

«Утверждаю»

Проректор по науке, международным
связям и инновациям Кыргызского
государственного университета им.

И. Арабаева, д.в.н., доцент

А.К. Чалданбаева

«13» февраля 2024 г.



ВЫПИСКА

из протокола №6 расширенного заседания кафедры прикладная информатика института новых информационных технологий (ИНИТ) КГУ им. И. Арабаева по предварительной апробации диссертационной работы Сабитова Баратбека Рахмановича на соискание ученой степени доктора физико-математических наук на тему «Искусственный интеллект в задачах цифрового сельского хозяйства» по специальности 05.13.16 - применение вычислительной техники, математического моделирования и математических методов в научных исследованиях.

от 13 февраля 2024 г.

г. Бишкек

1. **Председатель:** Керимов У.Т. – к.т.н., и.о. доц. кафедры прикладная информатика КГУ им. И. Арабаева, 05.13.16 – применение вычислительной техники, математического моделирования и математических методов в научных исследованиях;
2. **Секретарь:** Кашкароева А.А. - к.соц.н., доцент кафедры прикладная информатика КГУ им. И. Арабаева, 22.00.04 – социальная структура, социальные институты и процессы.

ПРИСУТСТВОВАЛИ:

3. Бийбосунов Б.И. - д.ф-м.н., д.т.н., профессор кафедры прикладная информатика КГУ им. И. Арабаева, 05.13.16 - применение вычислительной техники, математического моделирования и математических методов в научных исследованиях, 05.13.13 – вычислительные машины, комплексы, системы и сети;
4. Эргашев С.Ф. – д.т.н., профессор, начальник научного отдела по подготовке научных кадров Ферганского политехнического института, Республика Узбекистан, г. Фергана, 05.05.06 – энергоустановки на основе возобновляемых видов энергии (приглашенный);
5. Сатыбаев А.Ж. – д.ф-м.н. (05.13.06), профессор, зав. кафедрой информационные технологии и управление ОшТУ, 05.13.18 – математическое моделирование, численные методы и комплекс программ;
6. Ашырбаева А.Ж. – д.ф-м.н., профессор, зав. кафедрой прикладная информатика ОшТУ, 01.01.02 – дифференциальное уравнение динамической системы и оптимальное управление;
7. Курманбек уулу Т. – д.т.н., доцент кафедры прикладная информатика КГУ им. И. Арабаева, 05.13.16 - применение вычислительной техники, математического моделирования и математических методов в научных исследованиях;
8. Юсупов К.М. – к.т.н., доцент кафедры прикладная информатика КГУ им. И. Арабаева, 05.13.13 – вычислительные машины, комплексы, системы и сети;
9. Асанбекова Н. О. - к.ф-м.н., и.о.доц. кафедры прикладная информатика КГУ им. И. Арабаева, 05.13.16 – применение вычислительной техники, математического моделирования и математических методов в научных исследованиях;
10. Кудакеева Г.М. - к.т.н., доц., 05.13.06 – автоматизация технологических процессов и производств;

11. Султанбаева Г. – к.п.н., и.о.доц. кафедры прикладная информатика КГУ им. И. Арабаева, 13.00.02 – теория и методика обучения и воспитания (математика);
12. Эсенгулов У.А. - и.о.доц. кафедры прикладная информатика КГУ им. И. Арабаева;
13. Бузурманкулова А.А. - старший преподаватель кафедры прикладная информатика КГУ им. И. Арабаева;
14. Ачекеев К.С. - старший преподаватель кафедры прикладная информатика КГУ им. И. Арабаева;
15. Барганалиева Ж.К. - старший преподаватель кафедры прикладная информатика КГУ им. И. Арабаева;
16. Садырова М.Р. - старший преподаватель кафедры прикладная информатика КГУ им. И. Арабаева;
17. Садыкова Л. - старший преподаватель кафедры прикладная информатика КГУ им. И. Арабаева;
18. Султанакунова А. - преподаватель кафедры прикладная информатика КГУ им. И. Арабаева;
19. Смайылбек кызы Ч. - старший преподаватель кафедры прикладная информатика КГУ им. И. Арабаева;
20. Турарбекова Н.Т. - преподаватель кафедры прикладная информатика КГУ им. И. Арабаева;
21. Мукудин уулу К. - преподаватель кафедры прикладная информатика КГУ им. И. Арабаева;
22. Абдувасиева М.А. - преподаватель кафедры прикладная информатика КГУ им. И. Арабаева.

Всего 22 человека.

ПОВЕСТКА ДНЯ:

1. Обсуждение диссертационной работы соискателя Сабитова Баратбека Рахмановича на тему: «Искусственный интеллект в задачах цифрового сельского хозяйства», представленной на соискание ученой степени доктора физико-

математических наук по специальности 05.13.16 - применение вычислительной техники, математического моделирования и математических методов в научных исследованиях.

Научный консультант – доктор физико - математических наук, доктор технических наук, профессор Бийбосунов Б.И.

Назначенные рецензенты:

1. Сатыбаев А.Ж. – д.ф-м.н. (05.13.06), профессор, зав. кафедрой информационные технологии и управление ОшТУ, 05.13.18 – математическое моделирование, численные методы и комплекс программ;
2. Эргашев С.Ф. – д.т.н., профессор, начальник научного отдела по подготовке научных кадров Ферганского политехнического института, Республика Узбекистан, г. Фергана, 05.05.06 – Энергоустановки на основе возобновляемых видов энергии.

ВЫСТУПИЛИ: Председатель к.т.н., и.о. доц. У.Т. Керимов объявил, что кворум имеется и ознакомил с процедурой проведения заседания кафедры. А затем предоставил слово для доклада соискателю.

СЛУШАЛИ: Доклад Б.Р. Сабитова на тему: «Искусственный интеллект в задачах цифрового сельского хозяйства», которая в течение 30 минут доложил цель работы, научную новизну, задачи исследования, основные положения и результаты, полученные в диссертационной работе.

В диссертационной работе исследуется использования современных методов искусственного интеллекта применительно к задачам сельского хозяйства. Кыргызская Республика является аграрной страной, и задача поддержки сельского хозяйства с достижениями науки и современной технологии является первостепенной задачей сегодняшнего дня. Чрезвычайно важной задачей является поддержка продовольственной безопасности страны. В работе подробно рассмотрены процессы построения различных моделей и прогнозирования с применением машинного обучения, глубокого обучения и компьютерного зрения для задач сельского хозяйства.

Актуальность темы диссертации.

Сельское хозяйство считается основой экономики и источником занятости в аграрных и развивающихся странах, например, таких как Кыргызская Республика. При этом основным ее составляющим является категория урожайность и тесно связанное с ним болезни растений сельскохозяйственных культур. Урожайность, которая встречается во многих сельскохозяйственных проектах относится к сложным категориям для моделирования и прогнозирования. **Машинное обучение** является важным инструментом поддержки принятия решений для прогнозирования урожайности. Весьма актуально при этом иметь инструменты исследования, основанные на научных достижениях задач моделирования и прогнозирования урожайности сельскохозяйственных растений и ее защиты от различных болезней. Изучение множество факторов, влияющих на здоровый рост выращиваемой культуры, которая в свою очередь является залогом урожайности не простейшая задача. В диссертации данная крупная задача исследуется с помощью построения искусственного интеллекта на базе нейронных сетей различной архитектуры и алгоритмов машинного обучения.

Цели и задачи исследования в диссертационной работе.

Целью диссертационной работы являлась использование современных методов искусственного интеллекта применительно к задачам сельского хозяйства. Исследуется, современные алгоритмы машинного обучения и глубокого обучения на основе базы данных климатических параметров, например, температура, количество осадков, тип почвы и разного рода пестициды.

Основной задачей диссертационной работы является использование современных методов искусственного интеллекта применительно к задачам сельского хозяйства. Для исследования задач сельского хозяйства рассмотрены следующие цели и задачи:

-исследования прогнозирование урожайности сельскохозяйственных культур с применением машинного и глубокого обучения;

- построения линейных и нелинейных моделей для задач сельского хозяйства;

- анализ результативности выявления основных признаков моделей на базе алгоритмов машинного обучения;

- исследование и получение визуализаций обобщающих результатов, обученных алгоритмами машинного обучения – дерево решений, случайного леса, градиентного бустинга в виде графического представления;

- изучение интенсивности скопления данных-дисперсии и смещения, вокруг линии регрессии алгоритмов машинного обучения, построения моделей наилучшей производительности для оценки точности;

- анализ эффективности методов случайный лес, метода опорных векторов, градиентного бустинга для прогнозирования различных категорий задач сельского хозяйства;

- получены результаты применения ансамблевых методов к задачам прогнозирования урожайности;

- визуализация производительности алгоритмов методов Лассо и Риджа для прогноза урожайности;

- реализация продвинутых алгоритмов машинного обучения: метод опорных векторов, K – ближайших соседей, варианты градиентного бустинга и случайный лес;

- выявления некоторых скрытых особенностей факторов, влияющих на урожайность выбранного региона, например, влияние изменения климата и применения пестицидов;

- расширение базы данных и сбор климатических и сельскохозяйственных параметров для построения моделей для конкретного региона (в качестве региона исследования, было выбрано Иссык-Кульская область) с более расширенными диапазонами и категориями для прогнозирования;

Рассматриваемые проблемы, можно идентифицировать как ресурс для будущих исследований и разработок в области использования глубокого

обучения в сельском хозяйстве. Для решения поставленной задачи автором диссертации решены следующие задачи:

Новизна диссертационной работы

В диссертационной работе с использованием методов машинного обучения исследуется обширный круг задач в сельском хозяйстве.

С применением глубокого обучения и различных архитектур нейронных сетей в диссертации исследован сложный раздел компьютерного зрения болезни сельскохозяйственных растений. Исследовано, также применения методов системы машинного и компьютерного зрения, используемых для классификации и обнаружения болезней растений сельского хозяйства. Построены различные модели по проблемам использования машинного обучения (МО), глубокого обучения (ГО) и их развертывания, которые можно имитировать как искусственный интеллект (ИИ) для моделирования задач сельского хозяйства, в частности получены следующие результаты:

- создано база данных и проведены работы по сбору собственных данных, состоящих из климатических и сельскохозяйственных параметров;

- для конкретного региона (в качестве региона исследования, было выбрано Иссык-Кульская область) был осуществлена сбор по болезням фруктовых деревьев груши по 4 видам и яблони по 4 видам болезней, по сельскохозяйственным растениям картофеля по 3 видам и кукурузы по 4 видам болезней; при построении моделей проводилось для конкретного региона с более расширенными диапазонами и категориями для прогнозирования;

- при создании моделей машинного обучения использованы различные методы линейной и множественной регрессии, машины опорных векторов (SVM), градиентный спуск, градиентный бустинг, дерево решений, метод LASSO, метод регуляризации А.Н.Тихонова-Риджа и другие;

- основным аппаратом при исследовании в диссертации использованы набор различных алгоритмов, которые моделируют высокоуровневые абстракции в данных используя архитектуры нейронных сетей, состоящие из

множества нелинейных преобразований и которые при обучении моделей, обнаруживают скрытые связи между данными;

-выбор алгоритмов машинного обучения осуществлялся в соответствии постановкой задач, структуры базы данных и уровнем их сложности;

-разработаны искусственный интеллект основанные на Фреймворках Python по распознаванию болезней растений.

Для фермеров и сельхозпроизводителей для оптимального управления урожайностью и определения, и идентификации болезней растений с помощью веб приложений были созданы веб системы на Django и Flask.

Были рассмотрены различные современные методы глубокого обучения для распознавания болезней различных сельскохозяйственных растений на основе трансферного обучения. Технология глубокого обучения в данной работе используются совместно с методами компьютерного зрения.

С помощью сверточных нейронных сетей, включая ResNet50, Inception, MobileNet и VGG16,19 и AlexNet, EfficientNetV2S, были получены результаты распознавания листьев растений с наибольшей точностью до 97-99%. При использовании данной технологии автором получены ряд результатов по распознаванию болезней груши по 4 болезням.

В диссертационной работе построены приложения для прогнозирования различных болезней сельскохозяйственных растений. Тем не менее, существуют проблемы, такие как сбор набора данных, время обучения моделей, необходимое количество данных для обучения и тестирования новыми данными, поддержка оборудования, развертывание больших моделей на небольших устройствах, таких как веб сайты или телефоны Android.

Связь темы исследований с приоритетными научными направлениями, основными научно-исследовательскими работами.

Применение машинного обучения в области сельского хозяйства имеет сегодня важную роль в развитии новых технологий для данной отрасли. Она заключается в разработке точных и настраиваемых моделей машинного

обучения, которые могут работать быстро, автоматически анализировать большие и сложные данные и помогать оптимизировать сельскохозяйственные процессы, такие как классификация, рекомендации или прогнозы.

Использование технологий глубокого обучения с технологиями компьютерного зрения в задачах прогнозирования играет весьма важную роль. Последнее связано с разработкой моделей, отвечающих современным требованиям. Моделирование и прогнозирование занимает ведущее положение в научных исследованиях прикладных задач. Данное направление тесно связано с методами искусственного интеллекта и моделированием для большого круга прикладных задач.

Структура диссертации

Диссертационная работа состоит из **Введения**, где отражено актуальность, цели и задачи исследуемой проблемы, тенденция развития и влияние методов машинного, глубокого обучения и компьютерного зрения, для решения задач сельского хозяйства. Диссертация состоит из четырех глав, в которых отражено основное содержание, методология и методы исследования, а также полученные результаты по моделированию и прогнозированию урожайности и болезни сельскохозяйственных культур, заключения и списка литературы и приложений. Диссертация изложена на 252 страницах компьютерного текста, содержит 6 таблиц, 133 рисунков, 75 листингов, а также содержит 9 приложений написанные на языке программирования Python, в которых приведены таблицы с данными, результаты по моделированию с применением машинного и глубокого обучения, а также численные расчеты, полученные на основе градиентного спуска.

В первой главе диссертации исследовано современное состояние и анализ публикаций по применению машинного, глубокого обучения и компьютерного зрения, опубликованные в последние годы для задач сельского хозяйства. Конкретно для задач распознавания изображений сделан обзор научных публикаций и технологий распознавания болезней широкого круга растений. Выделены основные направления для моделирования и прогнозирования задач сельского хозяйства. Предложены основные алгоритмы машинного обучения и

методы глубокого обучения, основанные на нейронных технологиях. **Во второй главе** диссертации отражено полный анализ и исследования, а также методы и методологии машинного обучения, ориентированной для задач цифрового сельского хозяйства и математическое описание методов машинного обучения. Рассмотрено анализ построения моделей и методология применения алгоритмов машинного обучения для широкого круга прикладных задач сельского хозяйства в том числе. **Третья глава** диссертации посвящена построению моделей для задач цифрового сельского хозяйства алгоритмами и методами машинного обучения. Представлен обзор соответствующих результатов, полученных автором на основе мощных алгоритмов машинного обучения для прогнозирования урожайности и распознавания болезней растений по различным сельскохозяйственным культурам. Приведены численные результаты, выполненные с помощью методов оптимизации - градиентного спуска. Несмотря на множество публикуемых научных работ по этой области, численные подходы в задачах машинного обучения еще остается проблемой для многих сельскохозяйственных задач. Так как найти точный прогноз урожайности не удается во многих случаях, поэтому численные подходы для этой области необходимо расширить.

Обзор научных исследований показывает, что сейчас более внимательно изучают применение машинного и глубокого обучения в этой области, для более точного моделирования. Рассмотренные выше технологии машинного обучения и результаты, полученные автором по глубокому обучению, рассмотрены в **Главе 3**. Серьезному направлению исследования задач сельского хозяйства, одного из точных и многообещающих методов искусственного интеллекта является глубокое обучение, и методы распознавания объектов с помощью компьютерного зрения с различными архитектурами и технологиями нейронных сетей, которое занимает лидирующее положение, рассмотрено в **Главе 4** диссертационной работы. Изучению множества сложных нелинейных взаимосвязей между признаками и данными, влияющих на точное моделирование и прогнозирования урожайности и болезни растений по

изображениям растений на основе компьютерного зрения, которая включает мощные алгоритмы машинного обучения и проектирования нейронных сетей глубокого обучения различной архитектуры, а также с предварительно обученными нейронными сетями. При этом по второму направлению, использованы различные архитектуры нейронных сетей, включая трансферное обучение моделей, основой которых являются обученные модели на BigData.

Апробация диссертационной работы.

В диссертационной работе исследованы методы машинного, глубокого обучения, компьютерного зрения и построения моделей и прогнозирования для многочисленных задач сельского хозяйства. Основные результаты диссертации апробированы под непосредственным руководством автора во многих проектах МОиН за 2017 -2024 годы. Задачи и проблемы в ходе выполнения проектов включены в Государственные программы по искусственному интеллекту, как приоритетные научные направления Кыргызской республики и Министерства образования и науки (МОиН), финансируемых МОиН КР. Работа выполнена на кафедре Прикладной информатики Кыргызского государственного университета им.И.Арабаева:

1.«Разработка и создание новых информационных технологий и интеллектуальной экспертной системы для сферы АПК КР (инвестиционные процессы, инфраструктура и логистика АПК)», 2017 год.

2.«Моделирование и прогнозирование в сфере АПК КР с применением интеллектуальных систем, Python технологий и нейронных сетей», 2018-2021гг.

3.«Технологии внедрения искусственного интеллекта в систему общеобразовательного образования», 2021-2022 гг.

4. Искусственный интеллект в сельском хозяйстве (2023-2025 гг.)

5. Прогнозирование регионального целевого устойчивого развития аграрной и экологической систем в условиях изменения климата с применением искусственного интеллекта (2024-2026 гг.).

Научные результаты, полученные в исследовательской диссертационной работе, докладывались во многих международных и Вузовских конференциях

КР в частности:

-Международная научная конференция “Технологии и перспективы современного инженерного образования, науки и производства”, посвященная 45-летию ФПИ – КТУ им. И. Раззакова – Бишкек, 1999 г.

-Международная научно-практическая конференция. «Применения цифровых технологий в образовании: проблемы и перспективы». Вестник Кыргызского национального университета им. Ж.Баласагына, Труды. Бишкек. 2019 г.

-Международная научно-практическая конференция. «Научно-технологическое развитие АПК для целей устойчивого развития»

«Моделирование и прогнозирование задач сельского хозяйства на основе машинного обучения». Труды. Бишкек. 2022 г.

-Межвузовская научно-практическая конференция «Цифровые технологии в отраслях производства и социальной сфере», 27 окт. 2022, Астана, Дакка

- Международная научно-практической конференция. «Роль науки и инновационных технологий в устойчивом развитии горных территорий и экосистем». 27-28 октября 2022 г. Бишкек, Кыргызская Республика.

- Макеты для мониторинга атмосферы. XV Международная конференция ученых, 26 - 28 апреля 2023 года г. Бишкек, 398-402 стр. 2023_НС РАН_сборник XV Международной конференции [1].pdf

Значимость полученных результатов и практические реализации. Все основные научно-исследовательские работы соискателя имеют прикладной характер, и основные научные результаты имеют значительную внедренческую ценность, в Государственную программу по продовольственной безопасности страны, сельское хозяйство в целом и искусственному интеллекту. Многие модели построенные, в диссертации основываются на реальных данных и имеют ценное практическое значение в прогнозировании задач сельского хозяйства.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Основные результаты, полученные автором по моделированию и прогнозированию задач сельского хозяйства разделены на два больших раздела.

В первой части исследованы моделирование и прогнозирование задач

сельского хозяйства на основе машинного обучения. В данном разделе исследуются результаты, полученные автором по применению машинного обучения к задачам сельского хозяйства. При исследовании моделирования и прогнозирования урожайности использовались алгоритмы машинного обучения: множественная регрессия (LR), регрессия Лассо (Lasso R), стохастический градиентный спуск (SGD), дерево решений (RT), которые дают хорошие результаты для многих сельскохозяйственных задач. Используются алгоритмы К – ближайших соседей (KNN), случайный лес (RF), метод опорных векторов (SVR) и варианты градиентного бустинга (GBR). В базу данных, урожайность, которая зависит от многих факторов, включены зависимости от климата, погоды, почвы и обработки посевных площадей. Рассматриваются задачи классификации и регрессии с применением множества алгоритмов машинного обучения для задач урожайности. Подробно проанализированы алгоритмы случайного леса, градиентного бустинга и его варианты применительно к задачам различных культур.

Данные, используемые в работе, были извлечены из источников пяти районов Иссык-Кульской области Кыргызской Республики и представляют собой вносимые и обработанные поля, удобрения (азот, фосфор и калий), температуру, влажность, осадки и кислотность почвы для пяти наименований районов, а также урожайность картофеля для каждого из районов. Для удобства работы с библиотеками Python при сборе данных учитывались особенности данных каждого из регионов и была представлена в виде обобщенного .csv файла.

Обучение линейных моделей задач сельского хозяйства на основе машинного обучения

В диссертационной работе подробно изучается процесс построения линейных моделей, как основа нелинейного моделирования. Она является и основой всей теории машинного и глубокого обучения. С применением линейной регрессии построена модель для урожайности сельскохозяйственных культур, в которой мы стремимся прогнозировать p наблюдений переменной

отклика урожайности $Y = \text{harvest}$ с линейной комбинацией p переменных предикторов X и нормально распределенной ошибкой с дисперсией σ^2 , которое можно представить формулой

$$Y = X\omega + \varepsilon, \text{ где } \varepsilon \sim N(0, \sigma^2) \quad (1)$$

где в (1) $N(\mu, \sigma^2)$ нормальное распределение с математическим ожиданием $\mu = 0$ и дисперсией σ^2 , которая определяется следующей формулой

$$N(\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (2)$$

Зависимая переменная Y урожайности в качестве первоначального представления можно представить в виде модели множественной регрессии, которая имеет следующий вид

$$Y = \omega_0 + \omega_1 X_1 + \omega_2 X_2 + \dots + \omega_p X_p + \varepsilon, \quad (3)$$

в которой Y – зависимая переменная урожайность, X_p – независимые переменные составляющие множественной регрессии, ω_p – неизвестные весовые коэффициенты, а ε – ошибка модели определяемая формулой $\varepsilon \sim N(0, \sigma^2)$.

Выражение с условием минимума квадратичного функционала можно переписать в векторной форме:

$$\begin{aligned} L(X, \vec{y}, \vec{\omega}) &= \frac{1}{2n} \sum_{i=1}^n (y_i - \vec{\omega}^T \vec{x}_i)^2 = \frac{1}{2n} \|\vec{y} - X\vec{\omega}\|_2^2 = \\ &= \frac{1}{2n} (\vec{y} - X\vec{\omega})^T (\vec{y} - X\vec{\omega}) \rightarrow \min \text{ по } \omega \quad (4) \end{aligned}$$

Тогда из условия минимума функционала (4) следует

$$\vec{\omega} = (X^T X)^{-1} X^T \vec{y}, \quad (5)$$

Вообще говоря, расчеты, проведенные по алгоритму (6) в одномерном и по аналогичной формуле в многомерном случае, приводят к переобученным моделям и не обладают обобщающими свойствами, теряя при этом смысл для тестовых данных. Модель созданный данным алгоритмом будет сколь угодно интерполировать данные вместо экстраполяции при тестировании данных. Для решения данной проблемы вводится некоторая регулирующая функция $R(\vec{\omega})$, и формулируется как задача определения весовых коэффициентов $\vec{\omega}$ с преобразованным функционалом ошибок в виде (6)

$$\Omega(X, \vec{y}, \vec{\omega}) = L(X, \vec{y}, \vec{\omega}) + \lambda R(\vec{\omega}) \Rightarrow \min \quad (6)$$

где λ - называется коэффициентом регуляризации, E-единичная матрица.

В данном случае алгоритм нахождения коэффициентов $\vec{\omega}$ из условий (6) определяется по формуле (7)

$$\vec{\omega} = (X^T X + \lambda E)^{-1} X^T \vec{y}, \quad (7)$$

Ниже приведены реализация данного алгоритма с применением регулирующих алгоритмов задачи (6), (7). В качестве применения данного алгоритма ниже для конкретной задачи сельского хозяйства прогнозирования, влияния количества вносимых пестицидов с применением регуляризации Тихонова-Риджа, выглядит следующим образом.

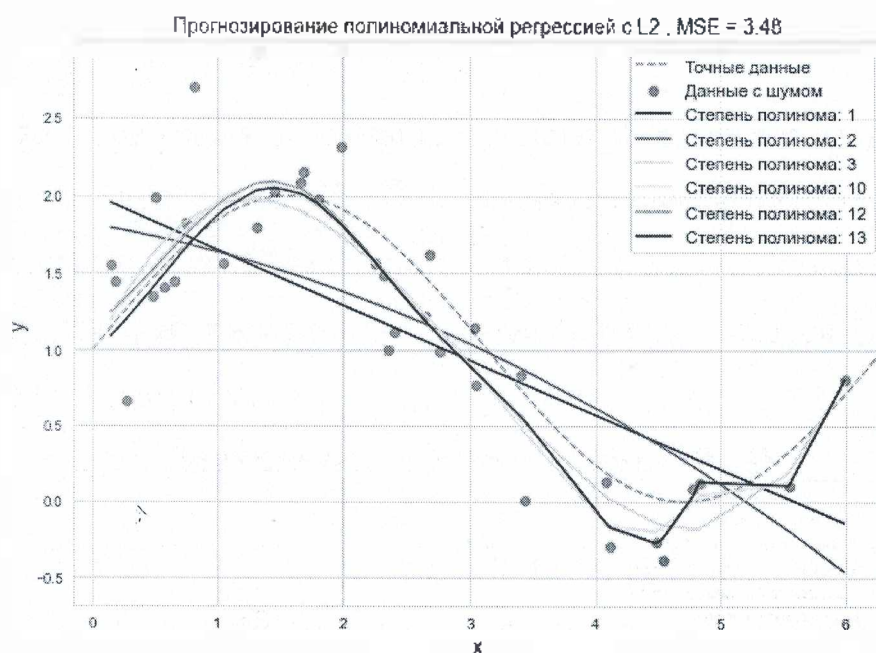


Рисунок 1. Прогноз урожайности полученная с применением регулирующего алгоритма.

Целевая переменная и погодные условия распределены по нормальному закону. Ниже представлены результаты, как урожайность связаны с другими переменными. Корреляционная матрица, которая отражает связи между целевой и независимыми переменными выглядит так.

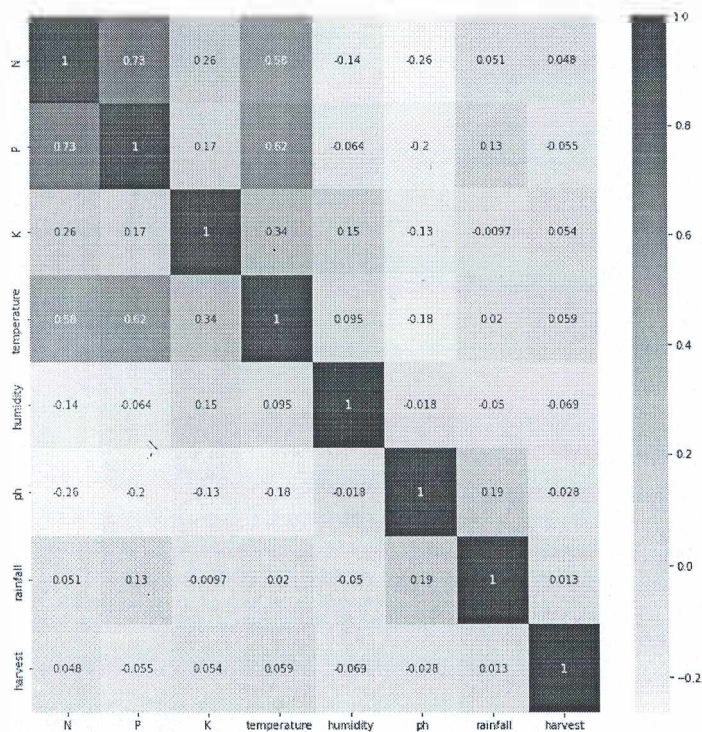


Рисунок 2. Корреляционная матрица связи целевой переменной и другими функциями.

Полиномиальная регрессия как расширение линейных моделей базисными функциями.

Результаты диссертации показали, что полиномы более высокой степени могут лучше соответствовать данным. Ниже представлен результат аппроксимации полиномами высокой степени для задачи о пестицидах, для которой в простейшем случае имеем.

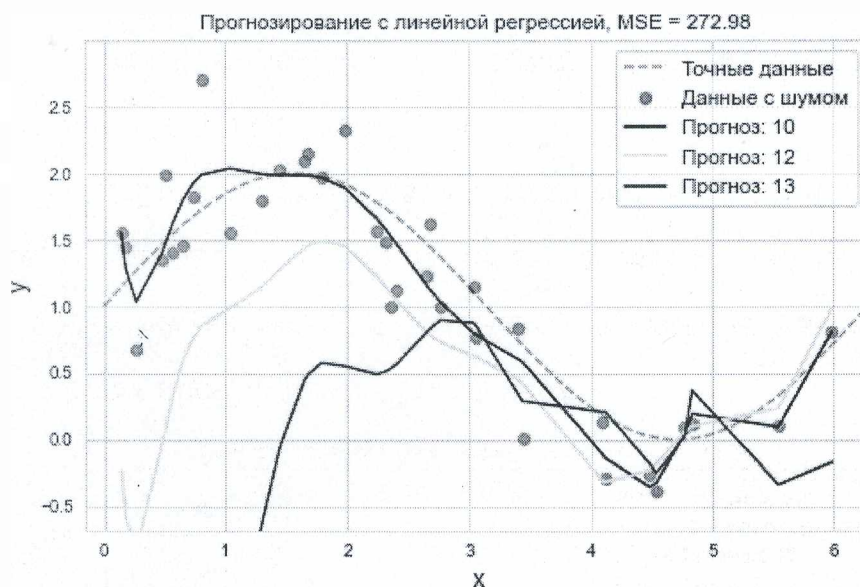


Рисунок 3. Прогнозирование с помощью полиномов высокой степени.

Но в то же время слишком большие степени могут показать, нежелательное колебательное поведение и особенно опасны для экстраполяции за пределы диапазона подобранных данных. В этом случае модель переобучается, и она не обладает способностью к обобщению. Вот коэффициенты аппроксимации в случае аппроксимации полиномом степени 13 видно на рис.4.

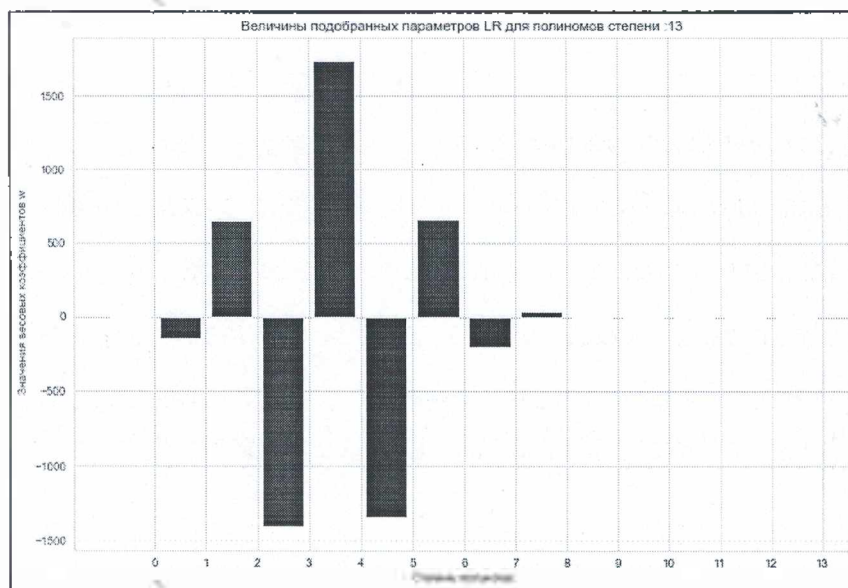


Рисунок 4. Коэффициенты регуляризации.

Алгоритм, построенный из условия минимума функционала $L(X, \vec{y}, \vec{\omega})$ по весовым ω называется алгоритмом Тихонова – Риджа. Параметр сложности $\lambda > 0$ контролирует величину стабилизирующего слагаемого: чем больше значение λ , тем больше величина стабилизирующего слагаемого и, следовательно, коэффициенты становятся более устойчивыми к коллинеарности. При увеличении параметра регуляризации, матрица $X^T X + \lambda E$ становится "менее сингулярной", т.е. регулярной, а задача становится более определенной. Для такой матрицы число обусловленности будет равно: $(e^{max} + \lambda) / (e^{min} + \lambda)$, e^x где — это собственные числа матрицы. Таким образом, увеличивая параметр регуляризации мы, уменьшаем число обусловленности. Прогнозирование влияния количества вносимых пестицидов с применением полиномиальной регрессии на почвенные характеристики с применением регуляризации Тихонова-Риджа, выглядит следующим образом.

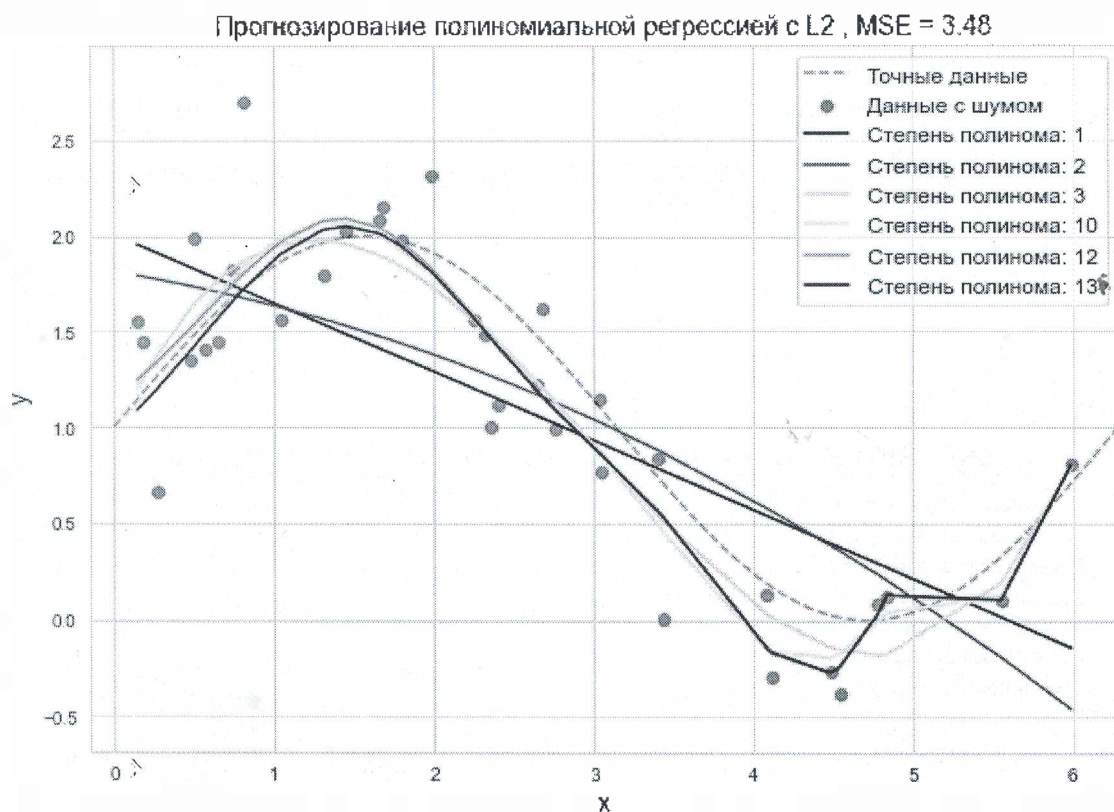


Рисунок 5. Результат применения технологии регуляризации для прогнозирования влияния пестицидов на урожайность картофеля.

Численные методы в задачах сельского хозяйства

В работе реализованы алгоритмы градиентного спуска и их численные результаты. Задача минимизации функционала издержек $J(\theta)$ на каждой итерации значение θ обновляется по следующему правилу

$$\theta^{n+1} = \theta^n - \alpha \frac{\partial}{\partial \theta} J(\theta^n), \quad (8)$$

где α — размер шага или скорость обучения. Если α слишком мало, алгоритм минимизации для сходимости требуется много времени, если α слишком велико, алгоритм может расходиться. Ниже приведена визуализация результатов расчета, по методу (8)

Пакетный градиентный спуск $\eta = 0.005$ & Эпохи = 500



Рисунок 6. Применения пакетного градиентного спуска к задаче регрессии
 Приведем результат точности и визуализации и процесс обучения реализации стохастического градиентного спуска- `StochasticGDLinearRegression`, со скоростью обучения $\eta=0.005$ и эпохой 20:

График обучения стохастического градиентного спуска на основе времени

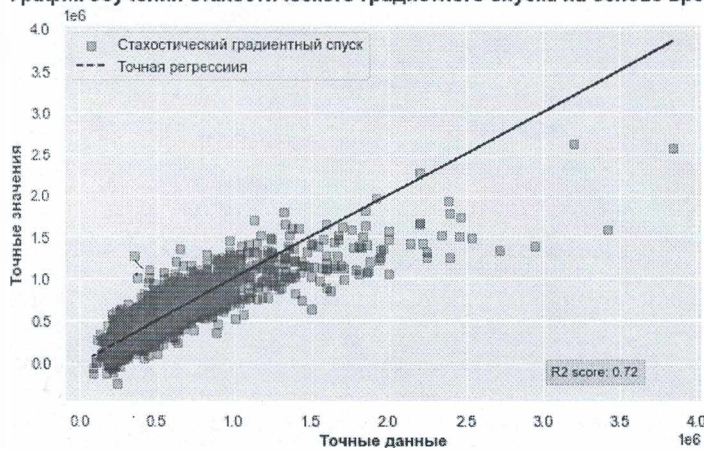


График обучения

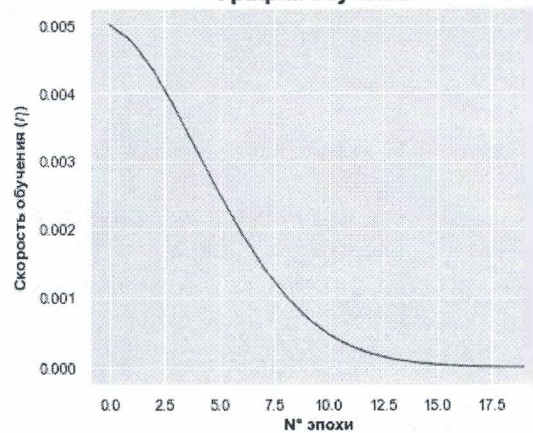


Рисунок 7. Применения стохастического градиентного спуска к задаче регрессии.

ЗАКЛЮЧЕНИЕ

В диссертационной работе проведено обширное исследование применения различных алгоритмов и методов машинного обучения и методов глубокого обучения для распознавания и классификации болезней сельскохозяйственных растений. Построены различные модели прогнозирования урожайности сельскохозяйственных культур. Используются также методы классификации и распознавания на основе компьютерного зрения, которые могут использоваться для обнаружения болезней растений, а также для помощи фермерам в автоматическом обнаружении всех видов болезней. В этом направлении применяются различные архитектуры нейронных сетей в глубоком обучении с применением технологий компьютерного зрения для задач болезни растений. Кроме того, были обобщены несколько методов/сопоставлений для распознавания симптомов заболевания. Здесь развитие технологий глубокого обучения в последние годы для выявления болезней листьев растений. Мы ожидаем, что эта работа станет полезным инструментом для ученых, занимающихся выявлением болезней растений. Также, также проводится сравнительное исследование между методами машинного и глубокого обучения. Несмотря на то, что в последние годы был отмечен значительный заметный прогресс, все еще остаются некоторые пробелы в исследованиях, которые необходимо устранить и внедрить эффективные методы обнаружения болезней растений.

1. Изучены прогнозирования задач сельского хозяйства на основе машинного обучения и ее алгоритмов. Построены модели для прогнозирования урожайности сельскохозяйственных культур.
2. На основе глубокого обучения с различными архитектурами нейронных сетей построены различные модели классификации и регрессии для урожайности.
3. Предложены новые модели, основанные на технологиях машинного обучения с алгоритмами многофакторного анализа, дерево решений, случайный лес, градиентный бустинг, стохастический градиентный

бустинг и распознавания образов для практических задач сельского хозяйства.

4. Разработаны нелинейные модели, информационные технологии и системы на базе глубокого обучения и нейронных сетей с элементами искусственного интеллекта.
5. В целях развития процессов цифровизации сельского хозяйства разработаны искусственные интеллекты распознавания болезней растений и управления урожайности для различных сельскохозяйственных культур. Данные технологии необходимы для дальнейшего развития сельского хозяйства и в целом для ее цифровизации.
6. В диссертации исследован современный подход прогнозирования задач сельского хозяйства на основе глубокого обучения с различными архитектурами нейронных сетей и технологии компьютерного зрения. Рассмотрены также прогнозирования урожайности с использованием нейронных сетей различной архитектуры и построены регуляризующие алгоритмы для уменьшения ошибки моделей.
7. Подробно рассмотрены методы обучения переобученных моделей с применением метода регуляризации. Отдельно было рассмотрены технологии интеграции алгоритмов машинного обучения и современного подхода в распознавании изображений сверточные нейронные сети. Методы CNN занимают особое место в прогнозировании многих задач сельского хозяйства. Одним из них является метод распознавания болезней растений, которое изучалось в данной диссертации. Следующим важным методом глубокого является трансферное обучения моделей с применением компьютерного зрения и больших данных. В данном направлении рассмотрены уже обученные модели на больших данных ImageNet. В частности, для трансферного обучения рассмотрены модели

трансферного обучения VGG16 и ResNet для прогнозирования болезней растений основанные на платформах больших данных PlantVillage.

8. Определено, что для определения болезней растений использованное с различными архитектурами нейронных сетей и основанные на глубоком обучении с технологиями компьютерного зрения ResNet значительно лучше работают для классификации изображений, когда некоторые параметры настраиваются и применяются такие методы, как планирование скорости обучения, отсечение градиента и уменьшение веса. Модель способна идеально предсказать каждое изображение в тестовом наборе без каких-либо ошибок.
9. Важной проблемой, является использования Фреймворков Python для построения веб систем, которые используют уже обученные модели для задач прогнозирования. В данном направлении построены системы искусственного интеллекта по определению урожайности и рекомендации по использованию пестицидов в сельском хозяйстве. Построены искусственные интеллекты по распознаванию болезней растений для широкого круга задач по рекомендации лечения болезней растений по результатам прогнозирования. Модели для прогнозирования задач распознавания болезней многих растений проверены на реальных тестовых данных Иссык-Кульской области.

В целом изученные проблемы и новые достижения в машинном и глубоком обучении показали результаты удовлетворяющие современные требования к прогнозированию и моделированию сложных нелинейных процессов.

При обсуждении диссертационной работы были заданы следующие вопросы:

Вопросы У.Т. Керимова - к.т.н., и. о. доц., КГУ им. И. Арабаева -

1. Коротко расскажите об актуальности темы диссертации?
2. Какие задачи в диссертационной работе являются необходимыми?

Ответы:

1. Кыргызская Республика (КР) является аграрной страной, сельское хозяйство является одним из основных видов деятельности жителей республики. Сельское хозяйство имеет, существенное значения, для выполнения продовольственной программы являясь экономически важной проблемой номер один.

Деятельность фермерско-крестьянских хозяйств напрямую связано с задачей выполнения продовольственной программы. Например, садоводство — это глобальный бизнес в мире с оборотом 5 триллионов долларов, который будет продолжать расти с помощью достижений искусственного интеллекта для увеличения урожая, борьбы с вредителями растений насекомыми, влияния климата, проверки почвы и улучшения условий выращивания. Из этого простейшего примера видно, что применение ИИ, которая в настоящее время имеет, большие обороты во всех отраслях и в настоящее время имеет важное значения для всей экономики страны. В данной диссертационной работе с помощью элементов искусственного интеллекта построены различные модели прогнозирования урожайности, с применением компьютерного зрения построены модели распознавания болезней сельскохозяйственных растений.

Основными задачами в диссертационной работе являются:

1. Прогнозирование и построение моделей для задач сельского хозяйства на основе машинного, глубокого обучения и компьютерного зрения.
2. Построены различные модели прогнозирования урожайности. Рассмотрены различные задачи распознавания болезней растений для регионов страны.

3. Созданы искусственный интеллект прогнозирования урожайности на конкретных видах растений. Построен искусственный интеллект распознавания болезней растений, выращиваемый в Иссык-Кульской области.

Вопросы: Юсупов К.М.- к.т.н., доцент КГУ им. И. Арабаева -

1. Какие методы применялись для построения моделей урожайности?

Ответы:

1. Для изучения взаимосвязи в данных для задач урожайности сельскохозяйственных культур в основном использованы методы машинного и глубокого обучения. Для изучения сложных взаимосвязей в данных и задач прогнозирования применялись продвинуты алгоритмы машинного обучения как метод ближайших соседей, случайный лес, градиентный бустинг и др. Построены алгоритмы машинного обучения случайный лес, метод опорных векторов и метод основанной на регуляризации переобученных моделей, на основе метода А.Н. Тихонова. С помощью этих алгоритмов обучены и построены различные модели для прогнозирования задач урожайности. Рассмотрены задачи линейной и многомерной регрессии, классификации болезней растений. Созданы программные пакеты для алгоритмов машинного обучения, с учетом важных признаков базы данных. С помощью глубокого обучения обучены модели для задач сельского хозяйства.

Вопрос: Эргашев С.Ф. – д.т.н., профессор - Какие параметры-признаки играли роль при построении моделей и прогнозирования?

Ответ: Основными параметрами при прогнозировании являются типы вносимых удобрений, почвенные характеристики изучаемого региона, климатические данные как суточные минимальные и максимальные значения температуры, влажность воздуха, осадки, кислотность почвы и др.

Вопрос: Кашкароева А.А.-к.соц.н., доцент, КГУ им.И.Арабаева –

1. В процессе разработки моделей и прогнозирования с помощью нейронных

технологий большую роль играет компьютерное зрение, какие технологии здесь применяются в диссертации?

2. Использовались ли технологии переноса обучения при распознавании болезней растений?

Ответы:

1) Глубокого обучения в технологиях компьютерного зрения имеет большое значение для задач прогнозирования. Особенность этого исследования использовать для раннего распознавания болезней сельскохозяйственных растений, когда имеются множество болезней, применяется великолепная технология сверточные нейронные сети- CNN в сочетании с технологиями компьютерного зрения. Еще одной особенностью является использование системы Keras с платформой Tensorflow.

2) Трансферное или перенос обучение обычно относится к процессу, в котором модель, обученная одной проблемой, каким-то образом используется для решения второй связанной проблемы. Трансферное обучение является ключевым моментом для обучения многих нейронных сетей в связи с отсутствием больших данных. Данная технология в диссертации применяется для задач сельского хозяйства, когда размер данных не так уж велико.

Вопрос: Сатыбаев А.Ж. – д.ф-м.н. (05.13.06), профессор, зав. кафедрой информационных технологий и управление ОшТУ – В методах глубокого обучения, для построения моделей и прогнозирования используют большое количество данных, как избежать этого момента?

Ответ: когда данные имеют ограниченное количество используют обычно технологии увеличения данных. Этот мощный аппарат глубокого обучения позволяет иногда увеличить данные в два раза и более раз. В диссертации данная технология использовалась, когда рассматривалась задача распознавания болезней растений с помощью методов компьютерного зрения.

Выступили:

Председатель заседания: Уважаемые коллеги! Если нет больше вопросов, предлагаю перейти к обсуждению диссертации.

В ходе обсуждения были следующие выступления: Керимов У.Т. - к.т.н., и. о. доц.; Юсупов К.М.- к.т.н., доцент; Эргашев С.Ф. – д.т.н., профессор; Кашкароева А.А.-к.соц.н., доцент; Сатыбаев А.Ж. – д.ф.-м.н. (05.13.06), профессор; Сабитов Б.Р.- к.ф.-м.н., доц.; Бийбосунов Б.И. - д.ф.-м.н., д.т.н., проф., научный консультант.

Во всех выступлениях были высказаны положительные отзывы и пожелания по диссертации Сабитова Б.Р.

Председатель: Слово для ответа по выступлениям предоставляется соискателю. Сабитову Б.Р.: Уважаемые коллеги! Разрешите мне поблагодарить вас за внимательное и объективное отношение к моей диссертационной работе. Я согласен с высказанными замечаниями, они будут учтены, от этого качество работы только улучшится.

Председатель: Уважаемые коллеги, на сегодняшнем заседании было задано достаточно вопросов, были выступления по обсуждаемой диссертационной работе. Если нет возражений, предлагаю подвести итоги и перейти к выработке решения нашего заседания и заключения по данной работе. Поступило предложение рекомендовать диссертационную работу к защите.

После обсуждения и обмена мнениями участники расширенного заседания кафедры прикладная информатика института новых информационных технологий КГУ им. И. Арабаева единогласно (голосовали: за – 22, против – нет, воздержались – нет) приняли следующее постановление

ПОСТАНОВИЛИ:

Диссертационная работа Сабитова Баратбека Рахмановича на тему: «Искусственный интеллект в задачах цифрового сельского хозяйства», представленной на соискание ученой степени доктора физико-математических

наук по специальности 05.13.16 - применение вычислительной техники, математического моделирования и математических методов в научных исследованиях, содержит научные и практические рекомендации, позволяющие решать ключевые проблемы сельского хозяйства Кыргызской Республики, как одно из важных проблем экономики страны выполнения продовольственной программы. Настоящая диссертационная работа представляет собой вполне завершённый этап научных исследований, её результаты могут служить необходимой базой и исходными позициями для проведения дальнейших исследований и разработок при использовании искусственного интеллекта для решения многих прикладных задач в том числе задач сельского хозяйства.

Работа отвечает требованиям ВАК, предъявляемым к диссертационным работам на соискание ученой степени доктора наук, и рекомендуется для представления в Диссертационный совет для защиты на соискание ученой степени доктора физико – математических наук по специальности 05.13.16 - применение вычислительной техники, математического моделирования и математических методов в научных исследованиях.

Председатель:

к.т.н., и.о. доц.

каф. прикладная информатика

КГУ им. И. Арабаева

Керимов У.Т.

Секретарь:

к.соц.н., доцент

каф. прикладная информатика



Кашкароева А.А.

13 февраля 2024 г.

Подписи заверяю:

И. АРАБАЕВ атындагы КЫРГЫЗ МАМЛЕКЕТТИК УНИВЕРСИТЕТИ
КЫРГЫЗСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ И. АРАБАЕВ
КОЛ ТАМГАСЫН ТАСТЫКТАЙМЫН
ПОДПИСЬ ЗАВЕРЯЮ