сирової клейковини (діапазон 24–28 %), індекс деформації клейковини (ІДК, діапазон 45–65 у.о.) та число падіння (діапазон 239–470 с). Аналіз даних проводився у програмному середовищі Orange Data Mining.

Застосування алгоритмів кластеризації. Для автоматизованого групування партій сировини було проведено порівняльний аналіз трьох методів кластеризації: K-means (K-середніх), методу Лувена (Louvain method) та ієрархічної кластеризації.

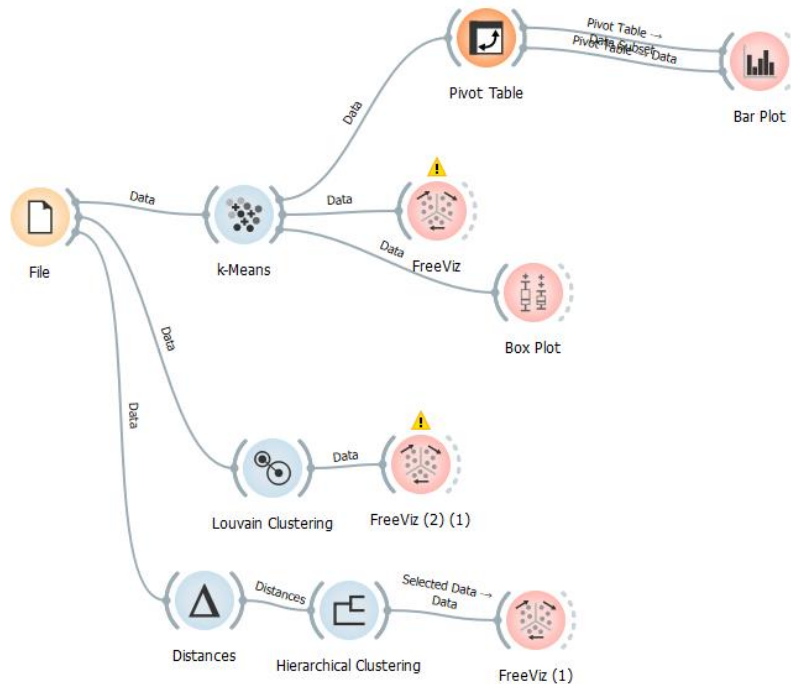
Принцип алгоритму K-means полягає в пошуку таких центрів кластерів та наборів елементів кожного кластера, коли сумарне квадратичне відхилення елементів кластерів від центрів цих кластерів буде найменшим:



$$V = \sum_{i=1}^k \sum_{x_j \in S_i} (x_j - \mu_i)^2$$

де  $k$  - число кластерів;  $S_i$  - отримані кластери;  $i = 1, 2, \dots, k$ ,  $\mu_i$  - центри мас векторів  $x_j \in S_i$ .

Цей алгоритм мінімізує дисперсії в межах кластера, що дозволяє чітко розмежувати партії борошна за подібними технологічними властивостями.



**Рис. 1** Схема моделювання процесів кластеризації у середовищі Orange Data Mining  
**Fig. 1** Scheme of modeling clustering processes in the Orange Data Mining environment

Як альтернативу розглянуто метод Лувена, який оптимізує модульність графа  $Q$  – міру, що відображає щільність зв'язків усередині спільнот порівняно зі зв'язками між ними:

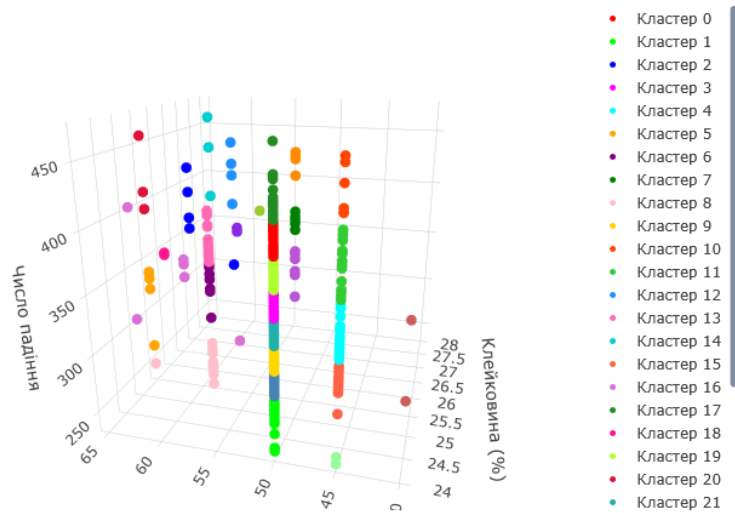
$$Q = \frac{1}{(2m)} \sum_{(i=1)}^N \sum_{(j=1)}^N \left[ A_{ij} - \frac{(k_i k_j)}{(2m)} \right] \delta(c_i, c_j).$$

$A_{ij}$  представляє вагу ребра між вузлами  $i$  та  $j$ ;  $k_i$  та  $k_j$  є сумою ваг ребер, приєднаних до вузлів  $i$  та  $j$ , відповідно;  $m$  є сумою всіх ваг ребер у графі;  $N$  – загальна кількість вузлів у графі;  $c_i$  та  $c_j$  є спільноти, до яких вузли  $i$  та  $j$  належать. Проте експерименти показали, що через велику відстань між центроїдами метод Лувена відносить більшість партій борошна до одного великого кластера, що нівелює технологічну цінність розподілу.

Ієрархічна кластеризація (з використанням критерію мінімальної дисперсії Уорда) також показала прийнятні результати, проте її часова складність робить її надто повільною для обробки великих промислових масивів даних у режимі реального часу. Таким чином, експериментальним шляхом встановлено, що алгоритм K-means є найбільш ефективним інструментом для нашої задачі.

Одним із найважливіших етапів дослідження було визначення оптимальної кількості кластерів ( $k$ ). Для цього застосовувався метод аналізу коефіцієнта силуету (Silhouette Score), який вимірює, наскільки об'єкт схожий на свій власний кластер порівняно з іншими кластерами. Значення силуету варіюється від -1 до +1. Високе значення вказує на те, що об'єкт добре узгоджується зі своїм кластером і погано з сусідніми. Експериментальним шляхом було визначено оптимальне розбиття масиву на технологічно значущі групи, що дозволило класифікувати всю вхідну сировину за ступенем її реологічної «сили» (від дуже сильного до слабого борошна).

Після ідентифікації кластерів було розроблено алгоритм інформаційного зв'язку між властивостями сировини та параметрами технологічного обладнання. Суть алгоритму полягає у тому, що кожному виділеному кластеру (його центроїду) ставиться у відповідність конкретна технологічна програма роботи тістомісильної машини. Наприклад, кластерам, що характеризуються високим вмістом клейковини та низьким ІДК (міцне борошно), автоматично призначається програма з подовженим часом інтенсивного механічного оброблення (замісу) для повноцінного розвитку білкового каркаса. Навпаки, для кластерів із показниками слабого борошна час замісу автоматично мінімізується для запобігання деструкції тіста.



**Рис. 2 Результати кластеризації показників якості борошна у тривимірному просторі**  
**Fig. 2 Results of clustering flour quality indicators in three-dimensional space**

Архітектура та реалізація інформаційної системи. На основі запропонованого математичного апарату в подальшому доцільно створити веб-орієнтований програмний модуль підтримки прийняття технологічних рішень. Для забезпечення гнучкої взаємодії між користувацьким інтерфейсом і серверною частиною (а в перспективі – з ERP та SCADA системами підприємства) пропонується використовувати архітектурний стиль RESTful API. Передбачається, що модуль оброблятиме дані у форматі JSON за допомогою методів HTTP.

Передбачається, що алгоритм системи автоматично порівнюватиме середні значення показників сировини у виділеному кластері з допустимими межами параметрами технологічних програм і формуватиме рекомендації: тривалість замісу на першій та другій швидкостях тістомісильної машини.

Серверну бізнес-логіку маршрутизації та генерації технологічних рекомендацій пропонується реалізувати на базі мови PHP. Для постійного зберігання історичних даних, результатів лабораторних аналізів та параметрів технологічних програм доцільно використовувати розроблену концептуальну структуру бази даних у СУБД MySQL.

Користувацький інтерфейс пропонується розробляти з використанням сучасних веб-технологій (HTML5, CSS3, фреймворк Bootstrap). Він має забезпечувати інтуїтивно зрозуміле введення нових результатів лабораторних аналізів, ініціювання процесу класифікації в один клік та наочну візуалізацію результатів. Передбачається, що візуальний модуль зможе будувати динамічні 3D-діаграми за допомогою бібліотеки Plotly.js. Це дозволить технологю візуально оцінити розташування нової партії борошна відносно існуючих кластерів та підтвердити запроповану алгоритмом програму замісу (розподіл часу між першою та другою швидкостями машини).

Кластер	Програма замісу	Тривалість 1-й етап (с)	Тривалість 2-й етап (с)	Загальна тривалість (с)	Вибрати вручну
0	Програма 7	162	168	330	Автоматично
1	Програма 5	162	138	300	Автоматично
2	Програма 1	192	168	360	Автоматично
3	Програма 5	162	138	300	Автоматично
4	Програма 3	168	102	270	Автоматично
5	Програма 7	162	168	330	Автоматично

**Рис. 3 Приклад реалізації інтерфейсу програмного модуля інтелектуального вибору технологічних режимів**  
**Fig. 3 Draft interface of the software module for intelligent selection of technological modes**



### Обговорення результатів

Аналіз отриманих результатів моделювання та апробації програмного модуля дозволяє зробити обґрунтований висновок про високу ефективність синергії методів машинного навчання та традиційних технологій хлібопечення. Перевагою розробленого алгоритму є безболісний перехід від статичних, жорстко зафіксованих у технологічних картах режимів замісу до динамічного адаптивного управління. На відміну від існуючих підходів, де рішення щодо часу замісу приймається оператором інтуїтивно (на основі органолептичної оцінки або пробних випічок), розроблена система забезпечує суворий математичний детермінізм, заснований на об'єктивних метриках.

Важливим аспектом, що вимагає уваги при промисловому впровадженні, є явище природного дрейфу концепту (concept drift) у характеристиках сировини. З огляду на те, що фізико-хімічні властивості пшениці та борошна можуть суттєво змінюватися залежно від регіону вирощування зерна, погодних умов під час жнив та термінів дозрівання борошна на складах, жорстка фіксація центротидів кластерів на тривалий час (роки) є недоцільною. Алгоритм K-means демонструє найвищу точність за умови періодичного перенавчання (оновлення) моделі. Ми рекомендуємо реалізувати в системі процедуру «ковзного вікна», яка б автоматично виключала з навчальної вибірки застарілі лабораторні дані (старші за один квартал) та перераховувала центроїди кластерів виключно на основі найсвіжшої інформації. Це забезпечить постійну релевантність системи управління.

Запропонована REST-архітектура успішно вирішує ще одну глобальну проблему сучасних хлібозаводів – відсутність інформаційного зв'язку між локальними підсистемами. Типові ERP-системи, масово впроваджені на виробництвах, працюють виключно з логістичними та фінансовими даними. Розроблений нами модуль заповнює цей вакуум, формуючи «цифровий міст» між аналітичною лабораторією підприємства та нижнім рівнем автоматизації (АСУ ТП).

### Висновки

1. Обґрунтовується доцільність застосування інструментів інтелектуального аналізу даних, зокрема алгоритму кластеризації K-means, для обробки багатовимірних масивів лабораторних показників якості борошна. Показано, що попереднє центрування та стандартизація вхідних ознак (вміст клейковини, ІДК, число падіння) можуть виступати необхідною умовою для коректного використання метричних алгоритмів.

2. Запропоновано підхід до прийняття технологічних рішень на основі залежності «центроїд кластера борошна – програма замісу тіста». Такий підхід потенційно дозволяє здійснювати адаптивний вибір параметрів замісу залежно від реологічних характеристик сировини та сприяти формуванню необхідних структурно-механічних властивостей тіста.

3. Розглянуто можливість створення веб-орієнтованого модуля підтримки вибору режимів замісу тіста. Використання архітектури RESTful API може забезпечити ефективний обмін даними у форматі JSON, достатню швидкість обробки запитів та інтеграцію відповідних алгоритмів у інформаційні системи підприємства.

### Список використаних джерел

- [1]. Cappelli A., Oliva N., Cini E. A systematic review of gluten-free dough and bread dough rheology, bread characteristics and improvement strategies // *Journal of Cereal Science*. – 2020. – Vol. 95.
- [2]. Hrušková M., Švec I. Evaluation of wheat flour rheological properties // *Czech Journal of Food Sciences*. – 2022. – Vol. 40(2). – P. 85–92.
- [3]. Pietrzak D., Różyło R. Quality characteristics of wheat flour // *Foods*. – 2019. – Vol. 8(8). – P. 349.
- [4]. Rosell C.M. The science of doughs and bread quality // *Food Research International*. – 2018. – Vol. 106. – P. 104–112.

### References

- [1]. A. Cappelli, N. Oliva, and E. Cini, “A systematic review of gluten-free dough and bread dough rheology,” *Journal of Cereal Science*, vol. 95, 2020.
- [2]. M. Hrušková and I. Švec, “Evaluation of wheat flour rheological properties,” *Czech Journal of Food Sciences*, vol. 40, no. 2, pp. 85–92, 2022.
- [3]. D. Pietrzak and R. Różyło, “Quality characteristics of wheat flour,” *Foods*, vol. 8, no. 8, p. 349, 2019.
- [4]. C. M. Rosell, “The science of doughs and bread quality,” *Food Research International*, vol. 106, pp. 104–112, 2018.

Отримана в редакції 02.02.2026. Прийнята до друку 16.02.2026. Розміщено в інтернеті 30 березня 2026.  
Received 02 February 2026. Approved 16 February 2026. Available in Internet 30 March 2026