

СИСТЕМА ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ПАРАМЕТРІВ ВИРОЩУВАННЯ АГРОКУЛЬТУР

Пухальський О. В., науковий керівник Голуб Б.Л.

Мета. Розробка системи інтелектуального аналізу параметрів вирощування агрокультур для оцінки та прогнозування їх урожайності на основі технологій Data Mining.

Об'єкт і предмет дослідження. Об'єктом дослідження є процес вирощування агрокультур у різних агроекологічних умовах. Предметом дослідження є методи інтелектуального аналізу даних для виявлення закономірностей між агрономічними параметрами та показниками врожайності.

Актуальність. У сучасних умовах аграрного виробництва важливого значення набуває застосування технологій інтелектуального аналізу даних (Data Mining) для оптимізації процесів вирощування сільськогосподарських культур. Використання великих обсягів аграрної інформації дозволяє виявляти приховані залежності між параметрами вирощування (тип ґрунту, кліматичні умови, регіональні особливості) та кінцевими результатами врожайності. Традиційні методи аналізу є недостатніми для ефективного прогнозування, що зумовлює необхідність впровадження сучасних аналітичних систем.

Для досягнення поставленої мети було застосовано комплекс методів інтелектуального аналізу даних.

На першому етапі здійснювалася класифікація записів за допомогою алгоритму 1-Rule, що дозволило виявити основні фактори, які найбільше впливають на рівень урожайності, базуючись на простих правилах однієї ознаки. Для підвищення точності класифікації використовувався метод наївного Байєса, що дає змогу враховувати одночасний вплив декількох параметрів (регіон, тип ґрунту, культура).

Наступним етапом був пошук асоціативних закономірностей між характеристиками вирощування та показниками урожайності з використанням алгоритму Apriori. Це дозволило сформулювати правила типу "якщо-то", що описують приховані залежності в аграрних даних.

На завершальному етапі проведено кластеризацію аграрних об'єктів методом K-Means з подальшим зниженням розмірності даних за допомогою методу головних компонент (PCA) для візуалізації отриманих груп.

Для реалізації процесу аналізу даних використовувалися інструменти Microsoft SQL Server, Power BI та Python (pandas, scikit-learn та mlxtend).

В рамках роботи було створено аграрну базу даних AgroAnalysisDB, яка містить інформацію про регіони вирощування, типи ґрунтів, кліматичні умови, культури та їх урожайність за період 2015–2021 років.

На основі методів класифікації вдалося виявити найважливіші фактори, що впливають на продуктивність агрокультур. Пошук асоціативних правил дозволив сформулювати залежності між умовами вирощування та ймовірністю досягнення високої врожайності. Кластеризація даних сприяла виділенню груп регіонів та культур зі схожими агрономічними характеристиками.

Візуалізація результатів кластеризації методом PCA наведена на рисунку 1.

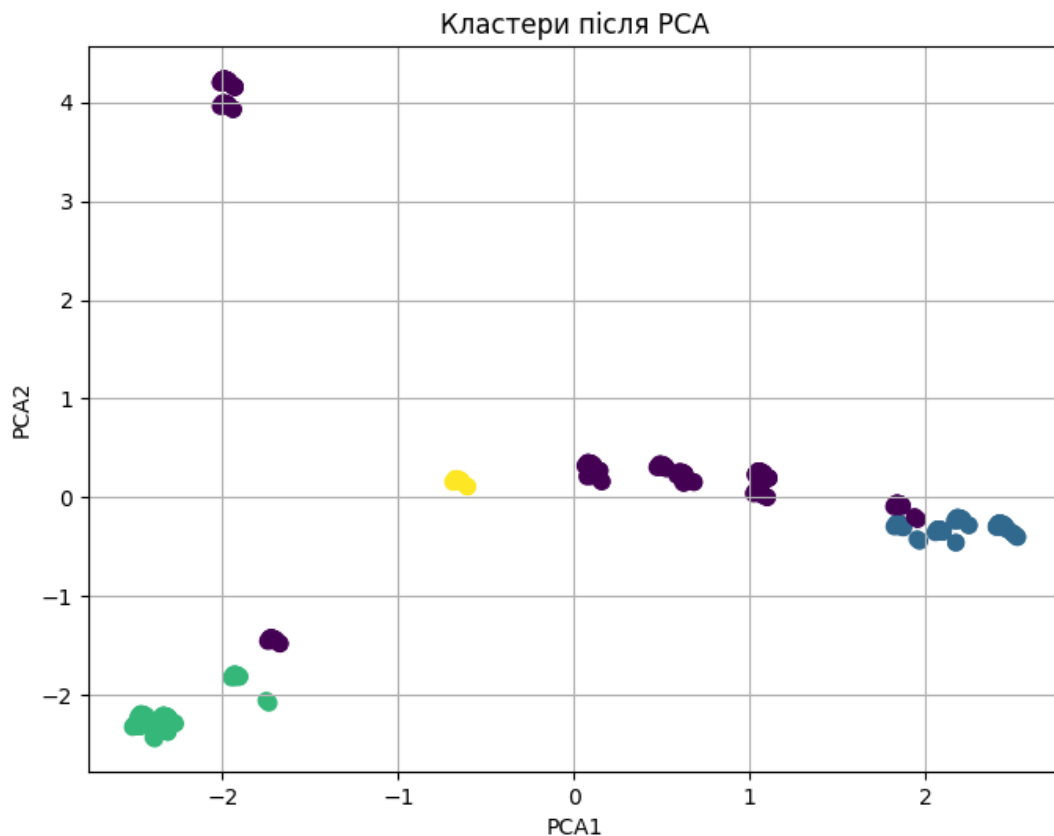


Рисунок 3 Візуалізація результатів кластеризації методом PCA

Висновки. Розроблена система інтелектуального аналізу параметрів вирощування агрокультур довела свою ефективність у виявленні ключових факторів, що впливають на урожайність. Застосування методів класифікації, асоціативного аналізу та кластеризації забезпечило можливість глибокого розуміння процесів вирощування культур, що дозволяє підвищувати ефективність агровиробництва, планувати агротехнічні заходи та мінімізувати виробничі ризики.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Державна служба статистики України. – [Електронний ресурс]. – Режим доступу: <https://ukrstat.gov.ua>
2. Tan P. N., Steinbach M., Karpatne A., Kumar V. Introduction to Data Mining. – 2nd ed. – Pearson, 2018. – 864 p.
3. Liu B. Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data. – Springer, 2011. – 632 p.
4. Scikit-learn Developers. User Guide for Scikit-learn: Machine Learning in Python. [Електронний ресурс]. – Режим доступу: https://scikit-learn.org/stable/user_guide.html

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І
ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**



**ТЕОРЕТИЧНІ ТА ПРИКЛАДНІ
АСПЕКТИ РОЗРОБКИ
КОМП'ЮТЕРНИХ СИСТЕМ**

**ЗБІРНИК НАУКОВИХ ПРАЦЬ ЗА МАТЕРІАЛАМИ
VII ВСЕУКРАЇНСЬКОЇ НАУКОВО-ПРАКТИЧНОЇ КОНФЕРЕНЦІЇ
СТУДЕНТІВ І АСПІРАНТІВ
*24 квітня 2025 року***

Київ 2025

УДК 004

Відповідальна за випуск: М.І. Лендел

Збірник наукових праць за матеріалами VII Всеукраїнської науково-практичної конференції студентів і аспірантів «ТЕОРЕТИЧНІ ТА ПРИКЛАДНІ АСПЕКТИ РОЗРОБКИ КОМП'ЮТЕРНИХ СИСТЕМ '2025», 24 квітня 2025 року, НУБіП України, Київ. – 452 с. (електронне видання)

Відповідальність за зміст публікацій несуть автори.

Передрук матеріалів, а також використання їх будь-якій формі допускається лише з дозволу авторів