

## ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ВИРІШЕННЯ ЗАДАЧІ НЕЙРОМЕРЕЖЕВОГО ВИЗНАЧЕННЯ ЯКОСТІ ВИНА

Вінницький національний технічний університет

### *Анотація*

Дана робота присвячена розробці програмного забезпечення для інформаційної технології нейромережевого визначення якості вина. У роботі було обґрунтовано вибір нейронної мережі багатопшаровий перцептрон для визначення якості вина, яка має 11 входів, 2 прихованих шари по 8 нейронів та вихідний шар із одного нейрона. У прихованих шарах обрано функцію активації ReLU та функцію активації Sigmoid у вихідному шарі. Для навчання цієї нейромережі використовується модифікація ADAM методу стохастичного градієнтного спуску. Для створення програми було використано мову програмування Python та спеціалізовані бібліотеки NumPy, Pandas та Matplotlib. Навчання нейромережі відбувалось з використанням бази даних для червоних вин, яка налічує 1599 записів. Набір даних було поділено на навчальну (75%) та тестову (25%) вибірки: обсяг навчальної вибірки склав 1199 записів, а тестової - 400 записів. Розроблена програма має достовірність класифікації (визначення якості вина) на тестовій вибірці 79,2%, а найкращий із 3-х методів-аналогів має достовірність класифікації на тестовій вибірці 73,5%, тобто достовірність класифікації збільшилась на 5,7%.

**Ключові слова:** класифікація, машинне навчання, нейронна мережа, двофакторний аналіз

### *Abstract*

This work is devoted to the development of software for information technology of neural network determination of wine quality. The work justified the choice of a multilayer perceptron neural network for determining wine quality, which has 11 inputs, 2 hidden layers of 8 neurons each and an output layer of one neuron. The ReLU activation function and the Sigmoid activation function in the output layer were selected in the hidden layers. To train this neural network, a modification of the ADAM stochastic gradient descent method is used. The Python programming language and specialized libraries NumPy, Pandas and Matplotlib were used to create the program. The neural network was trained using a database for red wines, which has 1599 records. The data set was divided into a training (75%) and a test (25%) sample: the volume of the training sample was 1199 records, and the test - 400 records. The developed program has a classification accuracy (determination of wine quality) on the test sample of 79.2%, and the best of the 3 similar methods has a classification accuracy on the test sample of 73.5%, i.e. the classification accuracy has increased by 5.7%.

**Keywords:** classification, machine learning, neural network, two-factor analysis

### **Вступ**

Кожен виробник вина стикається з необхідністю його оцінки та сертифікації. Оцінка вина дозволяє визначити його якість з погляду споживача, тобто наскільки гарний його смак, запах і т. д. Існує два основних методи, що використовуються при сертифікації вина. Перший ґрунтується на фізико-хімічному аналізі. Другим способом є сенсорний аналіз. Він заснований на використанні спеціально навченої людини – дегустатора. Залежність між фізико-хімічними властивостями вина та результатом дегустації є вкрай складною та заплутаною, і в даний час вона до кінця не визначена. З цього випливає, що за допомогою фізико-хімічного аналізу складно дати відповідь на запитання, наскільки добре буде вино з погляду покупця. З іншого боку, дегустатор визначає якість вина виходячи зі своїх суб'єктивних понять про «хороше» і «погане» вина. Отже, у сфері виробництва винної продукції існує потреба у дослідженні можливості створення деякого класифікатора, який міг би визначати якість вина за його хімічним складом, керуючись лише математичними алгоритмами.

Метою роботи є підвищення достовірності визначення якості вина за рахунок використання нейронної мережі.

### Результати досліджень

Було розглянуто три нейромережі: одношаровий перцептрон, багатошаровий перцептрон та RBF-неймережа та як найбільш перспективний для даної задачі було обрано багатошаровий перцептрон [1]. Розроблено архітектуру нейронної мережі багатошаровий перцептрон (рис. 1) для інформаційної технології визначення якості вина, яка має 11 входів, 2 прихованих шари по 8 нейронів та вихідний шар із одного нейрона. У прихованих шарах обрано функцію активації ReLU та функцію активації Sigmoid у вихідному шарі. Для навчання цієї нейромережі використовувалася модифікація ADAM методу стохастичного градієнтного спуску.

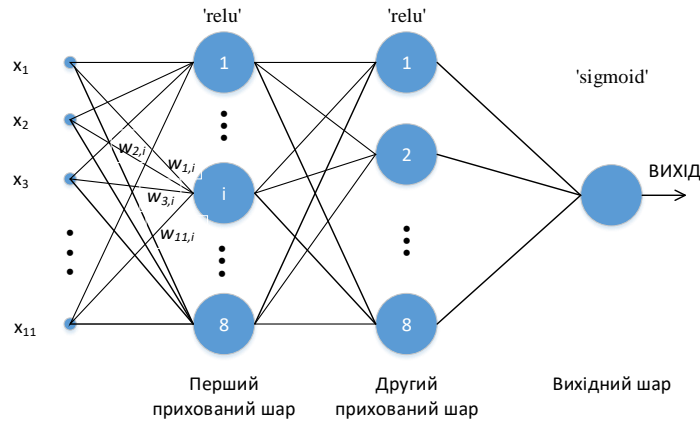


Рисунок 1 – Структура багатошарового перцептрона для інформаційної технології визначення якості вина

Відповідно до мети роботи та постановки задачі було розроблено алгоритм роботи інформаційної технології нейромережевого визначення якості вина, представлений на рис. 2.

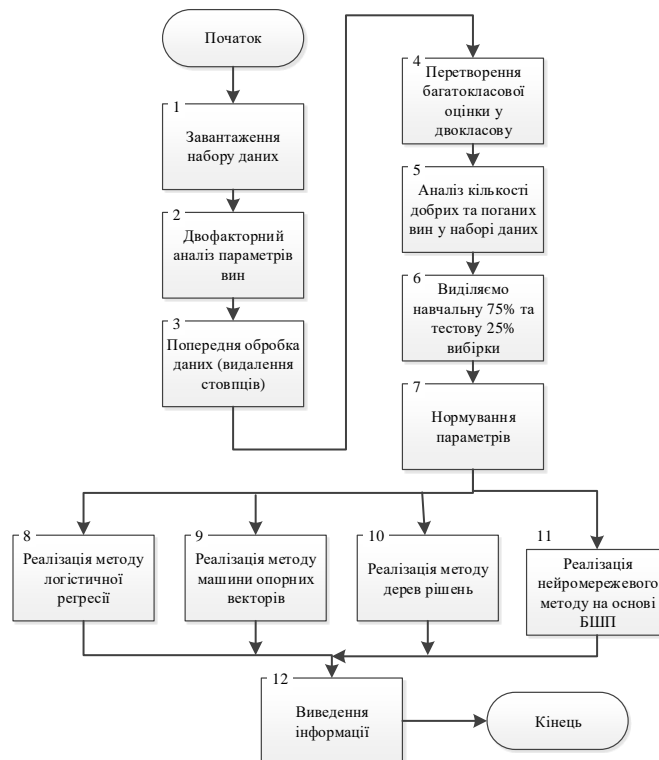


Рисунок 2 – Схема алгоритму роботи інформаційної технології нейромережевого визначення якості вина

Для програмної реалізації інформаційної технології нейромережевого визначення якості вина було обґрунтовано вибір мови програмування Python та спеціалізованих бібліотек NumPy, Pandas та Matplotlib.

Для навчання та тестування нейромережі використовувався набір даних, який охоплює вина, створені в Португалії в провінції Мінью з травня 2004 року по лютий 2007 року. Дані були надані дослідниками Paulo Cortez, Antonio Cerdeira, Fernando Almeida, Telmo Matos, Jose Reis у статті «Modeling wine preferences by data mining from physicochemical properties» [2].

Фрагмент записів набору даних показано на рис. 3.

	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	Sulphates	Alcohol	Quality
0	7.4	0.70	0.00	1.9	0.076	11.0	34.0	0.9978	3.51	0.56	9.4	5
1	7.8	0.88	0.00	2.6	0.098	25.0	67.0	0.9968	3.20	0.68	9.8	5
2	7.8	0.76	0.04	2.3	0.092	15.0	54.0	0.9970	3.26	0.65	9.8	5
3	11.2	0.28	0.56	1.9	0.075	17.0	60.0	0.9980	3.16	0.58	9.8	6
4	7.4	0.70	0.00	1.9	0.076	11.0	34.0	0.9978	3.51	0.56	9.4	5

Рисунок 3 – Фрагмент набору даних параметрів вина для навчання та тестування нейромережі

Набір даних містить параметри як червоного вина, так і білого вина. Набір даних червоного вина містить 1599 екземплярів, а набір даних білого вина містить 4898 екземплярів. Обидва набори даних мають 11 вхідних змінних (на основі фізико-хімічних тестів): фіксовану кислотність, летючу кислотність, лимонну кислоту, залишковий цукор, хлориди, вільний діоксид сірки, загальний діоксид сірки, щільність, рН, сульфати, алкоголь та 1 вихідну змінну (на основі сенсорних даних дегустаторів): якість. Сенсорні дані оцінюються за 11 класами якості від 0 до 10 (0 - дуже погано, 10 - дуже добре).

Чисельні результати роботи розробленого програмного забезпечення, отримані в результаті тестування, занесено до табл. 1 для наочності порівняння різних методів визначення якості вина (класифікації [3]).

Таблиця 1 – Порівняння достовірності класифікації різних методів визначення якості вина

Метод класифікації	Кількість зображень у навчальній вибірці	Кількість зображень у тестовій вибірці	Достовірність класифікації на навчальній вибірці	Достовірність класифікації на тестовій вибірці
Метод логістичної регресії	1199	400	74,8%	72,3%
Метод машини опорних векторів	1199	400	81%	73%
Метод дерев рішень	1199	400	100%	73,5%
Нейромережевий метод	1199	400	93,7%	79,2%

Із табл. 1 видно, що розроблений програмний продукт має достовірність класифікації (визначення якості вина) на тестовій вибірці 79,2%, а найкращий із 3-х методів-аналогів має достовірність класифікації на тестовій вибірці 73,5%.

Таким чином, можна зробити висновок, що нейромережевий метод визначення якості вина має порівняно з методами-аналогами збільшену на 5,7% достовірність класифікації. Тобто мета роботи досягнута – достовірність визначення якості вина підвищена.

### Висновки

Таким чином, було розроблено програмне забезпечення для інформаційної технології нейромережевого визначення якості вина. У роботі було обґрунтовано вибір нейронної мережі

багатошаровий перцептрон для визначення якості вина, яка має 11 входів, 2 прихованих шари по 8 нейронів та вихідний шар із одного нейрона. У прихованих шарах обрано функцію активації ReLU та функцію активації Sigmoid у вихідному шарі. Для навчання цієї нейромережі використовується модифікація ADAM методу стохастичного градієнтного спуску. Для створення програми було використано мову програмування Python та спеціалізовані бібліотеки NumPy, Pandas та Matplotlib. Навчання нейромережі відбувалось з використанням бази даних для червоних вин, яка налічує 1599 записів. Набір даних було поділено на навчальну (75%) та тестову (25%) вибірки: обсяг навчальної вибірки склав 1199 записів, а тестової - 400 записів. Розроблена програма має достовірність класифікації (визначення якості вина) на тестовій вибірці 79,2%, а найкращий із 3-х методів-аналогів має достовірність класифікації на тестовій вибірці 73,5%, тобто достовірність класифікації збільшилась на 5,7%.

#### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. “ Руденко О.В. Штучні нейронні мережі: Навчальний посібник / О.В.Руденко, Є.В.Бодянський. - Харків : ТОВ «Компанія СМІТ», 2006. — 404 с. - ISBN 966-8630-73-X.
2. P. Cortez, A. Cerdeira, F. Almeida, T. Matos and J. Reis. Modeling wine preferences by data mining from physicochemical properties. In Decision Support Systems, Elsevier, 47(4):547-553, 2009.
3. Інтелектуальний аналіз даних та машинне навчання. Частина 1. Базові методи та засоби аналізу даних (навчальний посібник) / Я. В. Іванчук, В. І. Месюра, А. А. Яровий, О. Д. Манжілевський – Вінниця : ВНТУ, 2021. – 69 с.

**Закревський Артур Олегович** — студент групи ІКН-23м, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, email: zakartur9@gmail.com

**Колесницький Олег Костянтинович** — канд. техн. наук, професор кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця. email: [kolesnytskyi@vntu.edu.ua](mailto:kolesnytskyi@vntu.edu.ua)

**Паночішин Юрій Миколайович** — канд. техн. наук, доцент кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця. email: y.panochyshyn@vntu.edu.ua

**Zakrevskyi Artur O.** – student of Intelligent Information Technologies and Automation Department, Vinnytsia National Technical University, Vinnytsia, e-mail: : zakartur9@gmail.com

**Kolesnytskyj Oleh K.** – Cand. Sc. (Eng.), Professor of the Computer Science Chair, Vinnytsia National Technical University, Vinnytsia. email: [kolesnytskyi@vntu.edu.ua](mailto:kolesnytskyi@vntu.edu.ua)

**Panochyshyn Yuriy M.** – Cand. Sc. (Eng.), Docent of the Computer Science Chair, Vinnytsia National Technical University, Vinnytsia. email: y.panochyshyn@vntu.edu.ua