

Наименование проекта: АР 19675312 Аналитическая система прогнозирования динамики численности вредителей зерновых культур в Казахстане на основе нейросетевой модели

Актуальность:

Актуальность данного проекта обусловлена тем, что мониторинг здоровья и ранняя диагностика от вредителей зерновых культур является важнейшей задачей устойчивого сельского хозяйства. Информация о ранней диагностике некоторых заболеваний растений может облегчить борьбу с вредителями за счет правильного выбора методов борьбы с ними для повышения урожайности зерна. Ручная идентификация нарушений в зерновых культурах может привести к неточным измерениям, а также занимает большое количество времени. Вышесказанная проблема требует вмешательства новейших технологий, в частности применения интеллектуальных алгоритмов в прогнозировании динамики численности вредителей растения. В мировой науке стало актуальным применение интеллектуальных систем в сельском хозяйстве. Но существующие системы раскрывают не все аспекты данного вопроса. Для решения проблемы прогнозирования динамики численности вредителей зерновых культур будет создана нейросетевая модель, на основе которого будет разработана аналитическая система прогнозирования динамики численности вредителей зерновых культур.

Цель:

Создание нейросетевой модели прогнозирования динамики численности вредителей зерновых культур для аналитической системы.

Ожидаемые и достигнутые результаты:

- за 2023 год: Проведен сбор данных (влияние на рост вредителей: климатические показатели, численность, листовая поверхность и так далее) для выявления вредителей (в частности хлебной полосатой блошки (*Phyllotreta vittula*)) зерновых культур. Для сбора данных были совершены командировки в областные центры РК: Талдыкорган, Семей, Оскемен, Кустанай, так же челнами исследовательской группы были получены данные с города Павлодара и Петропавловска. В данных, собранных с РГУ «Республиканский методический центр фитосанитарной диагностики и прогнозов», указывалось количество плотности расположения полосатых блошек на м². Все полученные данные были обработаны и приведены в один табличный формат (базы данных EXCEL). Также в составленную БД были внесены погодные данные за несколько лет, полученные из гидрометеорологического центра. Данные в БД составили около 100 тыс единиц, они будут служить входными данными для нейронной сети. Обработанные данные были нормализованы. Нормализованные данные преобразованы во входные данные и создана матрица взаимодействия данных для обучения на основе нейронной сети. Выполнено исследование

алгоритмов и методов прогнозирования в машинном обучении (регрессионные модели, деревья решений и ансамбли, кластеризация, случайный лес). Для анализа методов прогнозирования численности роста вредителей, был проведен обзор научной литературы (более 109 научных статей). Из анализа научных исследований был сделан вывод что, часто используются такие методы как линейная регрессия, случайный лес, дерево решений, кластеризация. В существующих исследованиях отсутствует прогнозирование роста численности полосатой блошки с применением машинного обучения, в которых использовался метод нейронных сетей. Метод прогнозирования нейронные сети в частности применялся для обучения данных о вредителях леса, комнатных растений и др. А также исследования по прогнозированию (вредителей леса, вредителей комнатных растений и так далее) на основе нейронных сетей показали хорошие результаты обученности данных. Выбор метода прогнозирования является ключевым фактором для достижения успешных результатов в машинном обучении, поэтому для проведения исследования была выбрана нейронная сеть. Результаты по анализу алгоритмов прогнозирования были опубликованы в статье «Методы прогнозирования в машинном обучении: обзор и сравнение» (авторы: Аканова А.С., Оспанова Н.Н., Шарипова С.Е., Анарбекова Г.А., Казанбаева А.С., журнал Вестник Торайгыров Университета №3, 2023 год, стр 20-33, журнал, рекомендован КОКСНВО, <https://vestnik-energy.tou.edu.kz/storage/journals/171.pdf>)

Был проведен анализ существующих аналитических систем для учета отрицательных и положительных сторон и их совершенствования в новой системе. Положительными сторонами аналитических систем можно выделить динамичность такого функционала системы, как добавление новых данных к уже существующим, при котором обучение продолжается и в то же время база данных обновляется. Отрицательные стороны аналитических систем – локализация данных. Системы предложенные учеными США и Канады рассчитаны только на насекомых, существующих только на их территории. А так же был выполнен анализ состояния аналитических систем по прогнозированию вредителей, в частности учеными Канады и США предлагаются информационные системы, в которых отсутствуют разделы по прогнозированию роста численности полосатой блошки.

Принимали участие в Международной научной конференции «Food Quality and Food Safety» (FQFS) (Качество и безопасность продуктов питания), которая состоялась с 20 по 22 сентября 2023 года в КАТИУ имени С.Сейфуллина По результатам исследования в области обработки и нормализации данных (подготовка данных к обучению) было осуществлено выступление на VIII — Международной научно-практической конференции «Информатика и прикладная математика», посвященная памяти известного ученого, основоположника казахстанской школы криптографии, доктора технических наук, профессора Бияшева Р.Г. (организаторами являются Институт информационных и вычислительных технологий КН МНВО РК, Казахским Национальным университетом им. аль-Фараби, Казахским

национальным исследовательским техническим университетом имени К. И. Сатпаева, Международным университетом информационных технологий, университет «Туран», Люблинский технический университет (г.Люблин, Польша), место проведения Алмата, **26 -27 октября 2023 г. На тему «ЕДА - подготовка данных для машинного обучения и прогнозирования динамики вредителей пшеницы»** ((авторы: Аканова А.С., Оспанова Н.Н., Шарипова С.Е., Анарбекова Г.А., Казанбаева А.С.,) <https://conf.iict.kz/ru/8th-ispcc-sam-ru/#registration>.

По результатам анализа методов прогнозирования была разработана модель многослойной нейросети были выбраны слои: Embedding, Dense, Dense, Dense, что дала оптимальную архитектуру для проведения обучения с результативными показателями прогноза динамики численности вредителей зерновых культур.

- за 2024год: На основе подобранных гиперпараметров нейронной сети будут получены обученные данные с показателями точности приближенная к единице и потери ошибки приближенная к нулю. Будет создана концептуальная модель и прототип аналитической системы. В результате будет разработана аналитическая система для прогнозирования динамики численности вредителей.

- за 2025 год: Будет проведена верификация аналитической системы с получением оптимального показателя влияния численности вредителя зерновых культур (в точности полосатой блошки (*Phyllotreta vittula*)). Будет протестирована, изменена и запущена аналитическая система после верификации. Будет получен патент на нейросетевую модель и свидетельство о внесении сведений в государственный реестр прав на объекты, охраняемые авторским правом на аналитическую систему.

Члены исследовательской группы:

руководитель проекта:

Аканова Акерке Сапаровна, $h = 3$ (*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>

исследовательская группа:

Оспанова Назира Нургазыевна – старший научный сотрудник, доцент $h = 4$ (*ORCID*: <https://orcid.org/0000-0003-0100-1008>, *Scopus Author ID*: <https://www.scopus.com/authid/detail.uri?authorId=57209645760>);

Байбусенов Курмет Серикович - старший научный сотрудник, $h=2$, (*ORCID*: <https://orcid.org/0000-0002-9957-3073>, *Scopus Author ID*: <https://www.scopus.com/authid/detail.uri?authorId=56242719400>)

Калдарова Мира Жорабековна – младший научный сотрудник, $h=1$, (*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>)

Шарипова Салтанат Еркиновна – младший научный сотрудник, h=1
(ORCID: <https://orcid.org/0000-0001-7267-3261>, Scopus Author ID:
<https://www.scopus.com/authid/detail.uri?authorId=57884433800>, ResearcherID:
JVZ-4150-2024)

Анарбекова Галия Абаевна – младший научный сотрудник,
(ORCID: <https://orcid.org/0000-0001-8960-8036>)

Список публикаций опубликованные в рамках предварительного исследования по теме данного проекта:

Аканова А.С., Оспанова Н.Н. Жасанды нейрожелі арқылы өсімдіктерді тану/Вестник ПГУ, №3, 2019, энергетическая серия/ <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. Вестник науки Казахского агротехнического университета им. С.Сейфуллина. – Астана, 2017. - № 4 (95). – С. 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>

Информация для потенциальных пользователей: Будет разработана аналитическая система прогнозирующая численность популяции вредителей зерновых культур.