облусунда социалдык инфраструктураны жакшыртуу - жашоо деңгээлин көтөрүүнүн маанилүү аспектилеринин бири. Жолдорду, мектептерди жана медициналык мекемелерди модернизациялоо калкка сапаттуу кызмат көрсөтүүлөрдү камсыз кылат жана ошонун негизинде калктын жашоо шарты жакшырат. Талас облусундагы калктын жашоо деңгээлин көтөрүү үчүн экономикалык, социалдык жана инфраструктуралык көйгөйлөрдү чечүү керек. Келечекте Талас облусунун экономикасын өнүктүрүү жаан ошонун неизинде калктын жашоо деңгээлин мындан ары жогорулатуу боюнча жакшы мүмкүнчүлүктөр бар экендиги талашсыз, бирок бул үчүн мамлекеттик, жеке жана коомдук секторлордун биргеликтеги аракеттери талап кылынат. Талас облусунун уникалдуу ресурстарын натыйжалуу пайдалануу жана инновацияларга таянуу калктын жашоо деңгээлин жакшыртууда негизги роль ойнойт.

Колдонулган адабияттар:

1. Статистикалык жыйнак: Кыргыз Республикасынын өнөр жайы, Басылмалар, 2019-2023жж., Бишкек, // https://www.stat.gov.kg/;
2. Электронный источник://https://barometr.kg/kg/zherj-altyn-kenin-ishtetch-alyans-altyn-kompaniyasy-talas-oblusun-nktrg-900-mln-som-blgn;
3. Статистикалык жыйнак: Кыргыз Республикасынын айыл чарбасы, Жылдык басылмалар, 2019-2023 жж., Бишкек // https://www.stat.gov.kg/;
4. Статистикалык жыйнак: Регионалдык статистика, Талас облусунун социалдык-экономикалык экономикалык абалы, Жылдык басылма, Талас, 2023//https://www.stat.gov.kg/;
5. Жээнбаев, А. "Талас облусунун айыл чарба потенциалы". Кыргыз агрардык университети, 2021.

УДК: 004.85
DOI 10.33514/BK-1694-7711-2024-2(1)-401-411

Сейтказиева Н. С., Иса кызы Барчын, Курманбек уулу Талантбек, Сабитов Б. Р.
И.Арабаев атындагы Кыргыз мамлекеттик университети, окутуучу,
И.Арабаев атындагы Кыргыз мамлекеттик университети, магистрант,
И.Арабаев атындагы Кыргыз мамлекеттик университети, техника илимдеринин доктору, доцент,
Ж.Баласагын атындагы Кыргыз улуттук университети, физика-математика илимдеринин кандидаты, доцент

Сейтказиева Н. С., Иса кызы Барчын, Курманбек уулу Талантбек, Сабитов Б. Р.
Кыргызский государственный университет имени И. Арабаева, преподаватель,
Кыргызский государственный университет имени И. Арабаева, магистрант,
Кыргызский государственный университет имени И. Арабаева, доктор технологических наук, доцент,
Кыргызский национальный университет Ж. Баласагына, кандидат физико-математических наук, доцент

Seytkazieva N. S., Isa kyzy Barchyn, Kurmanbek uulu Talantbek, Sabitov B. R.
Kyrgyz State University named after I. Arabaev, teacher,
Kyrgyz State University named after I. Arabaev, Master's student,
Kyrgyz State University named after I. Arabaev, Doctor of Technological Sciences, Associate Professor,
Kyrgyz National University named after J. Balasagyn, Candidate of Physical and Mathematical Sciences, Associate Professor

МАШИНАЛАРДЫ ҮЙРӨНҮҮ АЛГОРИТМДЕРИН КОЛДОНУУ МЕНЕН КОЛДОНМО ТАПШЫРМАЛАРДЫ МОДЕЛДӨӨ ЖАНА БОЛЖОЛДОО
МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ ПРИКЛАДНЫХ ЗАДАЧ С ПРИМЕНЕНИЕМ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

MODELING AND FORECASTING OF APPLIED PROBLEMS USING MACHINE LEARNING ALGORITHMS

Аннотация: Машина менен окутууну колдонуу менен ар кандай айыл чарба өсүмдүктөрүнүн түшүмүн болжолдоо маселеси изилденет. Айыл чарба проблемаларын болжолдоонун моделдери алынды. Негизги фундаменталдуу идея - бул четтөөлөрдү аныктоо үчүн аралыкты баалоону колдонуу. Регрессия жана классификация жакынкы кошуналарга негизделген жана иш жүзүндө биз жакынкы кошуналарды болжолдуу издөөнүн белгилүү ыкмаларын чогултабыз. Кошумчалай кетсек, биз жакынкы кошуна ыкмалары менен колдонуу үчүн аралыктарды үйрөнүү менен байланыштарды талкуулайбыз. Бул макалада машиналык окутуу ыкмаларын колдонуу менен айыл чарба көйгөйлөрүн талдоо жана прогноздоо үчүн ар кандай ыкмалар майда-чүйдөсүнө чейин каралат.

Аннотация: Исследуется задача прогнозирования урожайности различных сельскохозяйственных культур с применением машинного обучения. Получены модели для прогнозирования задач сельского хозяйства. Основная фундаментальная идея заключается в использовании оценки расстояния для выявления выбросов. Регрессии и классификации на основе ближайших соседей, так и на практике, для которой мы собираем известные методы приближенного поиска ближайших соседей. Кроме того, мы обсуждаем связи с расстояниями обучения для использования с методами ближайшего соседа.В данной работеподробно исследовано различные методы анализа и прогнозирования задач сельского хозяйства методами машинного обучения.

Abstract: The problem of forecasting the yield of various crops using machine learning is investigated. Models for forecasting agricultural tasks have been obtained. The basic fundamental idea is to use distance estimation to identify outliers. Regression and classification based on nearest neighbors, as well as in practice, for which we collect well-known methods of approximate nearest neighbor search. In addition, we discuss connections with learning distances for use with nearest neighbor methods.In this paper, various methods of analysis and forecasting of agricultural tasks using machine learning methods are studied in detail.

Негизги сөздөр: машиналык окутуу, алгоритм, технология, маалымат базалары, чечим дарагы, сызыктуу эмес модель, болжолдоо, жакынкы коңшулук ыкмасы.

Ключевые слова: машинное обучение, алгоритм, технология, базы данных, дерево решений, нелинейная модель, прогнозирование, метод ближайших соседей.

Keywords: machine learning, algorithm, technology, databases, decision tree, nonlinear model, forecasting, nearest neighbor method.

В данной работе рассмотрено основы применения машинного обучения к прикладным задачам. Исследуется алгоритмы машинного обучениядля задач сельского хозяйства. Для исследовательских целей мы изучим популярный алгоритм для прикладных целей метод ближайших соседей(KNN). Первоначально методы ближайшего соседа обнаруживают глобальные выбросы, а затем присваивают им оценки выбросов. В классификации KNN информация о расстоянии исследуется от точки до ее соседа, независимо от того, близко оно или нет. Данная технология является самым простым в решении многих прикладных задач классификации. Суть метода такова, посмотри на соседей, какие преобладают, таков и ты. В данном случае, если метрика расстояния между примерами введена достаточно удачно, то схожие примеры гораздо чаще лежат в одном классе, чем в разных. Для классификации каждого из объектов тестовой выборки необходимо последовательно выполнить следующие операции[1].

Прикладные задачи.

Рассмотрим задачу прогнозирования с применением методов машинного обучения деревья решений. Далее исследуем также, метод ближайших соседей. Рассматривается задача прогнозирования оттока фермеров из сельского хозяйства.В нашем случае переведем данные в DataFrame и далее проведем анализ данных. Region, aiylokmot, useudobr удалим из нашей базы

данных. Проведем следующие операции над данными для ее дальнейшего применения.Исследуется задачу прогнозирования на основе дерево решений, которая считается популярной в последнее время [9].Приведем теперь пример прогнозирования с применением дерево решений на примере базы данных по оттоку фермеров по выращиванию ячменя, картошки и сено. В качестве целевой переменной возьмем бинарную переменную churn, которая будет означать отток. Введем библиотеки следующие библиотеки Python.При работе с многими базами данных они уже установлены в Jupyter [2].

```
import pandas as pd
from sklearn.tree import DecisionTreeClassifier
```

На основе больших данных, собранных по Иссык-Кульским районам, предварительно анализируются для ее дальнейшего использования. Исследуем базу данных по нескольким айылокмоту Иссык-Кульской области данных 1457 фермеров по выращиванию ячменя, картофеля и заготовки сена. Для исследования будем широко использовать библиотеки Pandas. База представлен как .csvфайл data_uroj_full.csv:Для чтения данных мы используем следующую функцию. Читаем данные как .csv файл через библиотеку pandas [3].

```
In [ ]: data = pd.read_csv('data_uroj_full.csv')
```

Ниже представлено анализ данных этой базы, ее структура и данные. Отображать будем через оператор head из датафрейма data по следующему коду.

```
data.head()
```

	region	aiyl okmot	id fermer	use udobr	have technik	u yach	plosh yach	charge yach	call yach tech	u kart	plosh kart	charge kart	call kart man	call kart tech	u seno	charge seno	plosh seno
1344	Ак-Суйский	Ак-Булунский	1445	No	No	42.86	2.30	10.51	5	7.373937	2.797245	7.584771	5	5	251.047146	103.542863	1.938583
503	Тюпский	Ак-Булунский	604	No	No	27.99	2.64	6.85	4	9.565207	2.846719	11.747083	5	4	233.635638	219.391601	2.098621
1343	Ак-Суйский	Ак-Булунский	1444	Yes	Yes	42.59	2.02	11.60	3	7.551641	1.699100	7.798400	4	3	335.692275	171.143463	2.748249
1342	Ак-Суйский	Ак-Булунский	1443	Yes	Yes	32.26	2.62	9.86	4	5.151371	2.199749	7.895725	3	4	290.532677	162.930554	1.911739
1341	Ак-Суйский	Ак-Булунский	1442	No	No	34.28	1.71	12.77	5	6.151450	2.260367	10.031944	5	5	234.959034	172.397673	2.422604

База данных имеет разные типы данных, количественные, текстовые логические или категориальные. Давайте удалим неколичественные, ненужные поля базы данных. Они представляют только информативный характер. Они нам интерес не представляют. При необходимости их мы снова включим БД.

```
In [ ]: data.drop(['region', 'use udobr','aiyl okmot'], axis=1, inplace=True)
        data['have technik'] = data['have technik'].map({'Yes': 1,'No': 0})
```

Теперь отобразим общую информацию о типах данных нашей базе данных с количественными данными. Для этой цели выпишем еще один оператор pandas в следующем виде.
data.info()

Ниже представлено подробная информация о базе данных. Например, указано сколько всего признаков базы данных, показаны типы данных этих признаков. Имеется информация и о отсутствующих данных [4].

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1457 entries, 0 to 1456
Data columns (total 16 columns):
have technik       1457 non-null int64
u yach             1457 non-null float64
plosh yach         1457 non-null float64
charge yach        1457 non-null float64
call yach tech     1457 non-null int64
u kart             1457 non-null float64
plosh kart         1457 non-null float64
charge kart        1457 non-null float64
call kart man      1457 non-null int64
call kart tech     1457 non-null int64
u seno             1457 non-null float64
charge seno        1457 non-null float64
plosh seno         1457 non-null float64
call man seno      1457 non-null int64
call seno tech     1457 non-null int64
churn              1457 non-null bool
dtypes: bool(1), float64(9), int64(6)
memory usage: 172.2 KB
```

Для удобства нашу целевую переменную churn. в переводе означает отток, преобразуем в переменную в целый тип. Она первоначально имеет тип логический. Вот, код перевода.

```
In [ ]: y = data['churn'].astype('int')
```

В множество переменных включим все поля кроме churn

```
In [ ]: X = data.drop('churn', axis=1)
```

Прочитаем преобразованную базу данных, которые имеют только количественные записи. Используя Pandasотображаем данные нашей базы данных.

X.head()

	id fermer	have technik	u yach	plosh yach	charge yach	call yach tech	u kart	plosh kart	charge kart	call kart man	call kart tech	u seno	charge seno	plosh seno	call man seno	call seno tech
0	101	0	32.12	2.89	12.82	5	10.945878	2.179680	11.781597	5	5	268.504055	161.571124	2.220578	5	5
1	102	1	34.82	2.38	5.49	3	5.193232	2.222455	7.914565	6	3	314.300068	103.275246	2.034102	6	3
2	103	1	25.66	2.41	11.37	2	11.849800	1.834595	11.254027	4	2	181.898116	106.051818	2.548502	4	2
3	104	0	41.71	2.49	9.70	4	10.070142	2.533347	10.449367	3	4	200.899484	149.393145	1.882755	3	4
4	105	1	27.19	2.60	8.46	3	8.311040	2.702282	7.624980	5	3	339.289381	215.318173	1.908365	5	3

Ниже представлен количество записей и признаков изучаемой базыданных.

X.shape, y.shape

((1457, 16), (1457,))

Теперь настроим параметры дерева решений на процесс разбиения данных. Настроим параметры модели следующим образом.Для этого обучим модель применяя технологию разделения данных на мелкие подмножества. С применением библиотеки sklearnвыпишем следующий код [5].

```
In [ ]: cross_val_score
from sklearn.model_selection import train_test_split, cross_val_score
import numpy as np
X_train, X_valid, y_train, y_valid = train_test_split(X, y,
                                                      test_size=0.3,
                                                      random_state=17)
X_train.shape, X_valid.shape
((1019, 16), (438, 16))
```

Сделаем следующие разбиение 70% множества (X_train, y_train) будет для обучения и 30% на тестирования модели (X_holdout, y_holdout). Причем везде множество которую мы отложили потом запустим только для проверки модели.Эти данные в процесс обучения не допускаются, они будут только тестировать процесс после обучения модели. давайте отдельно рассмотрим и обучим 2 модели, это наш дерево решений и метод ближайших соседей [8].

В процесс дерево запускаем деревья с количеством 17 деревьев 0.8695112527769728

```
In [ ]: first_tree = DecisionTreeClassifier(random_state=17)
np.mean(cross_val_score(first_tree, X_train, y_train, cv=5))
```

Для настройки соседей пишем следующий код, которая использует общий алгоритм ближайших соседей.

```
In [ ]: from sklearn.neighbors import KNeighborsClassifier
first_knn = KNeighborsClassifier()
np.mean(cross_val_score(first_knn, X_train, y_train, cv=5))
```

0.7536994107988023

Для настройки глубины дерева используем параметрmax_depth.

Обучение ближайших соседей. Модели.

Качество прогнозов будем проверять с помощью простой метрики – доли правильных ответов. Сделаем прогнозы для отложенной выборки. Дерево решений справилось лучше: доля правильных ответов около 94% против 88% у kNN. Но это мы пока выбирали параметры наугад. Код для оценки моделей следующая.

0.8561643835616438

array([[323, 14],

[49, 52]], dtype=int64)

np.bincount(y_valid)

array([337, 101], dtype=int64)

Построим графическое изображение дерева. вызываем библиотеку 0.7876712328767124

Сначала изучим дерево решений для глубины 2 с параметром max_depth=2

Вот построенный модель

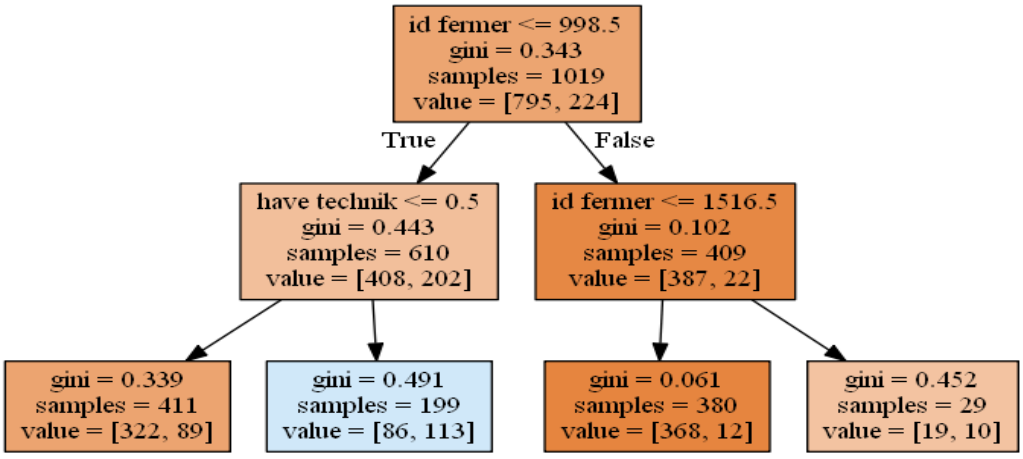


Рис.2. Построения дерево решений для глубины 2 с параметром max_depth=2

Для глубины 3 имеем следующий модель дерево решений. Вот код реализации дерево.
0.860730593607306

Таким образом ошибка составляет 86%, что неплохо.

Построим теперь графическое представление модели для данной глубины дерева. Имеем модель следующего вида.

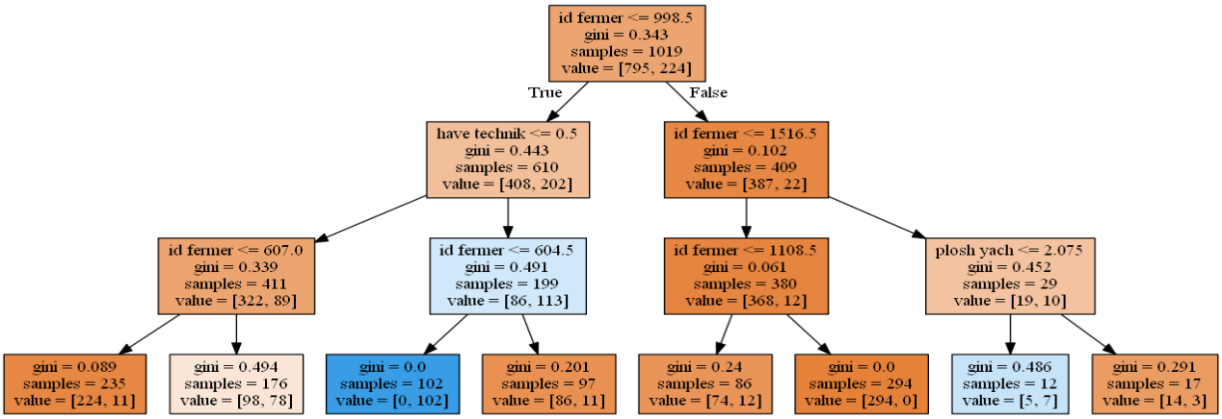


Рис.3. Построения дерево решений для глубины 2 с параметром max_depth=3

0.865296803652968

В данном случае точность модели превышает не много.Вот код визуализации.

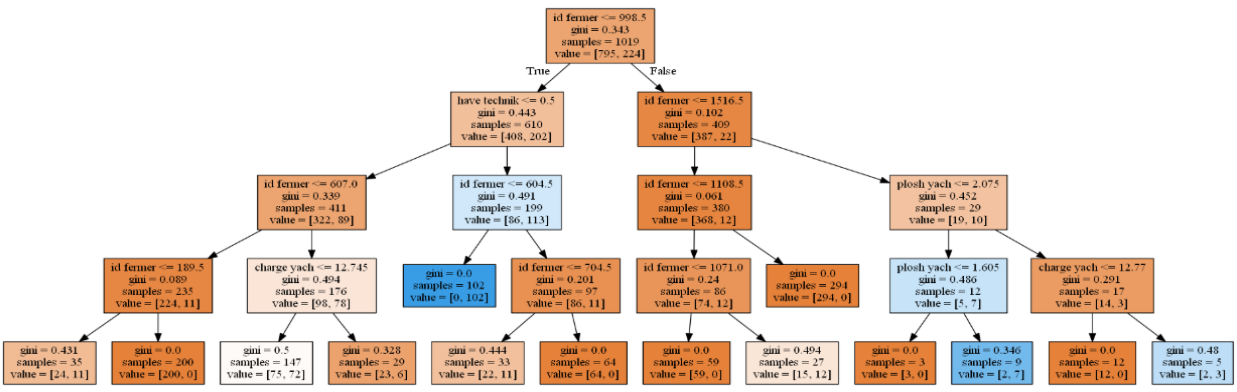


Рис.4. Построения дерево решений для глубины 2 с параметром max_depth=4

Углубляемся дальше возьмем дерево глубины уже 5.Вот результат дерево решений для глубины дерево max_depth=5
0.8767123287671232

Процент точности заметно улучшается. И составил почти 87%.Вот графическое представление модели.

```
export_graphviz(second_tree, out_file='tree_work_full_05.dot',
                feature_names=X.columns, filled=True,)
```

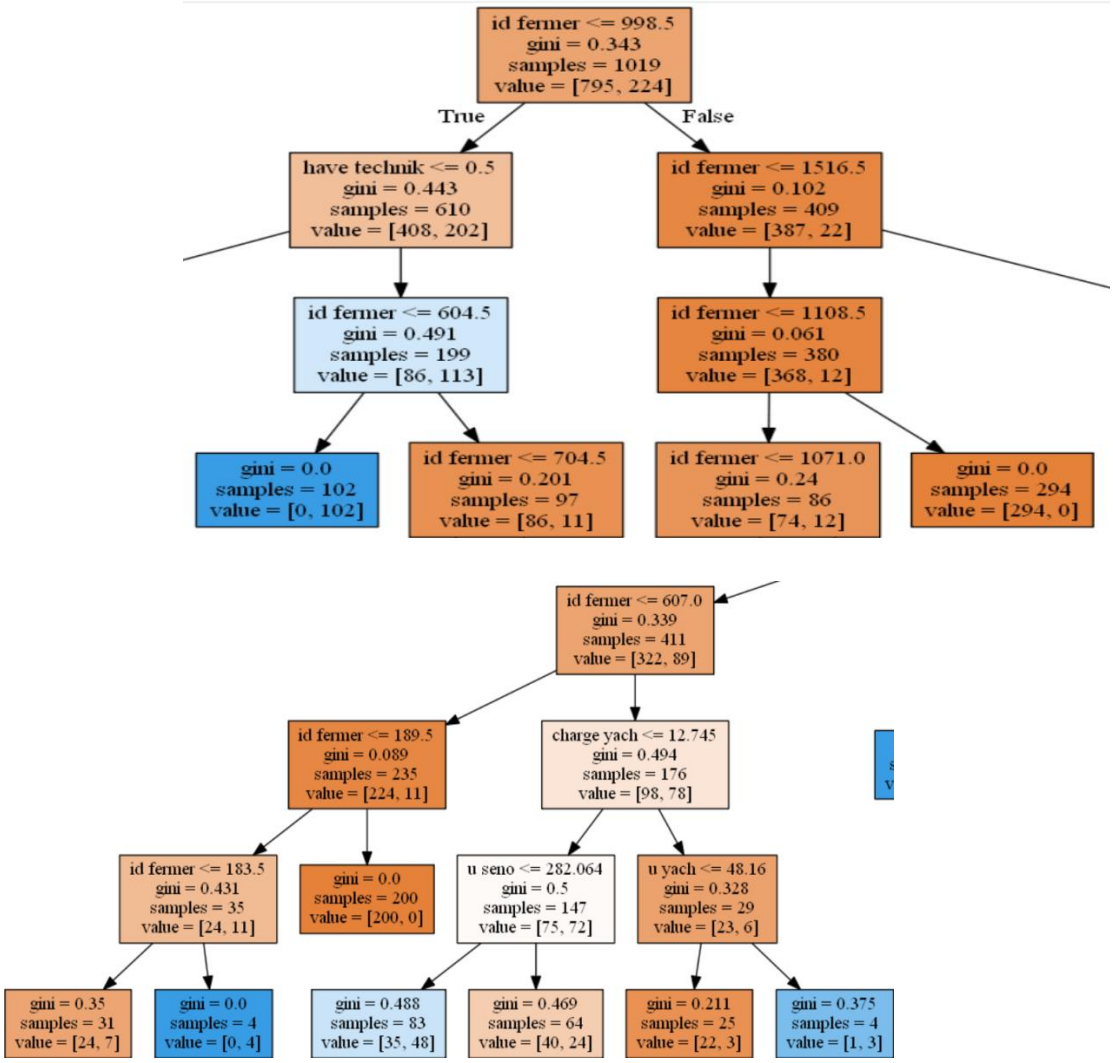


Рис.5. Построения дерево решений для глубины 2 с параметром max_depth=5

Для решения нашей проблемы построим предварительно дисперсию с шумом вокруг точного значения функции, которую мы выписали в виде математической формулы выше. Обычно в данном случае для ошибки в данных используют функцию нормального распределения [7].

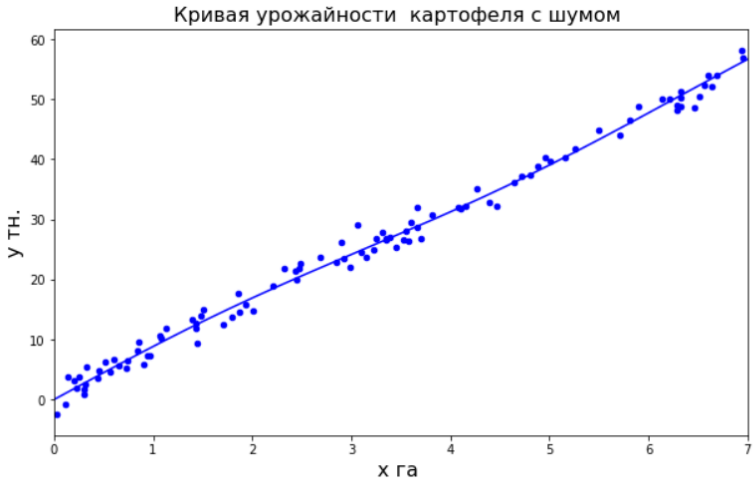


Рис.6. График разброса-данных с шумом вокруг точного значения

Теперь давайте выведем нашу кривую распределения урожайности картофеля для отдельно взятой фермерского хозяйства. Вообще применение дерево решений к нашему целевому переменному отток, будет выглядит следующей структурой-ступенчатая функция и прямая.Использование к данной задаче метода дерево решений будет иметь вид следующего графика [6]. Вот код представления дерево решений.

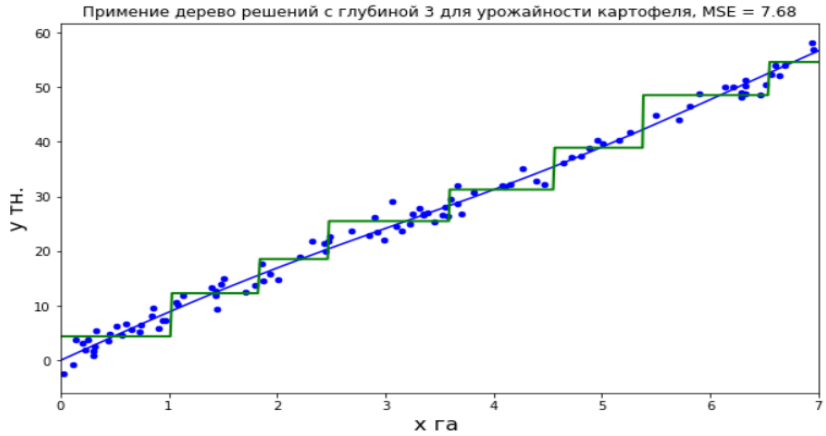


Рис.7. Применение дерево решений с глубиной равной 3 для задачи урожайности.

Прогнозная кривая урожайности картофеля с применением дерево решений глубины5, в данном случае мы просто увеличиваем количество деревьев.

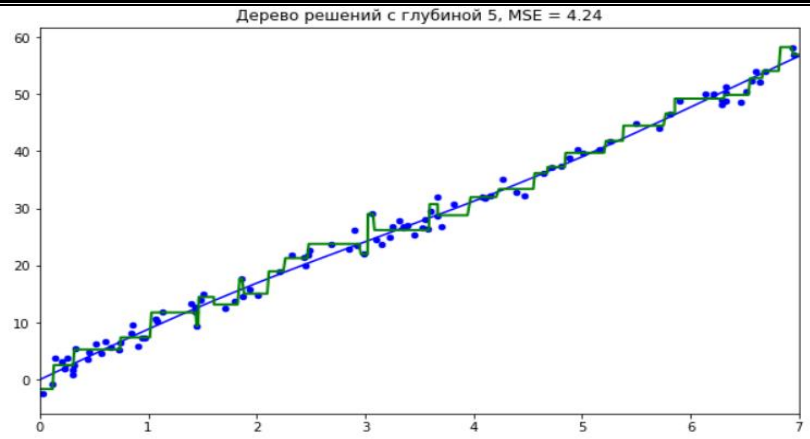


Рис.8. Применение дерево решений с глубиной равной 5 для задачи урожайности.

Представим прогнозная кривая урожайности картофеля с применением дерево решений глубины 5. Здесь явно просматривается, и мы видим, что точность вырастает заметно.

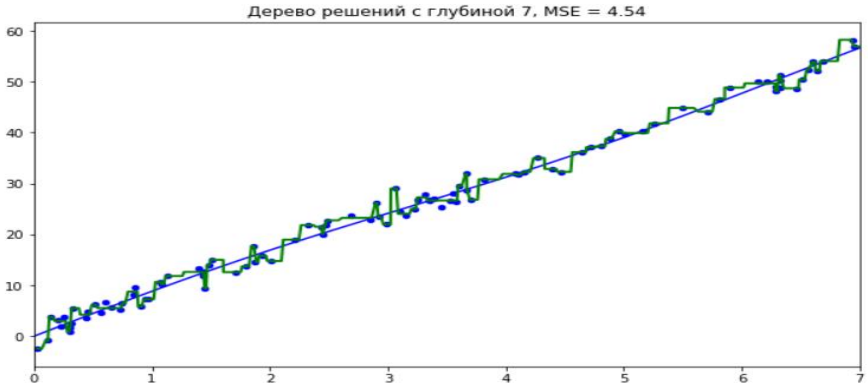


Рис.9. Применение дерево решений с глубиной равной 7для задачи урожайности.

Точность модели в данном случае почти вырос до 99%. Прогнозная кривая урожайности картофеля с применением дерево решений глубины 7 рассмотренная деревом решений.

Рассмотрим другую сложную нелинейную функцию, которая описывает урожайность кукурузы в зависимости от площади засева. Анализы построения нелинейной модели т.е. зависимость урожайности от площади засева, проведен с данными от 150 фермерских по урожайности хозяйств Иссык-Кульской.

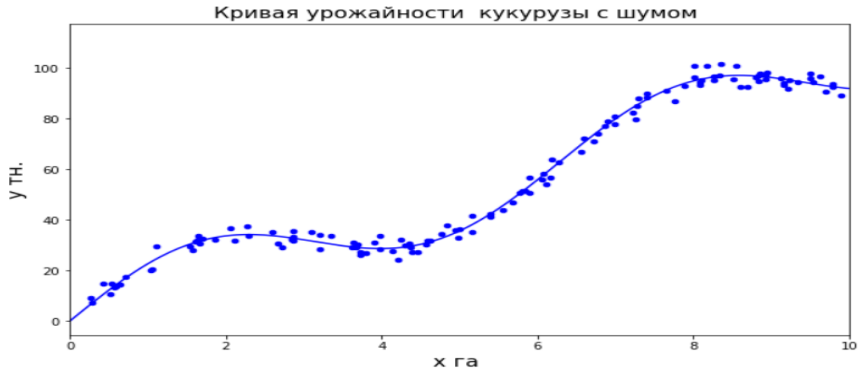
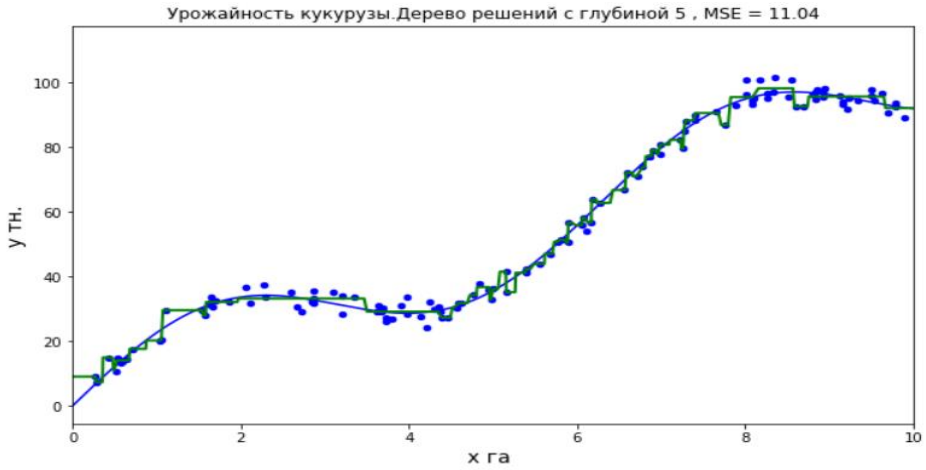


Рис.10. Построения нелинейной модели урожайности кукурузы с сложной функцией регрессии.

Таким образом в данном случае мы с участием шума с небольшим разбросом вокруг точного значения получили данные.

Вот другой пример, где в задаче регрессиидля дереварешений кривая урожайности картофеля с применением глубины дерево равное 5, имеет следующий код реализации и имеет следующий код.



.

Рис.11. Применение дерево решений к регрессионной задаче урожайности со сложной структурой данных.

В данном случае мы увеличили число деревьев и глубины дерева, для получения более точных значений модели, которые необходимы при тестировании модели.

Ниже представлен прогнозная кривая урожайности картофеля с применением дерево решений глубины 5.

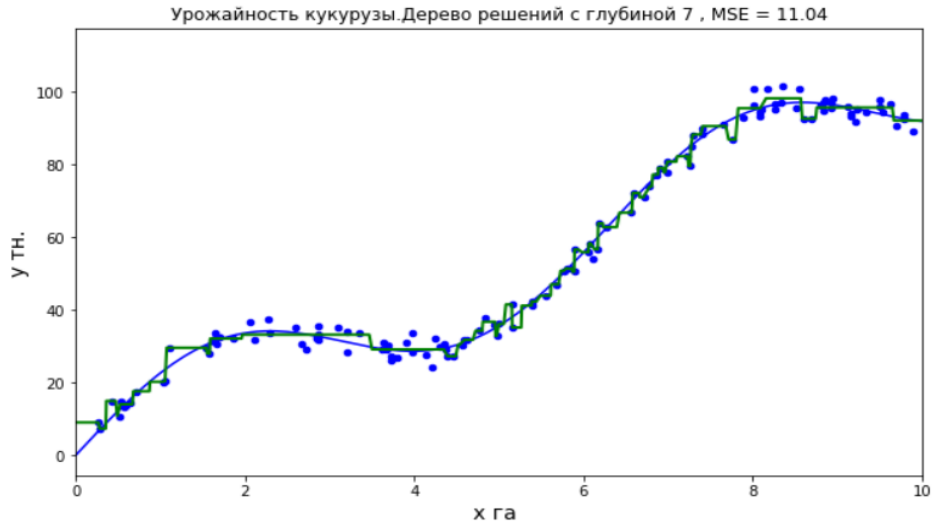


Рис.12. Применение дерево решений с большой глубиной

В данном случае мы видим, что точность очень высокая, даже в точках регрессии где сохраняется постоянная дисперсия. Прогнозная кривая урожайности картофеля с применением дерево решений глубины 7 дает максимально удовлетворительной решение для построения моделей со сложными исходными данными.

Один небольшой недостаток, на графике видно, что дерево решений аппроксимирует зависимость-целевую функцию прогнозирования в данных кусочно-постоянной функцией. Однако в этом случае сохраняется точность модели.

В данной статье подробно исследовано различные методы анализа и прогнозирования задач сельского хозяйства методами машинного обучения. Использовано дерево решений и метод ближайших соседей для широкого круга задач АПК. Данная технология решающих деревьев, хорошо описывают процесс принятия решения многих прикладных задач. Например, в весенний период перед фермером стоить задача, что сеять чтобы получить хороший прибыль. Целевым признаком в данной задаче является прибыль от полученного урожая в зависимости от площади засева, аренда сельхозтехники и использование наемного труда. Отток связан и от продажи своей продукции в осенний период по минимальной цене.

Список использованной литературы:

1. Хенрик Бринк, Джозеф Ричардс, Марк Феверолф. «Машинное обучение», 2017г.
2. Бастиан Шарден, Лука Массарон, Альберто Боскетти. «Крупномасштабное машинное обучение вместе с Python»,2017г.
3. Себастьян Рашка. «Python и машинное обучение», 2018 г.
4. Георгий Кухарев, Екатерина Каменская, Юрий Матвеев, Надежда Щеголева. «Методы обработки и распознавания изображений лиц в задачах биометрии».2018 г.
5. Петер Флах. «Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных». 2019г.
6. Андреас Мюллер, Сара Гвидо. «Введение в машинное обучение с помощью Python»,2018 г.
7. Петер Флах. «Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных», 2019 г.
8. Сабитов Б.Р., Сейтказиева Н.С., Кашкароева А.А.
9. «Бинарная задача классификации болезни растений с применением технологий глубокого обучения»
10.Наука, новые технологии и инновации Кыргызстана. 2022. № 7. С. 30-39.
11.Сабитов Б.Р., Сейтказиева Н.С., Кубанычбекова А.К. «Технологии интеграции машинного обучения с веб приложениями», Современные проблемы механики. 2022. № 47 (1). С. 38-46.

УДК: 338.483.1
DOI 10.33514/ВК-1694-7711-2024-2(1)-411-416

Султанов Ж. М., Базарбаева Г. Б., Жунусалиева А. К.
Ош мамлекеттик университети, ага окутуучу,
Ош мамлекеттик университети, окутуучу,
М.Адышев атындагы Ош технологиялык университети, окутуучу
Султанов Дж. М., Базарбаева Г. Б., Джунусалиева А. К.
Ошский государственный университет, преподаватель,
Ошский государственный университет, преподаватель,
Ошский технологический университет им. М. Адышева, преподаватель
Sultanov J. M., Bazarbaeva G. B., Dzhunusalieva A. K.
Osh State University, teacher,
Osh State University, teacher,
Osh Technological University named after. M. Adysheva, teacher

КЫРГЫЗСТАНДЫН ТҮШТҮК АЙМАГЫНЫНДАГЫ ТУРИСТТИК-РЕКРЕАЦИЯЛЫК РЕСУРСТАРДЫН ӨНҮГҮҮ ӨЗГӨЧӨЛҮКТӨРҮ
ОСОБЕННОСТИ РАЗВИТИЯ ТУРИСТСКО-РЕКРЕАЦИОННЫХ РЕСУРСОВ ЮЖНОГО РЕГИОНА КЫРГЫЗСТАНА