

ІНТЕЛЕКТУАЛЬНА ОБРОБКА ДАНИХ ВІД ХЛОРОФІЛ-ФЛУОРОМЕТРИЧНИХ СЕНСОРІВ

Вступ. Один із способів спостереження за станом рослинних об'єктів є індукція флуоресценції хлорофілу (ІФХ) – це випромінювання хлорофілу в червоному спектрі світла, що виникає при освітленні хлорофілу живої рослини в синьому спектрі світла. Для вимірювання ІФХ використовують спеціальні прилади, так звані хлорофіл-флуорометри. Ряд таких приладів розроблені в Інституті кібернетики імені В.М. Глушкова НАН України. Зокрема, розроблений портативний хлорофіл-флуорометр «Флоратест» та мережа бездротових сенсорів для вимірювання ІФХ [1].

Зростання досліджень ІФХ за останні роки як в Україні, так і за її межами призвело до накопичення значних обсягів даних, що дозволяє застосовувати моделі машинного навчання для обробки даних вимірювань. В даний час найпопулярнішим алгоритмом є штучні нейронні мережі [2]. Також, рідше, але використовуються й інші методи машинного навчання, такі як метод k-найближчих сусідів, дерева рішень і т. п. [3] Серед робіт присвячених обробці вимірювань ІФХ можна виділити такі:

- 1) прогнозування стану декоративних рослин у залежності від умов середовища утримання та змін в індукції флуоресценції хлорофілу [4];
- 2) таксономічне розрізнення рослин по кривих флуоресценції хлорофілу [5];
- 3) визначення водного дефіциту [6, 7];
- 4) визначення джерела стресу [8];
- 5) визначення зрілості рослин [9].

Вигляд кривої ІФХ та ряд основних геометричних параметрів, які використовують для опису даної кривої, показано на рис. 1.

З використанням розроблених в Інституті кібернетики імені В.М. Глушкова НАН України сенсорів (рис. 2, 3) проведено ряд експериментальних досліджень та зібрано ряд наборів даних. Зокрема, проведено експеримент з обприскування рослинних об'єктів гербіцидом і зареєстрована зміна ІФХ по днях після проведення експерименту.

В статті представлені результати застосування методів машинного навчання (нейронних мереж, методу опорних векторів, алгоритму XGBoost) для обробки вимірювань індукції флуоресценції хлорофілу на прикладі задачі класифікації рослин оброблених різними дозами гербіциду та задачі визначення необхідності штучного поливу.

Ключові слова: індукція флуоресценції хлорофілу, нейронна мережа, метод опорних векторів, алгоритм XGBoost.

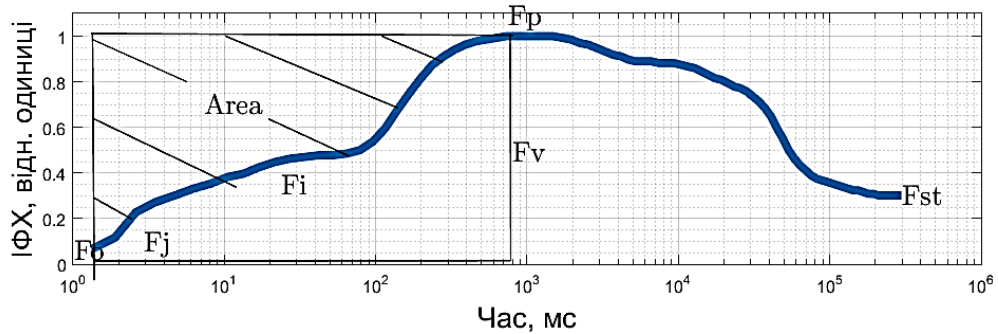


РИС. 1. Крива ІФХ (логарифмічна шкала)



РИС. 2. Прилад сімейства «Флоратест»



РИС. 3. Дослідний взірець бездротової сенсорної мережі

1. Розпізнавання впливу гербіциду

З метою вивчення можливості визначення достатньої дози внесення гербіциду з допомогою індукції флуоресценції хлорофілу був проведений відповідний експеримент в якому рослинні об'єкти було поділено на три групи (контрольна – без обробки гербіцидом і дві групи з різними дозами гербіциду). На зібраних даних проведено вивчення можливостей різними алгоритмами машинного навчання здійснити класифікацію рослин за відповідними групами в різні дні експерименту.

Нейромережевий підхід включав використання двошарової нейронної мережі з сигмоїдною функцією активації (гіперболічний тангенс) та нормалізована експоненційна (softmax) функцією на вихідному шарі нейронів. Використана нейронна мережа показала, що вже на 7-й день після обробки рослинних гербіцидом може навчитися розпізнавати рослини різних груп. Покращити результат дозволила нормалізація кривої, що усувало деякі відмінності між кривими вимірюваними різними флуорометрами. Найкращий результат показали мінімаксна нормалізація в межах $[-1, 1]$ та централізація значення навколо середнього (z – оцінка):

$$\tilde{x}_{ik} = 2 \frac{x_{ik} - F_{0,k}}{F_{m,k} - F_{0,k}} - 1, \quad (1)$$

$$\tilde{x}_{ik} = \frac{x_{ik} - \bar{x}_k}{\delta_{x_k}}, \quad (2)$$

де x_{ik} – елемент матриці в i -му рядку, k -му стовпчику; \tilde{x}_{ik} – нормалізоване значення флуоресценції у i -му рядку k -го стовпчика; \bar{x}_k – середнє значення k -го стовпця; δ_{x_k} – стандартне відхилення значень ІФХ в k -му стовпчику; $F_{m,k}$ – максимальне значення в k -му стовпчику; $F_{0,k}$ – мінімальне значення флуоресценції в k -му стовпчику.

Застосування даних способів нормалізації дозволили покращити середню точність розпізнавання нейронною мережею груп рослин оброблених гербіцидом на 7-й день після обприскування з 66,7 % до 80,9 %.

Нейронні мережі забезпечують раннє визначення впливу стресового фактору на стан рослини ще до появи на рослині зовнішніх ознак, а саме показали непогані результати на 7-й день після обробки гербіцидом в той час, як явна дія гербіциду проявляється в зовнішніх ознаках рослин на дванадцятий, чотирнадцятий день.

В подальшому згаданий набір даних було опробовано на методі опорних векторів (Support Vector Machine). Опробовано такі ядерні функції (ядра) як лінійна, поліноміальна, rbf та сігмоїдна. Найкращий результат продемонструвало поліноміальне ядро.

Як і з нейронними мережами, при застосуванні методу опорних векторів на різних днях обприскування, метод почав добре розрізняти дослідні групи на 7-й день після обприскування. Метод опорних векторів показав кращі результати ніж використовувана нейронна мережа, так на 7-й день після обприскування максимальна точність розпізнавання методом опорних векторів 1.0 (100 %), водночас, як нейронна мережа забезпечувала максимальну точність 0.8.

Також на згаданому наборі даних використано алгоритм XGBoost (Extreme Gradient Boosting) – популярний алгоритм машинного навчання, що реалізує модель градієнтного бустінгу, який предсталає собою альтернативу регресійним методам та нейронним мережам. Метод полягає у створенні ансамблю послідовно уточнюючих один одного дерев ухвалення рішень. N -модель дерева навчається на «помилках» ансамбля з $n-1$ моделей, відповіді моделей зважено сумуються. Була використана реалізація алгоритму XGBClassifier на мові python з бібліотеки scikit-learn.

В табл. 1 наведено результат XGBClassifier для задачі розпізнавання впливу гербіциду в різні дні після обприскування без нормалізації даних. Проводилось тисячу випробувань з різним розбиттям даних на тестову та навчальну вибірки. На тестову вибірку було відведено 30 % вимірних кривих ІФХ. Алгоритм найкраще справляється з класифікацією даних на 11-й день після обприскування та на 20-й день. Між тим доволі низький результат продемонстровано на 7-й день після обприскування. Саме цей день використовується як тестовий для різних алгоритмів класифікації.

В табл. 2 наведено аналогічні результати тестування класифікації кривих ІФХ при використанні мінімаксної нормалізації (1). Таким чином, після мінімаксної нормалізації вимірних кривих індукції флуоресценції хлорофілу, XGBClassifier зміг розпізнати криві із 100 % точністю на 7-й день після обприскування, на 11-й та 13-й день. На 20-й день після обприскування якість розпізнавання погіршилась. Тож XGBClassifier демонструє кращі результати ніж використана нейронна мережа. XGBClassifier також демонструє кращий результат за метод опорних векторів 18.07.2013 (13-й день після обприскування), де метод опорних векторів забезпечив точність лише 87,78 %.

ТАБЛИЦЯ 1. Результати XGBClassifier без попередньої обробки даних

Точність розпізнавання			Дата вимірювання
мін	середнє	макс	
0,0000	0,3839	0,7500	До обробки
0,1538	0,4685	0,8462	До обробки
0,0769	0,4289	0,7692	3-й день після обробки
0,0769	0,3582	0,6923	5-й день після обробки
0,0000	0,3366	0,6923	7-й день після обробки
0,1000	0,6089	1,0000	11-й день після обробки
0,1538	0,5279	0,8462	13-й день після обробки
0,1538	0,5630	1,0000	20-й день після обробки

ТАБЛИЦЯ 2. Використання мінімаксної нормалізації

Точність розпізнавання			Дата вимірювання
мін	середня	макс	
0,0833	0,5051	0,8333	До обробки
0,0769	0,5037	0,9231	До обробки
0,0000	0,3643	0,7692	3-й день після обробки
0,0000	0,3395	0,6923	5-й день після обробки
0,2308	0,5701	1,0000	7-й день після обробки
0,2000	0,6998	1,0000	11-й день після обробки
0,1538	0,5849	1,0000	13-й день після обробки
0,1538	0,4943	0,9231	20-й день після обробки

Використання Z-оцінки (2) погіршило результат.

Найкращі досягнені результати отримані різними алгоритмами машинного навчання на 7-й день після обприскування наведені в табл. 3, де А – точність розпізнавання у процентах.

ТАБЛИЦЯ 3. Найкращі результати, отримані різними алгоритмами класифікації

Алгоритм класифікації	А, %
Нейронна мережа	80.9 %
Метод опорних векторів	100 %
XGBoost	100 %

2. Визначення вологості ґрунту за допомогою індукції флуоресценції хлорофілу

Також був проведений експеримент з визначення необхідності штучного поливу. Протягом декількох тижнів здійснювалося вимірювання ІФХ рослин цинії, частина з яких поливалась постійно, а інша частина не поливалась. Алгоритмам машинного навчання ставилась задача визначити за кривою ІФХ рівень вологості (задача регресії). Як і в попередньому експерименті було застосовано нейронну мережу, метод опорних векторів та алгоритм XGBoost.

Використана нейронна мережа містила три шари нейронів з сигмоїдною функцією активації у прихованих шарах та лінійною функцією активації нейронів у вихідному шарі.

Один з результатів проведення регресії нейронною мережею показано на рис. 4.

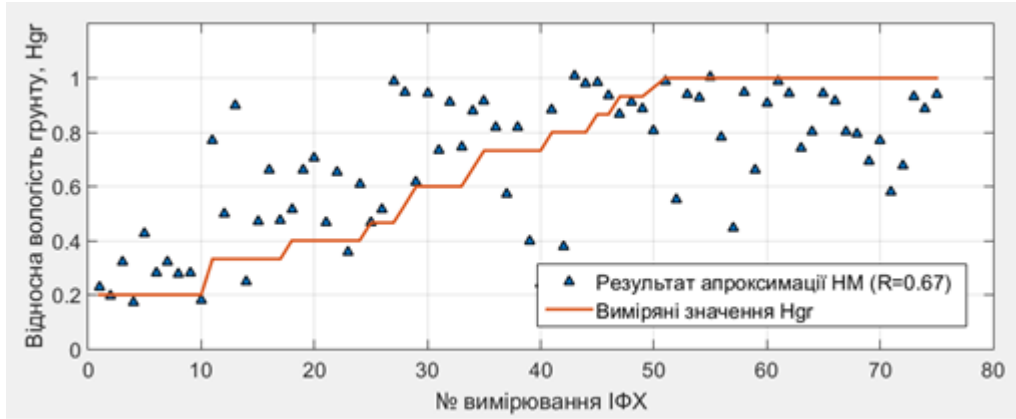


РИС. 4. Результати апроксимації нейронною мережею

Найкращий отриманий результат отриманий за допомогою згаданої нейронної мережі становив $R = 0.81$.

Використання методу опорних векторів показало, що найкращий результат демонструвало поліноміальне ядро четвертого степеня з використанням мінімаксної нормалізації даних. Також опробовано можливість подачі різної кількості дискретно виміряних точок кривої ІФХ від 10 до максимальних 90, що видають сенсори ІФХ взятих нелінійно за степеневою шкалою (ступінь 1/8). Найкращий результат SVR показало при використанні 70 точок.

Використання алгоритму XGBoost (використано реалізацію XGBRegressor з python бібліотеки scikit-learn) показало, що найкращі результати отримуються із використанням нормалізації даних за допомогою z-оцінки (2). Регресія проводиться з доволі високим результатом уже при взятті 10 дискретно виміряних точок вимірювання ІФХ взятих за степеневою шкалою.

Результати всіх трьох алгоритмів машинного навчання (коефіцієнт регресії та коефіцієнт кореляції) подані в табл. 4.

ТАБЛИЦЯ 4. Найкращі результати, отримані різними алгоритмами при проведенні регресії

Алгоритм регресії	R^2_{\max}	R_{\max}
Нейронна мережа	0,65	0,81
Метод опорних векторів	0,54	0,73
Алгоритм XGBoost	0,52	0,72

Висновки. Застосування різних алгоритмів машинного навчання при обробці кривих індукції флуоресценції хлорофілу показало, що для задачі класифікації рослин обприсканих гербіцидом найкращий результати демонструють метод опорних векторів та алгоритм XGBoost водночас нейронна мережа продемонструвала найгірші результати. Застосування згаданих методів для задачі визначення необхідності штучного поливу показало, що найкращі результати демонструє нейронна мережа, дещо гірший результат метод опорних векторів і найгірше з задачею справлявся алгоритм XGBoost.

Список літератури

1. Romanov V., Galelyuka I., Antonova H., Kovyrova O., Hrusha V., Voronenko O. Application of Wireless Sensor Networks for Digital Agriculture. *Proceedings of the 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems*. 2019. P. 340–344. <https://doi.org/10.1109/IDAACS.2019.8924267>
2. Samborska I., Alexandrov V., Sieczko L., Kornatowska B., Goltsev V., Cetner M., Kalaji H. Artificial neural networks and their application in biological and agricultural research. *Signpost open access journal of nanophotobiosciences*. Vol. 02. 2014. P. 14–30.
3. Silva J., Figueiredo A., Cunha J., Eiras-Dias J., Silva S., Vanneschi L., Mariano P. Using Rapid Chlorophyll Fluorescence Transients to Classify Vitis Genotypes. *Plants*. Vol. 9. Issue 2. 174. 2020. <https://doi.org/10.3390/plants9020174>
4. Chlorophyll fluorescence spectral discrimination by artificial neural network methods. DEFRA project code HH1530SPC. 2002.
5. Kirova M., Ceppi G., Chernev P., Goltsev V., Strasser R. Using Artificial Neural Networks for Plant Taxonomic Determination Based on Chlorophyll Fluorescence Induction Curves. *Biotechnology and Biotechnological Equipment. XI Anniversary Scientific Conference 120 Years of Academic Education in Biology 45 Years Faculty of Biology*. P. 941–946. <https://doi.org/10.1080/13102818.2009.10818577>
6. Goltsev V., Zaharieva I., Chernev P., Kouzmanova M., Kalaji H.M., Yordanov I., Krasteva V., Alexandrov V., Stefanov D., Allkhverdiev S.I., Strasser R.J. Drought-induced modification of photosynthetic electron transport in intact leaves: Analysis and use of neural network as a tool for a rapid non-invasive estimation *Biochimica et Biophysica Acta*. **1817** (8). 2012. P. 1490–1498. <https://doi.org/10.1016/j.bbabi.2012.04.018>
7. Rybka K., Janaszek-Mankowska M., Siedlarz P., Mankowski D. Machine learning in determination of water saturation deficit in wheat leaves on basis of Chl a fluorescence parameters. *Photosynthetica*. **57** (1). 2019. P. 226–230. <https://doi.org/10.32615/ps.2019.017>
8. Soja G., Soja A.M. Recognizing the Sources of Stress in Wheat and Bean by using chlorophyll fluorescence induction parameters as inputs for neural network models. *Phyton*. Special issue: «D. Grill». **45** (3). 2005. P. 157–168.
9. Xanthoula Eirini Pantazi, Dimitrios Moshou, Dimitrios Kasampalis and Pavlos Tsouvaltzis. Automatic Assessment of Phenotypes in lettuce plants by using Chlorophyll Fluorescence Kinetics and Machine Learning. *Proceedings International Conference of Agricultural Engineering. AgEng*. 2014. Zurich, 6-10.07.2014. P. 167–176.

Одержано 15.06.2022

Груша Володимир Михайлович,

кандидат технічних наук, молодший науковий співробітник
 Інституту кібернетики імені В.М. Глушкова НАН України, Київ.
vhrusha@gmail.com
<https://orcid.org/0000-0002-2497-0939>

UDC 004.942

Volodymyr Hrusha**Intelligent Processing of Data From Chlorophyll Fluorometric Sensors**

V.M. Glushkov Institute of Cybernetics of the NAS of Ukraine, Kyiv

Correspondence: vhrusha@gmail.com

Introduction. Chlorophyll fluorescence induction (CFI) is a monitoring method of plant objects. CFI is a radiation of chlorophyll in red spectrum during a chlorophyll lighting of alive plant in blue spectrum. Chlorophyll fluorometers – the special devices that are used for measurement of CFI. Series of such devices were developed in V.M. Glushkov Institute of Cybernetics of the NAS of Ukraine. In particular, fluorometer «Flo-ratest» and a network of wireless sensors were developed for CFI measurement. An accumulation of massive amount of measurements resulted into possibility to use intellectual methods like neural networks.

The purpose of the paper is to research the possibilities of machine learning methods (neural networks, support vector machine (SVM), XGBoost algorithm) for analysis of CFI curves that were measured by means of sensors developed in V.M. Glushkov Institute of Cybernetics of the NAS of Ukraine.

Results. Neural networks, SVM, XGboost ensure early detection of influence of stress factors on state of plants before appearance of external symptoms on plants that was showed on basis of data received during experiments with treatment of plants by herbicide. Analogically there was showed the possibility of using the machine learning methods for determination of soil humidity. The better methods for given tasks were determined. The study of possibilities to enhance the results of mentioned methods by means of normalization was conducted. The best results were demonstrated by z-score normalization and by minimax normalization to the range $[-1;1]$.

Conclusions. The application of different machine learning algorithm for processing CFI curves demonstrated that SVM and XGBoost better suit for task of classification plants treated by means of herbicide. Neural network demonstrated worst results. The application mentioned methods for task of determination of artificial watering necessity demonstrated that neural network shows better result, SVM shows worse result and XGBoost shows worst result.

Keywords: Chlorophyll fluorescence induction, neural network, support vector machine, algorithm XGBoost.