Name of the project:AP 19675312 Analytical system for forecasting the dynamics of the number of pests of grain crops in Kazakhstan based on a neural network model

Relevance:

The relevance of this project is due to the fact that health monitoring and early diagnosis of pests of grain crops is the most important task of sustainable agriculture. Information about the early diagnosis of certain plant diseases can facilitate pest control by choosing the right methods of combating them to increase grain yield. Manual identification of violations in grain crops can lead to inaccurate measurements, and also takes a lot of time. The above problem requires the intervention of the latest technologies, in particular the use of intelligent algorithms in predicting the dynamics of the number of plant pests. The use of intelligent systems in agriculture has become relevant in world science. But the existing systems do not disclose all aspects of this issue. To solve the problem of predicting the dynamics of the number of pests of grain crops, a neural network model will be created, on the basis of which an analytical system for predicting the dynamics of the number of pests of grain crops will be developed.

Purpose:

Creation of a neural network model for predicting the dynamics of the number of pests of grain crops for an analytical system.

Expected and achieved results:

- for 2023: The data collection was conducted to identify pests (particularly the striped flea beetle (Phyllotreta vittula)) in grain crops and assess their growth influence factors such as climate indicators, population, leaf area, and more. Data were gathered through expeditions to regional centers in Kazakhstan: Taldykorgan, Semey, Oskemen, Kostanay, and information was obtained from Pavlodar and Petropavlovsk. Data retrieved from the Republican Methodological Center for Phytosanitary Diagnostics and Forecasts included the density of striped flea beetle per m2. All compiled data, around 100,000 units, were processed and consolidated into a single Excel database. This database incorporated several years' worth of weather data obtained from the hydrometeorological center. The processed data were normalized and transformed into inputs, constructing an interaction matrix for neural network-based training. Research encompassed an investigation of algorithms and machine learning forecasting methods (regression models, decision trees, ensembles, clustering, random forests). A review of over 109 scientific articles on forecasting the growth of pests concluded that common methods included linear regression, random forests, decision trees, and clustering. Notably, there was a gap in employing machine learning, particularly neural networks, to forecast the growth of the striped flea beetle. This method had demonstrated effectiveness in training data for forest pests, indoor plants, among others, presenting promising results in forecasting. Given that the choice of forecasting method significantly influences machine learning outcomes, the neural network was selected for this research. Results analyzing forecasting algorithms were published in the article "Forecasting Methods in Machine Learning: Overview and Comparison" (authors: Akanova A.S., Ospanova N.N., Sharipova S.E., Anarbekova G.A., Kazanbayeva A.S., Journal: Bulletin of Toraygirov University No.3, 2023, pp. 20-33, recommended by CQASE ME RK, https://vestnik-energy.tou.edu.kz/storage/journals/171.pdf).

An analysis of existing analytical systems was conducted to evaluate their strengths, weaknesses, and potential enhancements in a new system. Positive aspects included the dynamic functionality of adding new data while continuing training and updating the database simultaneously. Negative aspects involved data localization; systems from the USA and Canada were tailored exclusively to insects existing within their territories. Furthermore, an evaluation of forecasting analytical systems revealed a lack of sections specifically forecasting the growth of

the striped flea beetle, notably absent in systems proposed by Canadian and American researchers.

Presentations were delivered at the "Food Quality and Food Safety" International Scientific Conference (FQFS) held from September 20th to 22nd, 2023, at S. Seifullin Kazakh Agrotechnical University. The research on data processing and normalization (data preparation for training) was presented at the VIII International Scientific-Practical Conference "Informatics and Applied Mathematics," dedicated to the memory of the renowned scientist and founder of the Kazakhstani cryptography school, Doctor of Technical Sciences, Professor Biyashev R.G., organized by the Institute of Information and Computational Technologies of SC SHEM RK, Al-Farabi Kazakh National University, Kazakh National Research Technical University named after K.I. Satpayev, International University of Information Technologies, Turan University, Lublin Technical University (Lublin, Poland), held in Almaty on October 26-27, 2023. The presentation topic was "FOOD – Data Preparation for Machine Learning and Forecasting the Dynamics of Wheat Pests" (authors: Akanova A.S., Ospanova N.N., Sharipova S.E., Anarbekova G.A., Kazanbayeva A.S., https://conf.iict.kz/ru/8th-ispc-csam-ru/#registration).

Following the analysis of forecasting methods, a multi-layer neural network model was developed, comprising layers: Embedding, Dense, Dense, Dense, forming an optimal architecture for training with effective indicators in forecasting the dynamics of grain crop pests.

- for 2024: Based on the selected hyperparameters of the neural network, trained data with accuracy indicators close to one and error loss close to zero will be obtained. A conceptual model and a prototype of the analytical system will be created. As a result, an analytical system will be developed to predict the dynamics of the number of pests.

- for 2025: Verification of the analytical system will be carried out to obtain the optimal indicator of the impact of the pest population of grain crops (exactly striped flea (*Phyllotretavittula*)). The analytical system will be tested, modified and launched after verification. A patent for a neural network model and a certificate of entering information into the state register of rights to objects protected by copyright for the analytical system will be obtained.

Members of the research group:

project manager:

Akanova Akerke Saparovna, h = 3 *(Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57212065112 ORCID: https://www.orcid.org/0000-0002-7178-2121, Researcher ID: G-8025-2018, ResearchGate: https://www.researchgate.net/profile/Akerke-Akanova-2*

research group:

Ospanova Nazira Nurgazievna– Senior Researcher, Associate Professorh = 4 *(ORCID: https://orcid.org/: 0000-0003-0100-1008, Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57209645760);*

Baibusenov Kurmet Serikovich - Senior Researcher, h=2, *(ORCID:https://orcid.org/ 0000-0002-9957-3073, Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56242719400)*

Kaldarova Mira Zhorabekovna - Researcher, h=1, *(ORCID:https://orcid.org/0000-0001-7494-9794), Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57218561944, ResearcherID - GPG-1180-2022: https://www.researchgate.net/profile/Mira-Kaldarova)*

Sharipova Saltanat Yerkinovna – Researcher, h=1, *(ORCID: https://orcid.org/0000-0001-7267-3261, Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=*57884433800)

Anarbekova Galiya Abayevna – Researcher *(ORCID: https://orcid.org/0000-0001-8960-8036)*

List of publications published as part of a preliminary study on the topic of this project:

Аканова А.С., Оспанова Н.Н. Жасанды нейрожелі арқылы өсімдіктерді тану // Bulletin of PSU, No. 3, 2019, energy series/http://vestnik-energy.tou.edu.kz/storage/journals/140.pdf

Байбусенов К.С., Ажбенов В.К., Сарбаев А.Т. Фитосанитарное прогнозирование популяционной динамики вредных нестадных саранчовых для обоснования и планирования защитных мероприятий в земледельческих районах Северного Казахстана // Bulletin of Science of the Kazakh Agrotechnical University named after S.Seifullin. – Astana, 2017. - № 4 (95). – pp. 28-35.

Anarbekova G., Ospanova N., Anarbekov D. Normalized input vectors: the primary stage of data preparation, News of National Academy of Sciences of the Republic of Kazakhstan, physical-mathematical series. Volume 2, Number 346 (2023), pp. 40–54. https://journals.nauka-nanrk.kz/physics-mathematics/article/view/5109

Information for potential users: An analytical system will be developed that predicts the population of pests of grain crops.