

ВИКОРИСТАННЯ ЧАСОВИХ РЯДІВ ТА НЕЙРОННИХ МЕРЕЖ LSTM ДЛЯ ІДЕНТИФІКАЦІЇ СТАНІВ ВУЛИКА

Жуков О.О.¹, Горбенко В.І.²

¹Запорізький національний університет, асистент, Україна

²Запорізький національний університет, к.ф.м.н, доцент, Україна

Анотація. Бджоли є надзвичайно важливими для екосистем та сільського господарства, проте їх популяції скорочуються через різноманітні загрози. Це зумовлює критичну потребу в ефективних системах моніторингу стану бджолиних сімей для своєчасного сповіщення пасічника. Традиційні методи мають обмеження, тому зростає інтерес до застосування нейронних мереж для цієї задачі. Метою даної роботи є розробка моделей на основі архітектури LSTM для ідентифікації різних станів вулика за допомогою аналізу часових рядів даних сенсорів (вага, температура). Було використано набір даних, що пройшов попередню обробку, включаючи агрегацію, інтерполяцію, розрахунок похідних ознак та сегментацію. Навчені дві LSTM-моделі - для точкових (роїння, підгодовування, забір меду) та глобальних станів (нормальний, загибель, весняний розвиток, медозбір) – продемонстрували високу точність (96% та 92% відповідно), підтверджуючи ефективність підходу.

Ключові слова: моніторинг бджіл, стан вулика, часові ряди, LSTM, нейронні мережі, машинне навчання, бджільництво, RNN.

Бджоли відіграють значну роль у підтримці біорізноманіття та забезпеченні продовольчої безпеки через запилення. Однак світові популяції бджіл зазнають значного скорочення через комплексні фактори, включаючи втрату середовища існування, пестициди та зміни клімату [1]. Це створює нагальну потребу в розробці ефективних систем моніторингу бджолиних колоній, що дозволить виявляти проблеми на ранніх стадіях та відповідних заходів. Існуючі методи моніторингу часто є інвазивними, трудомісткими або недостатньо інформативними. Тому використання технологій штучного інтелекту, зокрема нейронних мереж LSTM, для аналізу даних з сенсорів вулика є перспективним напрямком для безперервного та неінвазивного контролю стану бджолиних сімей. Попередні дослідження показували ефективність використання температурних даних для ідентифікації станів вулика. [2]

В якості джерела даних в дослідженні було використано відкритий датасет з Kaggle[3], що містить анотовані дані з вулика за 2019-2022 роки. Процес підготовки даних включав: 1) агрегацію похвилинних даних до 15-хвилинних (для точкових станів) та 4-годинних (для глобальних станів) інтервалів; 2) селективний експорт стабільних діапазонів; 3) інтерполяцію пропусків та згладжування; 4) розрахунок похідних ознак (відносна зміна ваги, співвідношення температур); 5) сегментацію даних з використанням рухомого вікна (6 годин для точкових, 96 годин для глобальних станів) та застосуванням буферу; 6) рандомізацію та балансування класів для уникнення зміщення; 7) розділення на навчальну (80%) та тестову (20%) вибірки. Було розроблено та навчено дві моделі на основі рекурентної архітектури LSTM. Перша модель для ідентифікації точкових станів (роїння, підгодовування, забір меду) була тренована з використанням відносної зміни ваги, місяця та години. Друга модель для класифікації глобальних станів (нормальний, загибель, весняний розвиток, медозбір) використовує відносну зміну ваги, співвідношення зовнішньої та внутрішньої температур, місяць та має модифіковану архітектуру нейронної мережі зі збільшеною кількістю нейронів. Обидві моделі навчались протягом 100 епох. Тестування на відкладеній вибірці показало загальну точність 96% для моделі точкових станів та 92% для моделі глобальних станів.

Висновки. У ході дослідження було успішно розроблено та протестовано дві моделі на базі нейронних мереж LSTM для ідентифікації точкових та глобальних станів бджолиного вулика за допомогою аналізу часових рядів. Продемонстровано ефективність запропонованого підходу: моделі досягли загальної точності 96% та 92% відповідно на тестових даних. Результати підтверджують перспективність використання LSTM та даних про вагу і температуру для створення автоматизованих систем моніторингу стану бджолиних сімей, що може сприяти своєчасному виявленню проблем та збереженню популяцій бджіл.

ЛІТЕРАТУРА/ REFERENCES

1. Overview of Bee Pollination and Its Economic Value for Crop Production / S. A. M. Khalifa et al. // Insects. 2021. Vol. 12, iss. 8. Art. 688. URL: <https://doi.org/10.3390/insects12080688>
2. Kviesis A. Application of neural networks for honey bee colony state identification / A. Kviesis, A. Zacepins // 2016 17th International Carpathian Control Conference (ICCC) (High Tatras, Slovakia, 29 May–1 June 2016) : матеріали. Piscataway : IEEE, 2016. P. 413–417. URL: <https://doi.org/10.1109/CarpathianCC.2016.7501133>
3. Senger D., Gruber C., Kluss T., Johannsen C. Weight, temperature and humidity sensor data of honey bee colonies in Germany, 2019–2022 / D. Senger, C. Gruber, T. Kluss, C. Johannsen // Data in Brief. 2024. T. 52. Art. 110015. URL: <https://doi.org/10.1016/j.dib.2023.110015>

USING TIME SERIES AND LSTM NEURAL NETWORKS FOR IDENTIFICATION OF HIVE STATES

Zhukov O.O., Horbenko V.I.

Abstract. *Bees are extremely important for ecosystems and agriculture, yet their populations are declining due to various threats. This creates a critical need for effective systems to monitor honeybee colony health, allowing beekeepers to react promptly. Traditional methods have limitations, leading to growing interest in artificial intelligence applications. This work aims to develop models based on the LSTM architecture to identify different beehive states by analyzing time-series data from sensors (weight, temperature). Data from Kaggle was used, undergoing thorough preparation including aggregation, interpolation, derivative feature calculation, and segmentation. Two LSTM models were trained – one for point states (swarming, feeding, honey harvesting) and another for global states (normal, colony collapse, spring build-up, honey flow). They demonstrated high accuracy (96% and 92% respectively), confirming the approach's effectiveness.*

Keywords: *bee monitoring, hive state, time series, LSTM, neural networks, machine learning, beekeeping.*