Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное учреждение

высшего образования

«Уральский федеральный университет

имени первого Президента России Б.Н. Ельцина»

Институт радиоэлектроники и информационных технологий – ИРИТ-РтФ

Кафедра информационных технологий и систем управления

Заказчик проекта

Улитко В.А.

**ОТЧЕТ ПО ПРОЕКТНОМУ ПРАКТИКУМУ**

**«Интеллектуальное материаловедение. Определение версии численной модели по фазовой диаграмме»**

Куратор

проекта Кошелев А.А.

Исполнители проекта:

Студент группы РИМ-140903 Еремин Ю. Д.  
Студент группы РИМ-140901 Роговой М. С.

Студент группы РИМ-140901 Рязанов В. В.  
Студент группы РИМ-140903 Хуснуллин В. Р.

Екатеринбург

2025

**СОДЕРЖАНИЕ**

[ВВЕДЕНИЕ 3](#_Toc188222029)

[ГЛАВА 1 МОДЕЛЬ ИЗИНГА 5](#_Toc188222030)

[ГЛАВА 2 ПОСТАНОВКА ЗАДАЧ И ИХ ИНТЕРПРЕТАЦИЯ 9](#_Toc188222031)

[ГЛАВА 3 МОДЕЛЬ МАШИННОГО ОБУЧЕНИЯ 10](#_Toc188222032)

[3.1 Входные данные 10](#_Toc188222033)

[3.2 Слепые данные 11](#_Toc188222034)

[3.3 Анализ качества моделей 12](#_Toc188222035)

[ГЛАВА 4 МОДЕЛЬ МАШИННОГО ОБУЧЕНИЯ ПО ВЕРСИЯМ 0–1 17](#_Toc188222036)

[4.1 Метрики для оценки качества модели 17](#_Toc188222037)

[4.2 Построение ROC-кривой и вычисление площади 18](#_Toc188222038)

[4.3 Результаты перекрестной проверки 19](#_Toc188222039)

[4.4 Важность признаков 20](#_Toc188222040)

[4.5 Построение дерева решений 22](#_Toc188222041)

[Выводы по главе 4 24](#_Toc188222042)

[ГЛАВА 5 МОДЕЛЬ МАШИННОГО ОБУЧЕНИЯ ПО ВЕРСИЯМ 0–2 26](#_Toc188222043)

[5.1 Значение ROC AUC и матрица ошибок 26](#_Toc188222044)

[5.2 Результаты перекрестной проверки (cross-validation) 27](#_Toc188222045)

[5.3 Важность признаков 28](#_Toc188222046)

[5.4 Дерево решений 30](#_Toc188222047)

[Выводы по главе 5 31](#_Toc188222048)

[ГЛАВА 6 АНАЛИЗ ЭФФЕКТИВНОСТИ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ПО МИНИМАЛЬНОМУ НАБОРУ ПАРАМЕТРОВ 33](#_Toc188222049)

[ГЛАВА 7 РАЗРАБОТКА API ДЛЯ РАБОТЫ С МОДЕЛЬЮ МАШИННОГО ОБУЧЕНИЯ 38](#_Toc188222050)

[ЗАКЛЮЧЕНИЕ 46](#_Toc188222051)

### ВВЕДЕНИЕ

Высокотемпературные сверхпроводники (ВТСП) — это уникальные материалы, способные проводить электрический ток без сопротивления при температурах, значительно превышающих температуры перехода в сверхпроводящее состояние у классических сверхпроводников. В отличие от традиционных сверхпроводников, которые требуют охлаждения до температур, близких к абсолютному нулю (минус 273,15 °C), ВТСП демонстрируют сверхпроводимость при температурах выше температуры кипения жидкого азота (минус 196 °C). Это делает их перспективными для практического применения в энергетике, транспорте, медицине и других областях. Изучение ВТСП имеет огромное значение для развития технологий, так как их внедрение может революционизировать системы передачи и хранения энергии, сделать медицинскую диагностику более доступной, а транспорт – более эффективным.

Одной из ключевых особенностей ВТСП является их сложная кристаллическая структура, которая играет важную роль в возникновении сверхпроводимости. Например, в купратных ВТСП (материалах на основе оксидов меди) наблюдается слоистая структура, где слои меди и кислорода чередуются с другими элементами. Такая структура создаёт условия для сильных электронных корреляций и спиновых флуктуаций, которые, как предполагается, ответственны за высокотемпературную сверхпроводимость. Однако механизм сверхпроводимости в ВТСП до конца не изучен, что делает теоретическое моделирование их свойств крайне важным.

Для понимания физических процессов в ВТСП можно использовать упрощённые модели, такие как модель Изинга, которая описывает взаимодействие между спинами в системе. Модель Изинга, несмотря на свою простоту, позволяет изучать коллективные явления, фазовые переходы и критическое поведение, что делает её полезным инструментом для анализа сложных систем, включая ВТСП. В частности, с помощью модели Изинга можно исследовать магнитные свойства ВТСП, такие как антиферромагнитное упорядочивание, которое часто наблюдается в купратах при температурах выше критической. Это позволяет лучше понять, как взаимодействие между спинами и зарядами влияет на сверхпроводимость.

Таким образом, изучение ВТСП с использованием моделей типа Изинга открывает новые возможности для теоретического анализа их свойств и поиска материалов с ещё более высокими критическими температурами. Это может привести к созданию сверхпроводников, работающих при комнатной температуре, что станет прорывом в науке и технологиях. В данном отчёте мы рассмотрим особенности кристаллической структуры ВТСП, их связь с магнитными свойствами и возможность использования модели Изинга для теоретического описания этих материалов.

# ГЛАВА 1 МОДЕЛЬ ИЗИНГА

В качестве физической модели данной работы рассматривается модель Изинга, упомянутая выше, на квадратной решетке (Рисунок 1.1). Данная модель представляет собой математическую модель, широко используемую в статистической физике для описания намагничивания материалов. Она основана на концепции дискретных спинов, которые могут принимать значения или , расположенных на узлах регулярной решётки. Энергия системы описывается гамильтонианом по формуле 1.1:

|  |  |
| --- | --- |
|  | (1.1) |

где:

 — полная энергия системы,

 — константа взаимодействия между ближайшими соседними спинами,

​ — константа взаимодействия между вторыми соседними спинами,

 и  — значения спинов на узлах  и , которые могут принимать значения  или ,

 — внешнее магнитное поле,

— обозначает сумму по всем парам ближайших соседей на решётке,

— сумма по всем узлам решётки.

Ниже представлена «случайная» квадратная решетка размером 10×20, где каждый узел представляет собой спин в модели Изинга.

\begin{figure}
 \includegraphics{ising}\end{figure}

Рисунок 1.1 – Пример случайного расположения спинов на квадратной решетке

Ближайшие соседи — это спины, которые находятся непосредственно рядом с данным спином по горизонтали и вертикали. Для каждого спина на квадратной решётке есть 4 ближайших соседа (если спин не находится на границе решётки).

Вторые соседи — это спины, которые находятся на расстоянии одного узла по диагонали или через один узел по горизонтали/вертикали. Для каждого спина на квадратной решётке есть 8 вторых соседей (если спин не находится на границе решётки).

В данной работе будут рассматриваться 3 версии модели Изинга:

1. Версия **"**0**"** – двумерная модель Изинга в отсутствие внешнего магнитного поля. В этой версии учитывается только взаимодействие между ближайшими соседними спинами, а остальными членами гамильтониана пренебрегают. Таким образом, гамильтониан для версии "0" записывается в соответствии с уравнением 1.2:

|  |  |
| --- | --- |
|  | (1.2) |

1. Версия **"**1**"** – двумерная модель Изинга с учётом внешнего магнитного поля. В этой версии, помимо взаимодействия между ближайшими соседними спинами, добавляется член уравнения, описывающий влияние внешнего магнитного поля на систему. Гамильтониан для версии "1" записывается соответствии с уравнением 1.3:

|  |  |
| --- | --- |
|  | (1.3) |

1. Версия **"**2**"** — двумерная модель Изинга с взаимодействием первых и вторых соседей и внешним магнитным полем. То есть формула для расчета численной модели в точности совпадает с формулой (1.1).

Таким образом, глобальная цель – поиск высокотемпературных сверхпроводников с уникальными свойствами при температуре, выше существующих на сегодняшний день, в данной работе, ограничилась в нахождении математической модели Изинга, в которой спиновое взаимодействие в системе привело бы к искомым сверхпроводящим свойствам.

При этом количественное определение спиновых взаимодействий экспериментально напрямую не представляется возможным. Однако косвенно, через привычные параметры физических систем, такие как энергия , теплоемкость , намагниченность , магнитная восприимчивость , антиферромагнитная часть намагниченности и антиферромагнитная часть магнитной восприимчивости , можно установить связь между спиновыми взаимодействиями и наблюдаемыми термодинамическими свойствами.

В данной работе используется модель Изинга как инструмент для интерпретации экспериментальных и расчётных данных. На основе параметров системы (), с помощью моделей машинного обучения должна определяться версия модели Изинга, которая наилучшим образом будет описывать наблюдаемые свойства. Это позволит не только классифицировать систему, но и предсказывать параметры спиновых взаимодействий, которые могут быть ключевыми для возникновения сверхпроводимости.

Таким образом, модель Изинга служит мостом между наблюдаемыми термодинамическими параметрами и скрытыми спиновыми взаимодействиями, что открывает новые возможности для поиска материалов с уникальными свойствами, включая высокотемпературные сверхпроводники.

# ГЛАВА 2 ПОСТАНОВКА ЗАДАЧ И ИХ ИНТЕРПРЕТАЦИЯ

Модель Изинга, несмотря на свою простоту, является мощным инструментом для описания сложных физических систем, таких как магнетики и высокотемпературные сверхпроводники. Однако анализ таких систем часто требует обработки большого объёма данных, включая термодинамические параметры (энергия, теплоёмкость, намагниченность, магнитная восприимчивость и другие). В этой связи возникает задача автоматизации анализа данных и определения версии модели Изинга, которая наилучшим образом описывает наблюдаемые свойства системы.

В данной работе, заказчиком были поставлены следующие цели:

1. Научиться определять версию численной модели с помощью методов машинного обучения по расчётным данным;
2. После обучения модели машинного обучения по всем признакам, определить минимально достаточный набор расчётных данных для определения версии численной модели;
3. Попробовать определить версию по экспериментальным данным для изинговских магнетиков по параметрам .

Для достижения поставленной цели были сформулированы задачи:

1. Выбор модели машинного обучения и успешного её обучения для различных версий модели Изинга (версии "0", "1" и "2").
2. Получение и анализ различных метрик для каждой из моделей (0–1 и 0–2), а также оценка важностей признаков.
3. Определить минимальные набор признаков, необходимого для классификации версии модели Изинга.
4. Создание сервиса для классификации численной версии модели по параметрам физической системы, загружаемых пользователем.

Таким образом, работа направлена на создание инструмента для анализа и классификации различных версий модели Изинга, а также для автоматического определения версии модели на основе входных данных.

# ГЛАВА 3 МОДЕЛЬ МАШИННОГО ОБУЧЕНИЯ

В данной работе перед моделью машинного обучения (ММО) стоит задача классификации численных версий модели Изинга. Однако для того, чтобы понять, какую именно ММО необходимо применить для выполнения поставленных целей, следует оценить следующие аспекты:

1. Входные/выходные данные.
2. Сложные нелинейные зависимости между признаками.
3. Устойчивость к переобучению.
4. Работа с несбалансированными и разнородными данными.
5. Гибкость и настройка гиперпараметров.
6. Интерпретируемость результатов.

## 3.1 Входные данные

Для обучения модели машинного обучения, заказчиком были отправлены .csv файлы, которые имели следующую структуру, показанную на рисунке 3.1.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| L | h | T | E | C | FM | Xfm | AFM | XAFM |

Рисунок 3.1 – Входные данные для обучения и тестирования ММО.

Здесь: «L»  – размерность системы (в данной работе L = const = 32 атома), «h» – версия численной модели, «T»  – температура системы,   
«E»  – энергия системы, «C»  – теплоемкость системы, «FM»  – ферромагнитная часть намагниченности, «Xfm»  – ферромагнитная часть магнитной восприимчивости, «AFM»  – антиферромагнитная часть намагниченности, «XAFM»  – антиферромагнитная часть магнитной восприимчивости

Таким образом целевой переменной является h, которая определяет версию 0 или 1 для модели 0–1, либо 0 или 2 для модели 0–2.

Признаками, на основе которых модель делает предсказания, являются параметры системы: «E», «С», «FM», «Xfm», «AFM», «XAFM».

## 3.2 Слепые данные

Для определения версии численной модели физической системы, заказчиком были отправлены .csv файлы, которые имели такие же признаки, как и по рисунку 3.1, однако структура отличается:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| L | sample | T | E | C | FM | Xfm | AFM | XAFM |

Рисунок 3.2 – Слепые данные для предсказания версии численной модели

Согласно рисунку 3.2 уже нет целевого признака , так как задача ММО добавить колонку и написать для каждой строчки «Predicted\_Target». Однако появилась колонка «sample» – которая группирует строчки одного целевого признака.

То есть для каждого существует n-ое количество строчек, которые на самом деле имеют один и тот же ключ «h». Поэтому необходимо добавить ещё 2 колонки для вывода результатов: «Predict\_Proba» – метод, отвечающий за вероятность принадлежности объекта к каждому классу, а также «Predicted\_Target\_For\_Sample» – конечная версии численной модели для всего семпла, рассчитанная путем подсчета большего количества предсказанных строк, то есть если в большинство строчек имеют «Predicted\_Target» равен 1, то «Predicted\_Target\_For\_Sample» равен 1, иначе равен 0.

Структура вывода результатов для слепых данных показана на рисунке 3.3.

|  |  |  |  |
| --- | --- | --- | --- |
| sample | Predicted\_Target | Predict\_Proba | Predicted\_Target\_For\_Sample |

Рисунок 3.3 – Вывод результатов определения версии численной модели по строкам и по семплам для слепых данных

## 3.3 Анализ качества моделей

Для проведения анализа производительности различных моделей машинного обучения был использован обширный набор методов, включая логистическую регрессию, методы опорных векторов (SVM), случайные леса (Random Forest), градиентный бустинг (Gradient Boosting), XGBoost, нейронные сети (MLP) и другие алгоритмы. Каждая модель была протестирована на задаче бинарной классификации с использованием пяти ключевых метрик: Accuracy, Precision, Recall, F1 и ROC-AUC. Это позволило получить полное представление о сильных и слабых сторонах каждого метода. Анализ эффективности моделей по метрикам Accuracy, Precision, Recall, F1 и ROC-AUC показаны на рисунках 3.4–3.8.

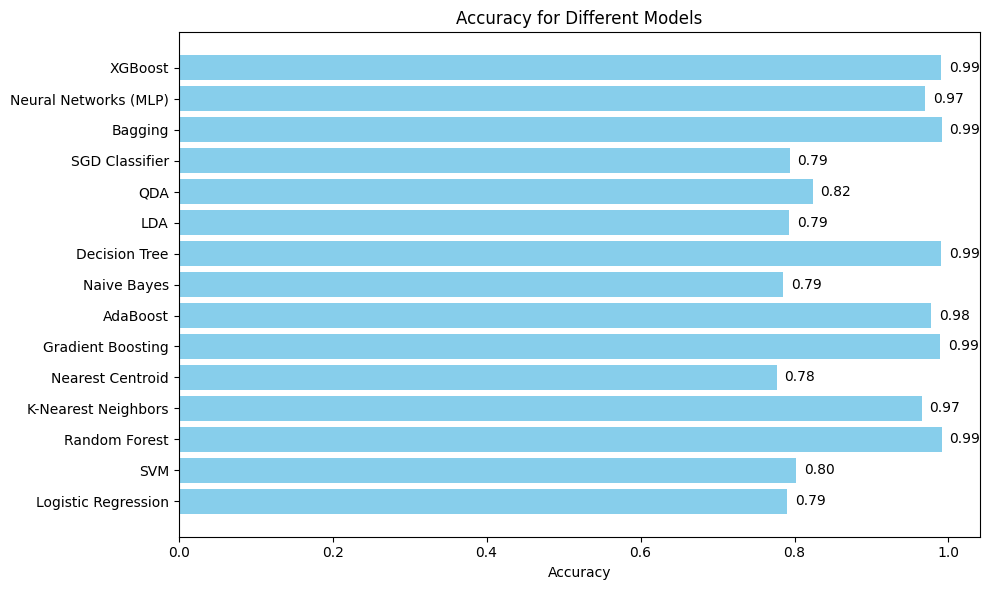


Рисунок 3.4 – Результат анализа эффективности моделей метрикой Accuracy

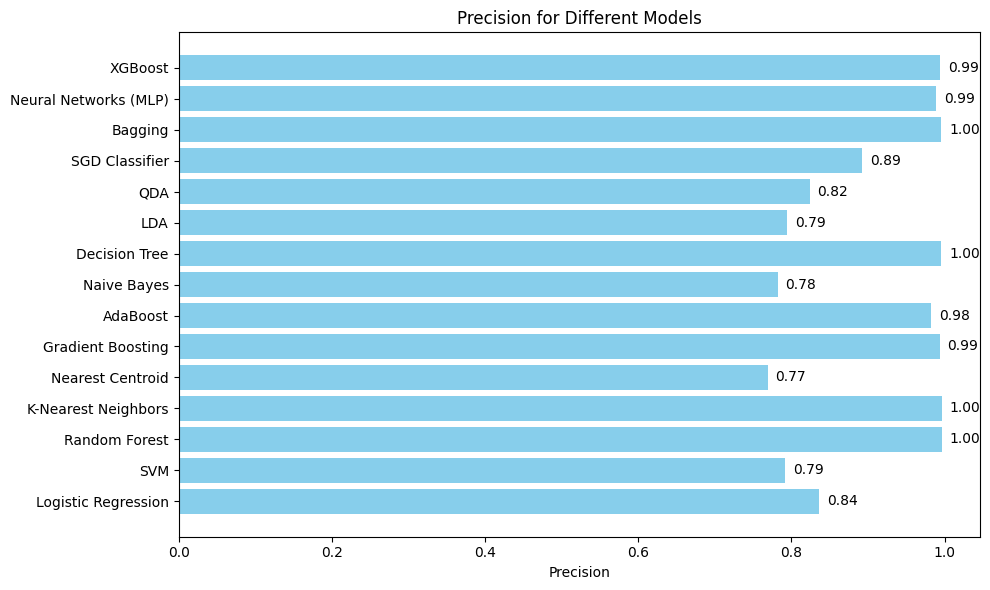


Рисунок 3.5 – Результат анализа эффективности моделей метрикой Precision

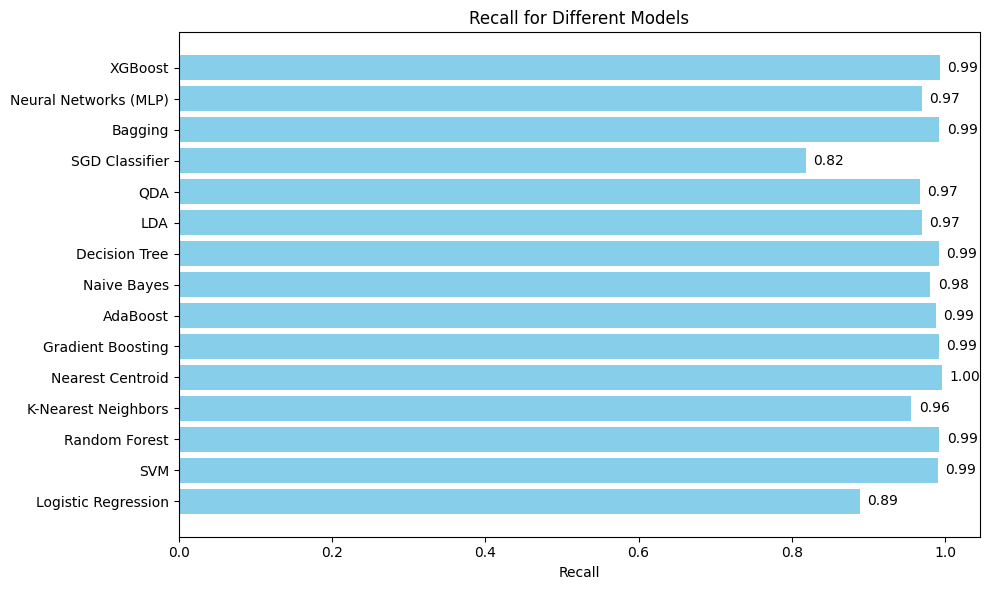


Рисунок 3.6 – Результат анализа эффективности моделей метрикой Recall

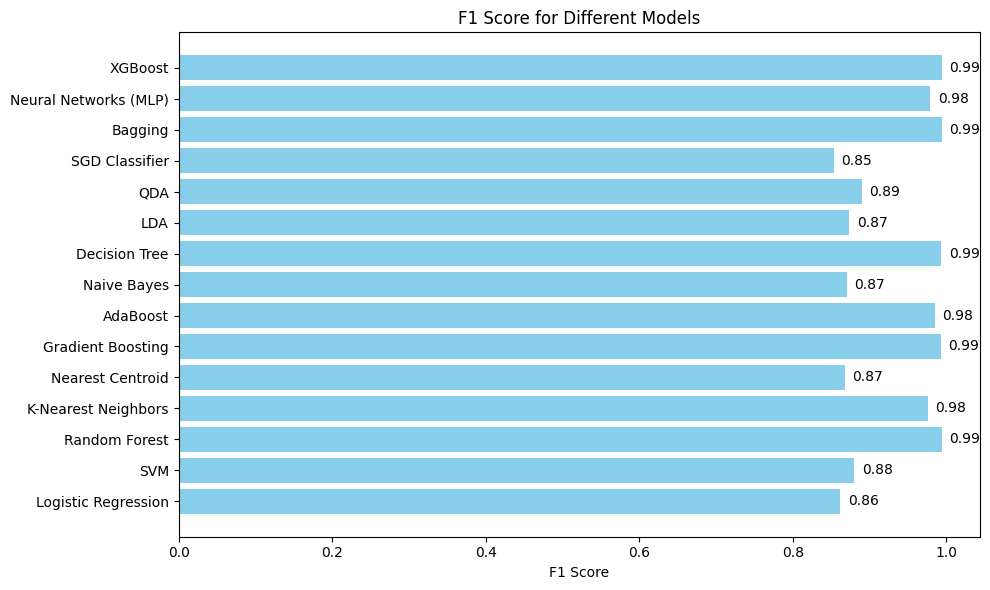


Рисунок 3.7 – Результат анализа эффективности моделей метрикой F1

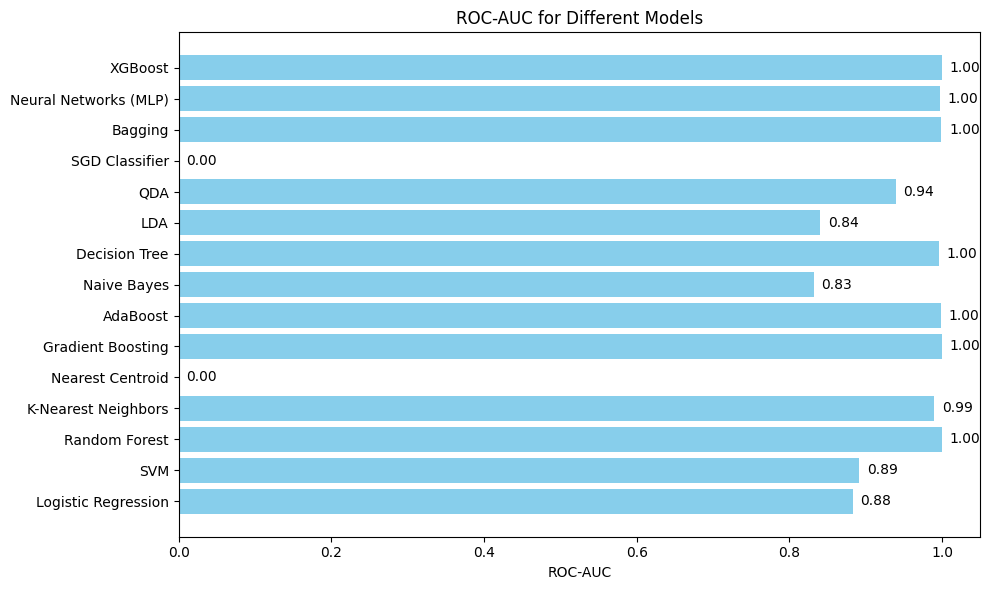


Рисунок 3.8 – Результат анализа эффективности моделей метрикой ROC-AUC

При анализе учитывались следующие факторы:

1. Распределение классов: Перед началом экспериментов была проведена проверка распределения классов в обучающей и тестовой выборках. Этот этап важен для выявления возможного дисбаланса классов, который мог бы повлиять на результаты тестирования. При необходимости могли быть применены такие методы, как взвешивание классов или ресэмплинг, однако в данной работе дисбаланс не был критичным.
2. Метрики оценки: Использование нескольких метрик помогло учесть разные аспекты производительности моделей. Например, высокая точность (Accuracy) некоторых моделей могла быть вызвана доминирующим классом, однако метрики F1-score и ROC-AUC позволяли оценить более сбалансированное качество классификации.
3. Особенности работы моделей: Некоторые алгоритмы, например, SGD Classifier и Nearest Centroid, не поддерживают расчет вероятностей (predict\_proba), из-за чего не удалось рассчитать ROC-AUC. Это могло внести определенные ограничения в сравнительный анализ, однако итоговый выбор был сделан с учетом остальных метрик.
4. Настройка гиперпараметров: Для большинства моделей использовались стандартные параметры, за исключением XGBoost, где была задана метрика logloss для оценки на этапе обучения, что позволило адаптировать процесс оптимизации под задачу бинарной классификации, и нейронных сетей (MLP), для которых были скорректированы настройки для повышения качества. Этот аспект подчеркивает потенциальные возможности для дальнейшей оптимизации других методов.

Несмотря на то, что точность многих моделей, таких как Random Forest, Gradient Boosting и Decision Tree, оказалась высокой (около 99 %), XGBoost был выбран по следующим причинам:

1. Устойчивость к переобучению: Благодаря механизмам регуляризации (L1 и L2), XGBoost демонстрирует устойчивость к переобучению, что особенно важно для работы с нелинейными зависимостями и разнородными данными.
2. Гибкость и настройка гиперпараметров: В отличие от некоторых других моделей, XGBoost предоставляет широкий спектр гиперпараметров для настройки, что позволяет адаптировать модель под особенности задачи.
3. Эффективность работы с несбалансированными данными: XGBoost использует взвешивание классов и оптимизацию по логарифмической потере, что делает его устойчивым к дисбалансу классов.
4. Скорость обучения: Благодаря оптимизации вычислений (например, параллельной обработке), XGBoost обеспечивает высокую скорость обучения даже на больших наборах данных.
5. Интерпретируемость: Использование XGBoost позволяет анализировать важность признаков, что упрощает интерпретацию результатов и получение дополнительных инсайтов.

Учитывая все вышеперечисленные аспекты, мы приняли решение использовать XGBoost в качестве основной модели для решения задачи бинарной классификации. Этот выбор обусловлен её высокой производительностью, гибкостью и надежностью в условиях поставленной задачи.

# ГЛАВА 4 МОДЕЛЬ МАШИННОГО ОБУЧЕНИЯ ПО ВЕРСИЯМ 0–1

Для разделения данных по версиям 0 и 1 была обучена модель машинного обучения на основе модели XGboost. Набор данных делили в соотношении 80 % для обучения модели и 20 % для тестирования модели. Качество обучения модели оценивали по следующим метрикам: precision, recall, F1 и ROC AUC. Для понимания работы модели машинного обучения были выявлены наиболее определяющие признаки, а также оценено их влияние на выбор класса моделью машинного обучения.

## 4.1 Метрики для оценки качества модели

Для оценки качества работы модели были использованы следующие метрики:

1. Accuracy (точность): показывает долю правильно классифицированных объектов от общего числа примеров. Высокое значение accuracy указывает на общее превосходное качество модели.
2. Precision (точность): измеряет долю истинно положительных объектов среди всех предсказанных положительных. Это важно для задач, где ложные срабатывания могут быть критичны.
3. Recall (полнота): определяет долю истинно положительных объектов, которые модель смогла обнаружить из общего числа реальных положительных примеров.
4. F1: гармоническое среднее precision и recall, полезно в задачах, где важен баланс между этими метриками.
5. ROC-AUC (Area Under the Curve): измеряет качество модели в различении классов независимо от порога классификации. Значение ROC-AUC, близкое к 1, говорит о практически идеальном разделении классов.

Эти метрики позволяют всесторонне оценить эффективность модели и понять её поведение как на тренировочных данных, так и при кросс-валидации. Далее мы рассмотрим их подробнее.

## 4.2 Построение ROC-кривой и вычисление площади

ROC-кривая и площадь под кривой (AUC) показывают способность модели различать классы. Как видно из графика ROC-кривой для нашей модели (рисунок 4.1) площадь под ней равна 0,9735, что свидетельствуют о высоком качестве разделения объектов по классам на 0 и 1, что делает модель очень надежной для решения поставленной задачи.

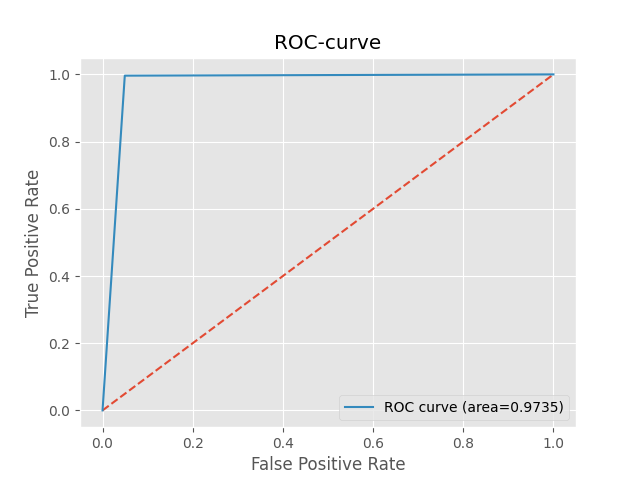


Рисунок 4.1 – ROC-кривая

## 4.3 Результаты перекрестной проверки

Перекрестная проверка (кросс-валидация) — это метод оценки производительности модели машинного обучения, который позволяет проверить, насколько хорошо модель обобщает данные. Основная идея заключается в том, чтобы разделить данные на несколько частей (фолдов), обучить модель на части данных и протестировать её на оставшейся части. Этот процесс повторяется несколько раз, чтобы каждый фолд использовался как тестовая выборка. Другими словами, кросс-валидация показывает стабильность и надежность модели.

На рисунке 4.2 изображены результаты 5-фолдовой кросс-валидации со значениями метрик Precision, Recall, F1, ROC-AUC. На графике видно, что метрики практически неизменны между фолдами, за исключением незначительных колебаний Precision и Recall в отдельных фолдах. Это может быть связано с особенностями распределения данных между фолдами.

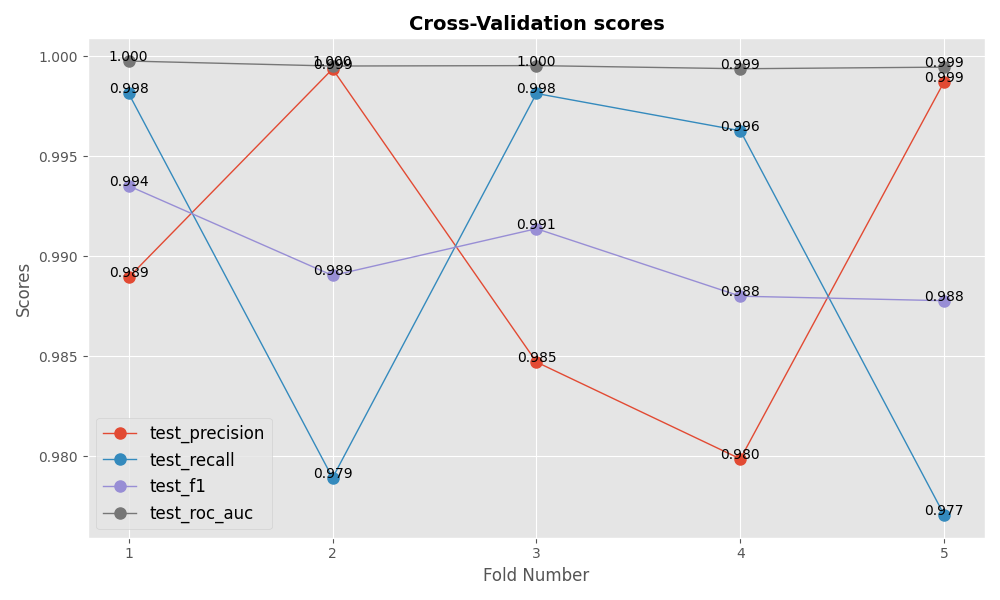


Рисунок 4.2 – Результаты перекрестной проверки модели машинного обучения для классов 0 и 2

График перекрестной проверки подтверждает устойчивость модели на различных подвыборках, модель имеет очень высокие и стабильные значения ROC-AUC по всем фолдам. Это указывает на то, что модель отлично справляется с задачей классификации и хорошо обобщает данные.

Средние значения метрик по результатам перекрестной проверки показаны в таблице 4.1.

Таблица 4.1 – Средние значения метрик

|  |  |
| --- | --- |
| **Метрика** | **Среднее значение** |
| Accuracy | 0.990 |
| Precision | 0,990 |
| Recall | 0,989 |
| F1 | 0,990 |
| ROC AUC | 0.999 |

## 4.4 Важность признаков

В ходе работы была проведена оценка признаков, которые оказывают наибольшее влияние на результат классификации. Результат оценки показан на рисунке 4.3. График SHAP (SHapley Additive exPlanations) позволяет детально проанализировать влияние каждого признака на предсказания модели (рисунок 4.4).

Изображение выглядит как текст, снимок экрана, линия, диаграмма

Автоматически созданное описание

Рисунок 4.3 – График важности признаков при классификации версий 0 и 1

Изображение выглядит как текст, снимок экрана, График, линия

Автоматически созданное описание

Рисунок 4.4 – SHAP-график важности признаков при классификации версий 0 и 2

По графику важности признаков видно, что FM, E и AFM являются наиболее значимыми для модели.

На основе анализа SHAP значений можно сделать следующие выводы:

1. Xfm и E оказывают наибольшее влияние на предсказания модели. Высокие значения этих признаков (отмечены красным цветом) чаще связаны с предсказаниями класса 0. Низкие значения AFM и Xafm (отмечены синим цветом) связаны с отрицательными предсказаниями (Class 0).

2. Смешанное влияние: Признаки FM и C демонстрируют смешанное влияние. В зависимости от их значений они могут как увеличивать, так и уменьшать вероятность предсказания Class 1.

3. Интерпретация SHAP значений: SHAP значения позволяют понять, как каждый признак влияет на предсказание модели для конкретного образца. Например, высокие значения FM увеличивают вероятность предсказания Class 1, а низкие значения AFM уменьшают её. График SHAP также подтверждает, что модель учитывает сложные взаимосвязи между признаками, что делает её более гибкой и точной.

## 4.5 Построение дерева решений

Дерево решений использует критерий Gini и Entropy для разделения данных на два класса: Class 0 и Class 1. Gini и Entropy: Оба критерия показывают, насколько хорошо разделены данные в каждом узле (рисунки 4.5 и 4.6).

Большинство образцов (всего образцов в узле 11008) принадлежат к Class 1 (8064 образца), но есть и образцы из Class 0 (2944 образца). Значение Gini = 0,392 указывает на то, что разделение данных в этом узле умеренно нечистое, то есть классы смешаны, но один из классов (Class 1) всё же доминирует. Чем ниже Gini, тем лучше разделение в узле. Если Gini близко к 0, это означает, что узел хорошо разделяет классы, и модель может принимать более уверенные решения.

Интерпретация решающего дерева подтверждает ключевые закономерности модели:

1. В качестве предикатов используются важные признаки, в том числе в корне – FM, E, AFM что соответствует анализу важности признаков.
2. Вторичные признаки, такие как C и Xfm, также участвуют в уточнении классов.

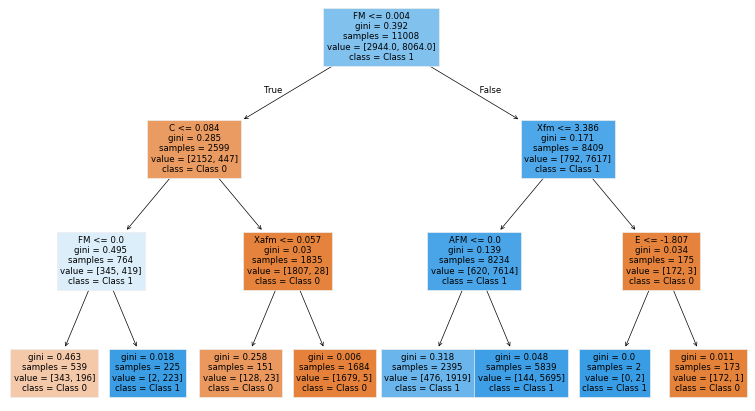


Рисунок 4.5 – Дерево решений для критерия Gini

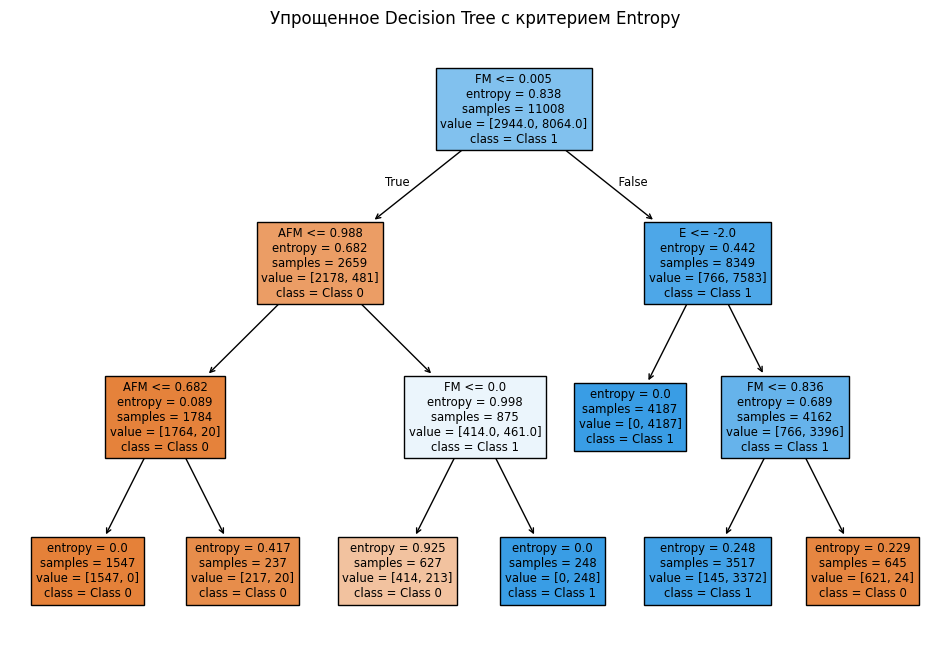


Рисунок 4.6 – Дерево решений для критерия Entropy

## Выводы по главе 4

На основе проведенного анализа можно сделать вывод, что модель XGBoost, используемая в проекте, демонстрирует высокую эффективность в задаче бинарной классификации. Модель успешно справляется с разделением классов, что подтверждается метриками качества, такими как Accuracy, Precision, Recall, F1-score и ROC-AUC, все из которых находятся на уровне, близком к идеальному. Это указывает на то, что модель не только точно классифицирует объекты, но и хорошо обобщает данные, что подтверждается результатами кросс-валидации.

По данной главе можно сделать следующие выводы выводы:

1. Высокая точность модели: Значение площади под ROC-кривой 0,9735 свидетельствуют о практически идеальном разделении классов. Это делает модель надежной для использования в задачах, где критически важно минимизировать ошибки классификации.

2. Стабильность модели: Значения метрик по результатам   
кросс-валидации показывают, что модель устойчива к изменениям в данных, что подтверждается низким отклонением метрик по фолдам.

3. Баланс между Precision и Recall: Модель демонстрирует высокие значения как Precision, так и Recall, что указывает на её способность минимизировать ложные срабатывания и пропуски.

4. Интерпретируемость модели: Анализ SHAP значений и важности признаков позволяет понять, какие факторы наиболее значимы для предсказаний. Такими признаки оказались FM, E и AFM они имеют наибольшее влияние на модель, что согласуется с результатами анализа дерева решений.

Таким образом, модель XGBoost версии 0-1 является эффективным инструментом для решения задачи классификации в рамках данного проекта. Её высокая производительность и стабильность делают её пригодной для дальнейшего использования и масштабирования на более крупных наборах данных.

# ГЛАВА 5 МОДЕЛЬ МАШИННОГО ОБУЧЕНИЯ ПО ВЕРСИЯМ 0–2

Для разделения данных по версиям 0 и 2 была обучена модель машинного обучения на основе модели XGboost. Набор данных делили в соотношении 80 % для обучения модели и 20 % для тестирования модели. Качество обучения модели оценивали по следующим метрикам: precision, recall, f1 и ROC AUC. Для понимания работы модели машинного обучения были выявлены наиболее определяющие признаки, а также оценено их влияние на выбор класса моделью машинного обучения.

## 5.1 Значение ROC AUC и матрица ошибок

Для оценки качества классификации был построен график ROC-кривой и вычислено значение площади под кривой, значение которой составило 0,9972, что говорит о высокой производительности модели. График ROC-кривой показан на рисунке 5.1.

Рассмотрим матрицу ошибок (confusion matrix) классификации на тестовых данных, которая приведена в таблице 5.1.

Таблица 5.1 – Матрица ошибок (версии 0 и 2)

|  |  |  |
| --- | --- | --- |
|  | **Предсказана версия 0** | **Предсказана версия 2** |
| **Действительная версия 0** | 32900 | 30 |
| **Действительная версия 2** | 115 | 23964 |

Полученная матрица ошибок показывает высокую долю правильно предсказанных классов из тестовой выборки. Метрика accuracy составила 99,75 %.

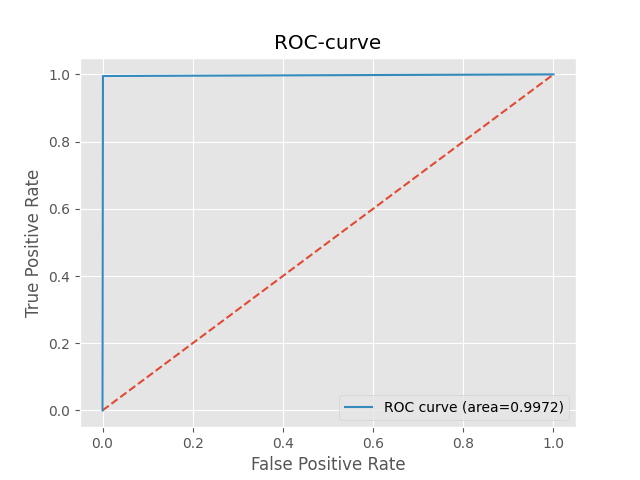


Рисунок 5.1 – График ROC-кривой для модели машинного обучения для классов 0 и 2.

## 5.2 Результаты перекрестной проверки (cross-validation)

При перекрестной проверке набор данных разбивался на 5 частей. Оценку результатов перекрёстной проверки по каждому разбиению проводили путем вычисления метрик: precision, recall, f1 и ROC AUC. Результаты перекрёстной проверки показаны на рисунке 5.2. Средние значения метрик по результатам перекрестной проверки показаны в таблице 5.2.

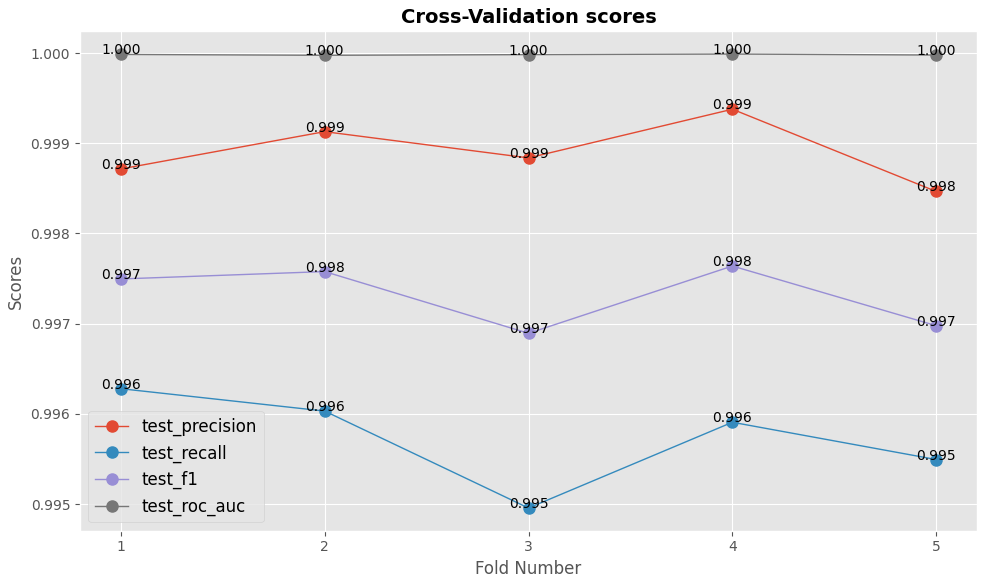


Рисунок 5.2 – Результаты перекрестной проверки модели машинного обучения для классов 0 и 2

Таблица 5.2 – Средние значения метрик

|  |  |
| --- | --- |
| **Метрика** | **Среднее значение** |
| Precision | 0,999 |
| Recall | 0,996 |
| F1 | 0,997 |
| ROC AUC | 1,000 |

Как видно по рисунку 5.2 значения метрик precision, recall, F1 более 0,99 для каждого из разбиений, что говорит о высоком качестве классификации моделью на разных данных.

## 5.3 Важность признаков

В ходе работы была проведена оценка признаков, которые оказывают наибольшее влияние на результат классификации. Результат оценки показан на рисунке 5.3.

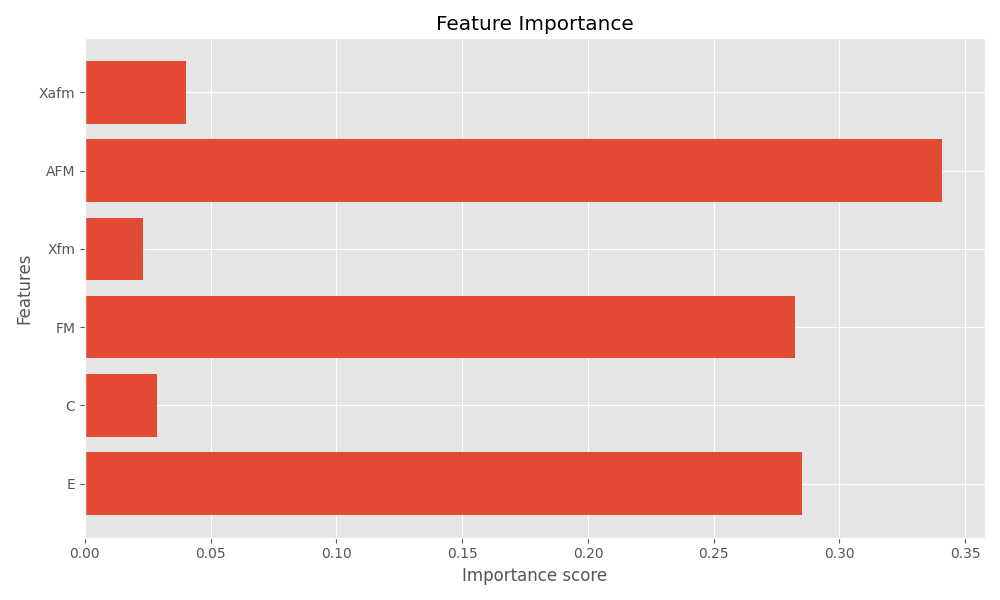


Рисунок 5.3 – График важности признаков при классификации версий 0 и 2

Наибольшее влияние на результат классификации версии модели сверхпроводника 0 или 2 оказывают признаки «AFM», «FM» и «E». Понимание того, какие признаки оказывают наибольший вклад в классификацию, может быть использовано для сокращения числа признаков в будущем и проверки производительности модели машинного обучения на сокращенном наборе признаков.

Также был построен SHAP-график важности признаков при определении версии, который приведен на рисунке 5.4.

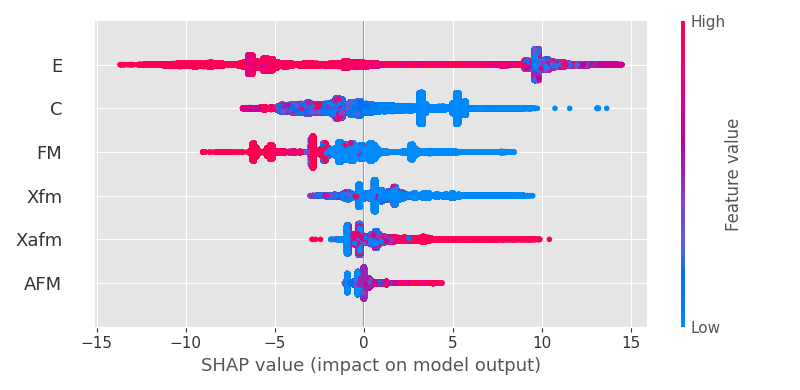


Рисунок 5.4 – SHAP-график важности признаков при классификации версий 0 и 2

По полученному графику можно сделать следующие выводы по параметрам, которые оказывают наибольшее влияние на результат классификации:

1. Чем ниже значение признака «FM», выше вероятность принадлежности объекта к второй версии модели сверхпроводника;
2. Чем выше значение признака «AFM», тем выше вероятность принадлежности объекта к второй версии модели сверхпроводника;
3. Высокие значения признака «E» характерны для обоих классов, однако объекты с низким значением признака «E» модель машинного обучения с большой вероятностью относит к объектам второй версии модели сверхпроводника.

## 5.4 Дерево решений

Для визуализации процесса классификации было построен график дерева решений с глубиной 3, который приведен на рисунке 5.5. Как видно из рисунка на 2 уровне модель уже определяет принадлежность 82435 объектов к версии 0. Также можно заметить, что в предикатах в основном используются признаки, оказывающие наибольшее влияние на результат классификации: «AFM», «FM» и «E».

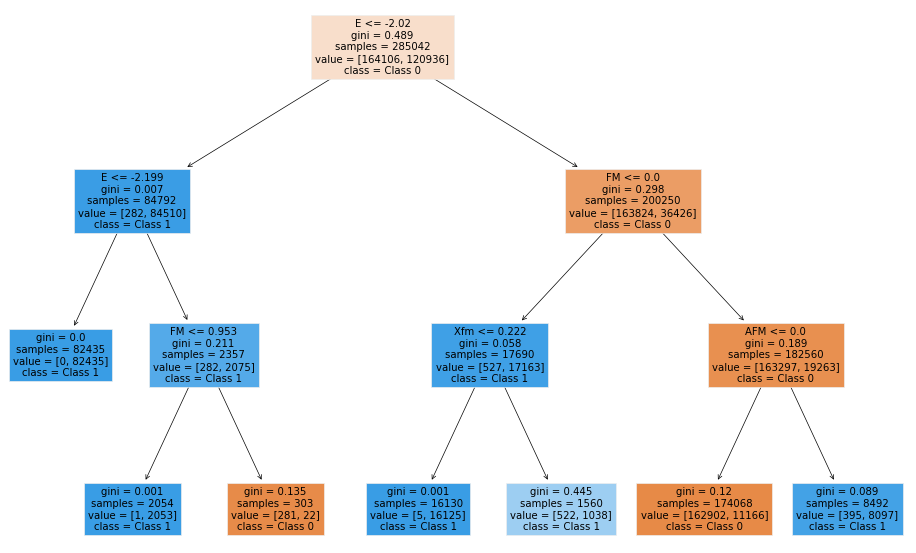


Рисунок 5.5 – Дерево решений для классификации версий 0 и 2

## Выводы по главе 5

1. Обученная модель машинного обучения показала высокую производительность для классификации объектов по версиям модели сверхпроводника 0 и 2. Значение ROC AUC составило 0,9972, значение метрики accuracy составило 99,75 %.
2. Метрики качества модели precision, recall, f1 и ROC AUC, полученные перекрестной проверкой путем разбиения набора данных на 5 частей (фолдов), показали значения выше 0,99, что говорит о высоком качестве классификации моделью на разных данных.
3. Определены признаки, которые оказывают наибольшее влияние на результат классификации версии модели сверхпроводника 0 или 2. Такими признаками оказались «AFM», «FM» и «E».

# ГЛАВА 6 АНАЛИЗ ЭФФЕКТИВНОСТИ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ПО МИНИМАЛЬНОМУ НАБОРУ ПАРАМЕТРОВ

В ходе работы одной из поставленных задач было определение минимально возможного набора параметров, которые можно использовать для эффективной классификации версий физической модели. Основной целью было выяснить, можно ли сократить количество параметров, используемых для обучения модели, без значительной потери точности классификации.

С физической точки зрения, наиболее логичными кандидатами для минимального набора параметров являются теплоёмкость (C) и магнитные свойства системы, такие как намагниченность (FM) или магнитная восприимчивость (Xfm). Эти параметры отражают ключевые термодинамические и магнитные характеристики системы, что делает их важными для анализа фазовых переходов и критического поведения, характерного для модели Изинга.

Для проверки гипотезы о том, что минимальный набор параметров (C и FM или C и Xfm) может быть достаточным для классификации версий модели Изинга, был проведён анализ с использованием модели машинного обучения XGBoost. Результаты анализа эффективности модели приведены в таблице 6.1 и на рисунках 6.1–6.4.

Таблица 6.1 – Результаты анализа эффективности модели по минимальному количеству признаков

|  |  |
| --- | --- |
| **Классификация 0 - 1** | **Классификация 0 - 2** |
| **Параметры C - FM** | **Параметры C - FM** |
| 1. Метрики качества модели:  * Accuracy: 95.10 % * Confusion Matrix: * Precision (Class 0): 0.94 * Precision (Class 1): 0.95 | 1. Метрики качества модели:  * Accuracy: 81.82 % * Confusion Matrix: * Precision (Class 0): 0.86 * Precision (Class 1): 0.77 |

Продолжение таблицы 6.1

|  |  |
| --- | --- |
| * Recall (Class 0): 0.87 * Recall (Class 1): 0.98 * F1-score (Class 0): 0.90 * F1-score (Class 1): 0.97 * ROC-AUC: 0.99  1. Результаты кросс-валидации:  * Accuracy: 95.64 % * Precision: 0.96 * Recall: 0.98 * F1-score: 0.97   ROC-AUC: 0.99 | * Recall (Class 0): 0.81 * Recall (Class 1): 0.83 * F1-score (Class 0): 0.83 * F1-score (Class 1): 0.80 * ROC-AUC: 0.90  1. Результаты кросс-валидации:  * Accuracy: 81.97 % * Precision: 0.77 * Recall: 0.84 * F1-score: 0.80   ROC-AUC: 0.90 |
| **Параметры C - Xfm** | **Параметры C - Xfm** |
| 1. Метрики качества модели:  * Accuracy: 91.10 % * Confusion Matrix: * Precision (Class 0): 0.86 * Precision (Class 1): 0.93 * Recall (Class 0): 0.79 * Recall (Class 1): 0.95 * F1-score (Class 0): 0.83 * F1-score (Class 1): 0.94 * ROC-AUC: 0.96  1. Результаты кросс-валидации:  * Accuracy: 91.22 % * Precision: 0.93 * Recall: 0.95 * F1-score: 0.94 * ROC-AUC: 0.97 | 1. Метрики качества модели:  * Accuracy: 93.87 % * Confusion Matrix: * Precision (Class 0): 0.96 * Precision (Class 1): 0.91 * Recall (Class 0): 0.93 * Recall (Class 1): 0.95 * F1-score (Class 0): 0.94 * F1-score (Class 1): 0.93 * ROC-AUC: 0.99  1. Результаты кросс-валидации:  * Accuracy: 93.75 % * Precision: 0.91 * Recall: 0.95 * F1-score: 0.93 * ROC-AUC: 0.99 |

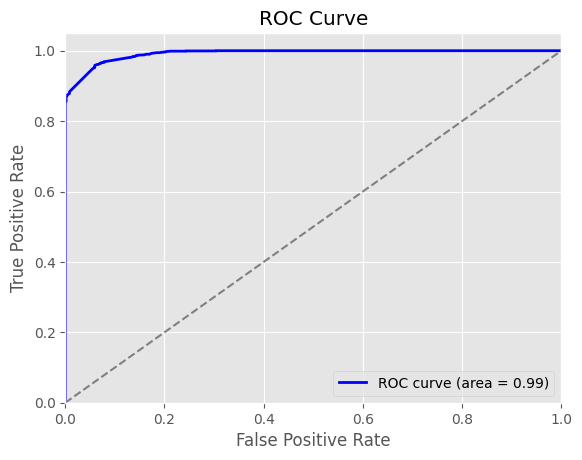


Рисунок 6.1 – ROC-кривая для классификации 0-1 по параметрам C и FM

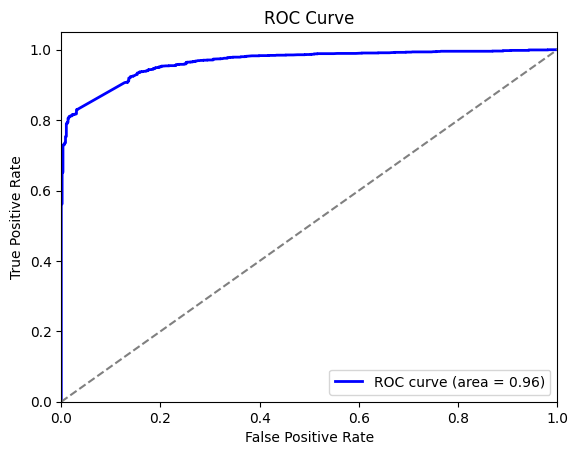


Рисунок 6.2 – ROC-кривая для классификации 0-1 по параметрам C и Xfm

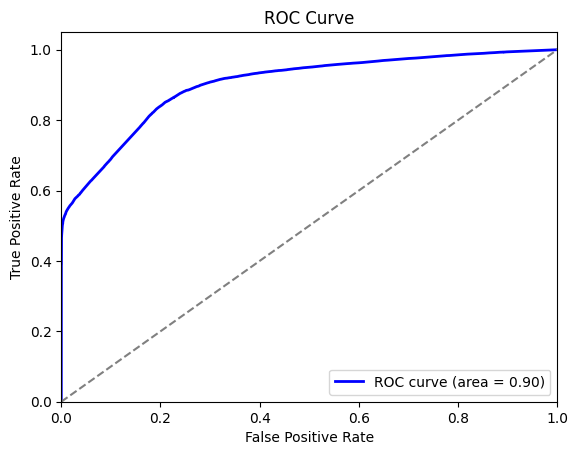


Рисунок 6.3 – ROC-кривая для классификации 0-2 по параметрам C и FM

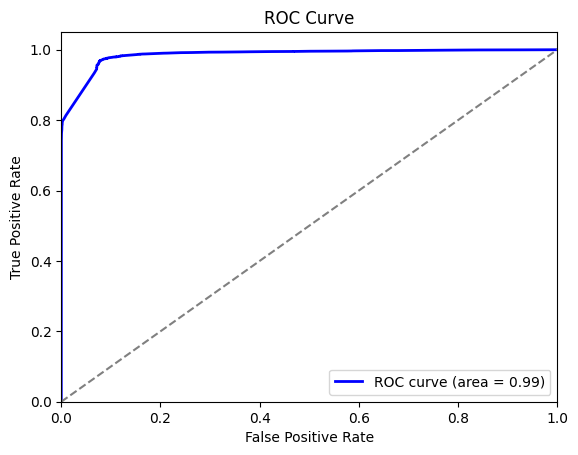


Рисунок 6.4 – ROC-кривая для классификации 0-2 по параметрам C и Xfm

Как видно из приведённых результатов анализа, при классификации версий 0 и 1, использование минимального набора параметров (C и FM или C и Xfm) показало высокую эффективность. Наилучшие результаты были достигнуты при использовании параметров C и FM, где точность (accuracy) составила 95.10 %, а средняя точность по результатам кросс-валидации — 95.64 %. Среднее значение ROC-AUC при кросс-валидации составило 0.99, что свидетельствует о высокой устойчивости модели к различным данным. При использовании параметров C и Xfm точность немного снизилась до 91.10 %, но средняя точность по результатам кросс-валидации осталась на уровне 91.22 %, а среднее значение ROC-AUC — 0.97, что также подтверждает эффективность этого набора параметров.

Для классификации версий 0 и 2 наилучшие результаты были достигнуты при использовании параметров C и Xfm. Точность (accuracy) составила 93.87 %, а средняя точность по результатам кросс-валидации — 95.64 %. Среднее значение ROC-AUC при кросс-валидации составил 0.99, что говорит о высокой устойчивости модели. При использовании параметров C и FM точность (accuracy) снизилась до 81.82 %, а средняя точность по результатам кросс-валидации — до 81.97 %. Это указывает на то, что намагниченность менее эффективна для классификации этих версий по сравнению с магнитной восприимчивостью.

# ГЛАВА 7 РАЗРАБОТКА API ДЛЯ РАБОТЫ С МОДЕЛЬЮ МАШИННОГО ОБУЧЕНИЯ

В рамках данной работы также был разработан API, который предоставляет возможность получать результаты предсказаний модели и оценки производительности (Accuracy, Precision, Recall, F1-score, ROC-AUC) на данных, загружаемых пользователем. Полный исходный код, документация, примеры использования и инструкции по развертыванию приложения размещены на платформе GitHub по ссылке <https://github.com/Maddenss/AI_phase_diagram_API>.

Проект имеет следующую структуру. Модель машинного обучения загружается из файла xgboost\_model.pkl, а для обучения и тестирования используются данные из файла Main\_Data.csv. Код API реализован в файле main.py, который связывает модель с эндпоинтами приложения (рисунок 7.1).

Для реализации API был выбран веб-фреймворк FastAPI, который представляет собой современный инструмент для разработки RESTful сервисов. FastAPI обеспечивает высокую производительность, поддержку асинхронных операций и автоматическую генерацию интерактивной документации с использованием OpenAPI и Swagger. Одним из ключевых преимуществ является его простота в использовании, что позволяет быстро разрабатывать и развертывать API, а также отличная интеграция с различными библиотеками для машинного обучения.

С подробной структурой проекта можно ознакомиться по ссылке <https://github.com/Maddenss/AI_phase_diagram_API>.

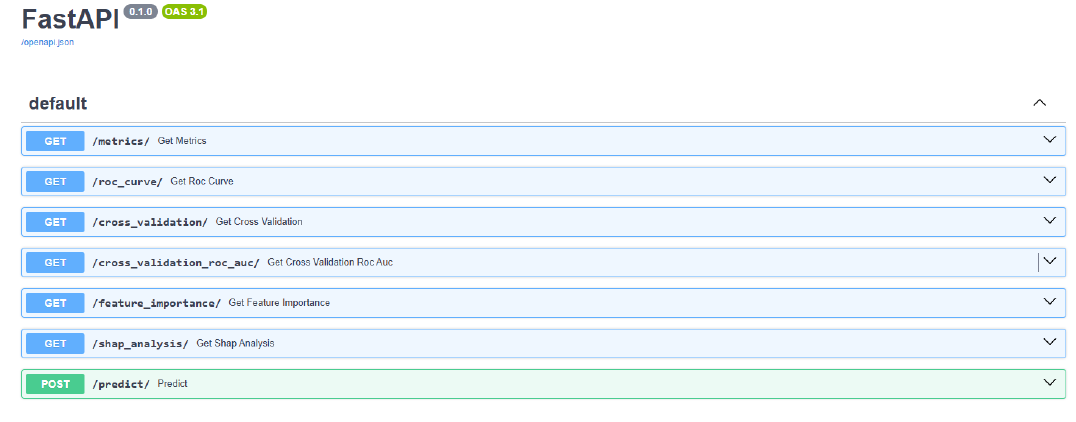


Рисунок 7.1 – Локальная версия приложения через FastAPI

Универсальность и стабильность обеспечиваются использованием Docker для контейнеризации приложения. Docker позволяет упаковать все необходимые компоненты — такие как код, библиотеки и зависимости — в единый контейнер, который можно запускать на любом сервере или локальной машине, независимо от операционной системы. Это обеспечивает стабильную и воспроизводимую среду выполнения, минимизируя риски, связанные с различиями в конфигурациях и зависимостях. Одним из основных преимуществ использования Docker является возможность изолировать приложение от хостовой системы, что упрощает управление зависимостями и устраняет проблемы с совместимостью. Кроме того, Docker ускоряет процесс развертывания, обеспечивая быстрый запуск и масштабирование приложения. Также контейнеры Docker могут быть легко интегрированы в процессы CI/CD (непрерывной интеграции и доставки), что способствует автоматизации тестирования и обновлений проекта.

API также был реализован через сервисе Streamlit, который представляет собой удобный и быстрый инструмент для создания веб-приложений с интерактивными визуализациями данных. Streamlit позволяет разработчикам без усилий создавать интерфейсы для моделей машинного обучения, аналитики и обработки данных, используя только Python: нет необходимости в сложной настройке фронтенда или написании HTML/CSS/JavaScript кода. Кроме того, Streamlit поддерживает динамическую визуализацию и обновление данных в реальном времени, что особенно полезно для анализа и демонстрации результатов работы моделей машинного обучения. Это делает Streamlit отличным инструментом для быстрого прототипирования и разработки интерактивных приложений, который позволяет значительно ускорить процесс разработки и предоставляет простое решение для пользователей, которым нужно взаимодействовать с результатами моделей через графический интерфейс (рисунок 7.2).

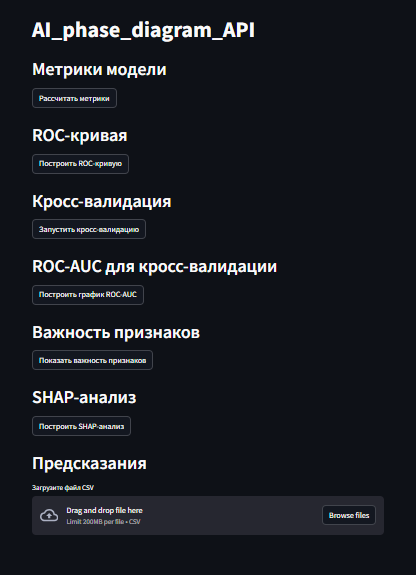


Рисунок 7.2 – Версия приложения, реализованная через Streamlit

API включает следующие эндпоинты:

1. /metrics/ — возвращает ключевые метрики модели: Accuracy, Precision, Recall, F1-score, ROC-AUC (рисунок 7.3).

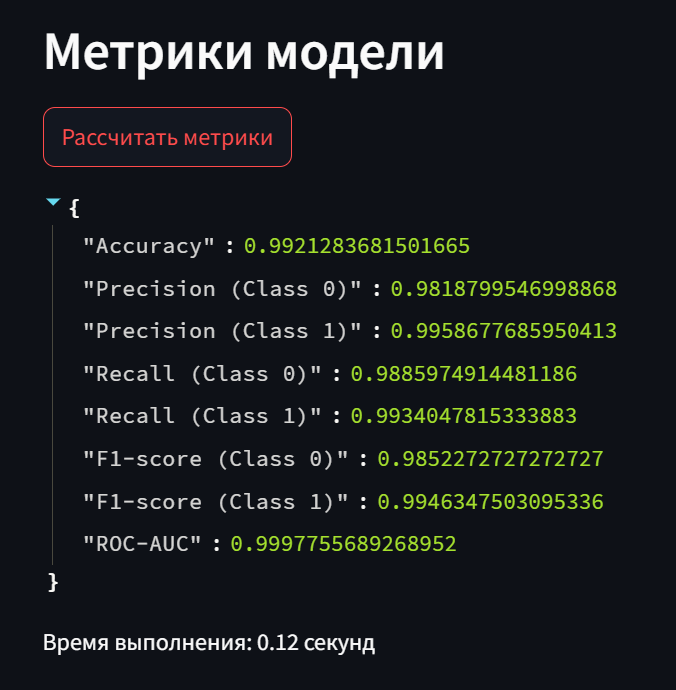


Рисунок 7.3 – Демонстрация вычисления метрик в сервисе Streamlit

2. /roc\_curve/ — генерирует и возвращает ROC-кривую в виде изображения (рисунок 7.4).

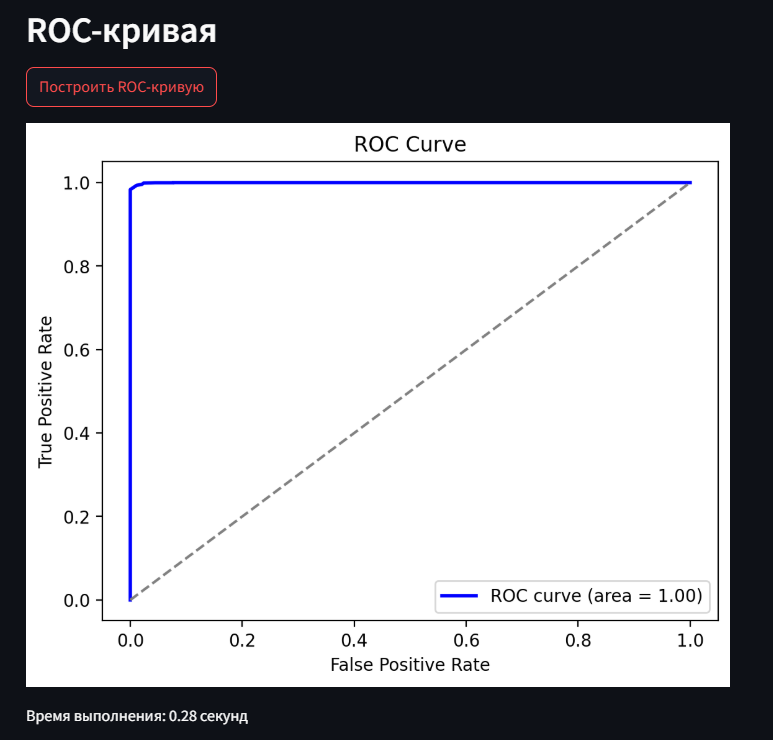


Рисунок 7.4 – Демонстрация генерациии ROC-кривой в сервисе Streamlit

3. /cross\_validation/ – выполняет 5-кратную кросс-валидацию, возвращая результаты для каждого фолда и среднее значение ROC-AUC (рисунок 7.5).



Рисунок 7.5 – Демонстрация выполнения 5-ти кратной кросс-валидации в сервисе Streamlit

1. /cross\_validation\_roc\_auc/ — строит график ROC-AUC по результатам кросс-валидации (рисунок 7.6).

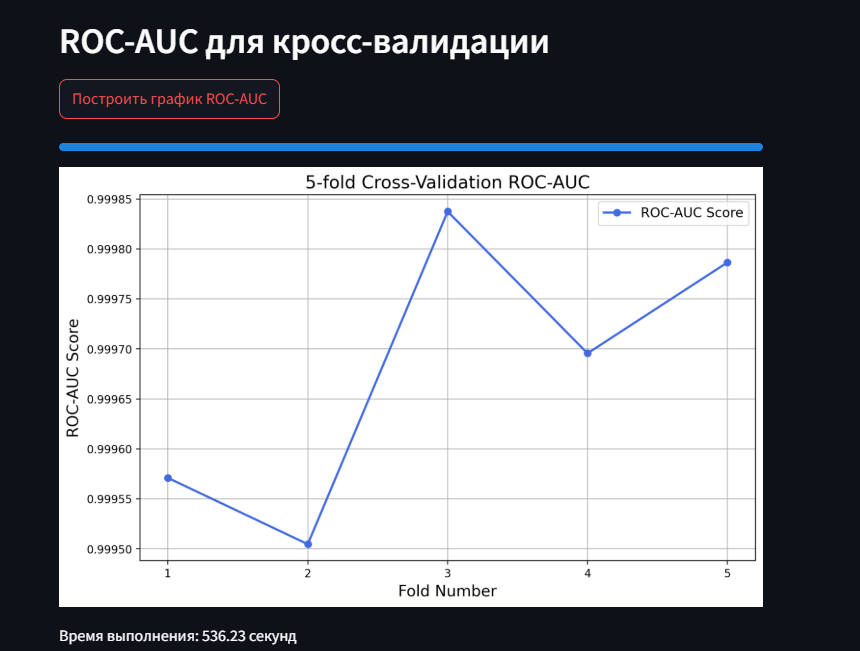


Рисунок 7.6 – Демонстрация построения ROC-кривой для кросс-валидации в сервисе Streamlit

5. /feature\_importance/ — генерирует график важности признаков модели (рисунок 7.7).

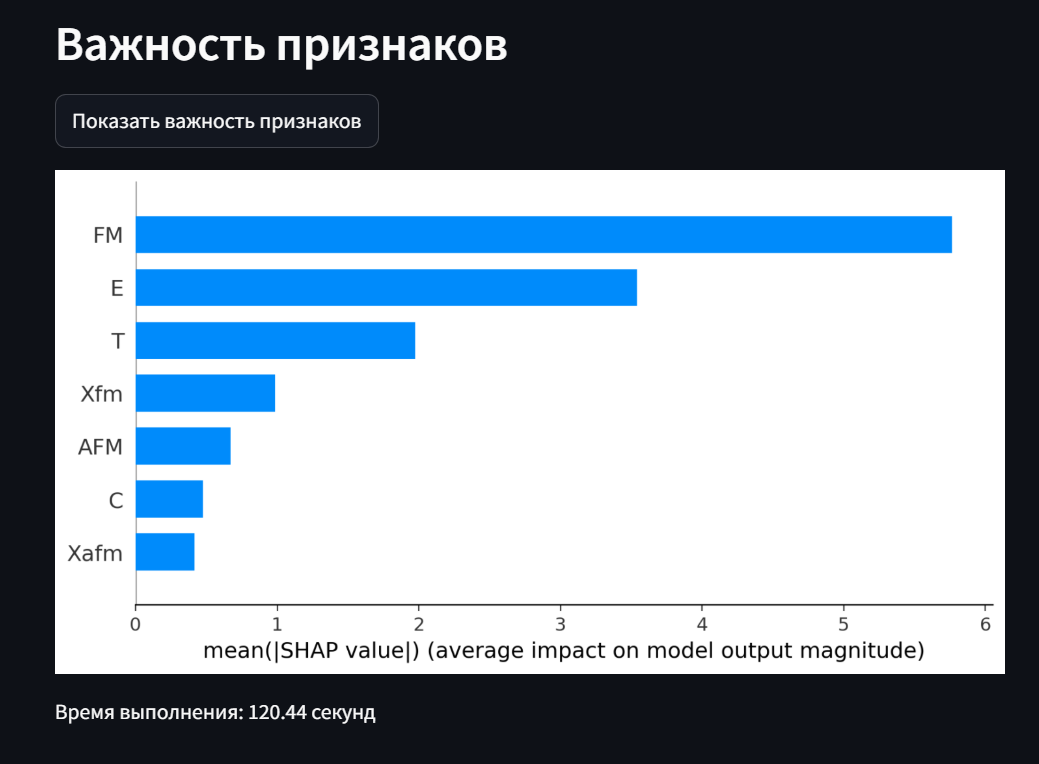


Рисунок 7.7 – Демонстрация анализа важности признаков в сервисе Streamlit

6. /shap\_analysis/ — визуализирует анализ SHAP для интерпретации модели (рисунок 7.8).

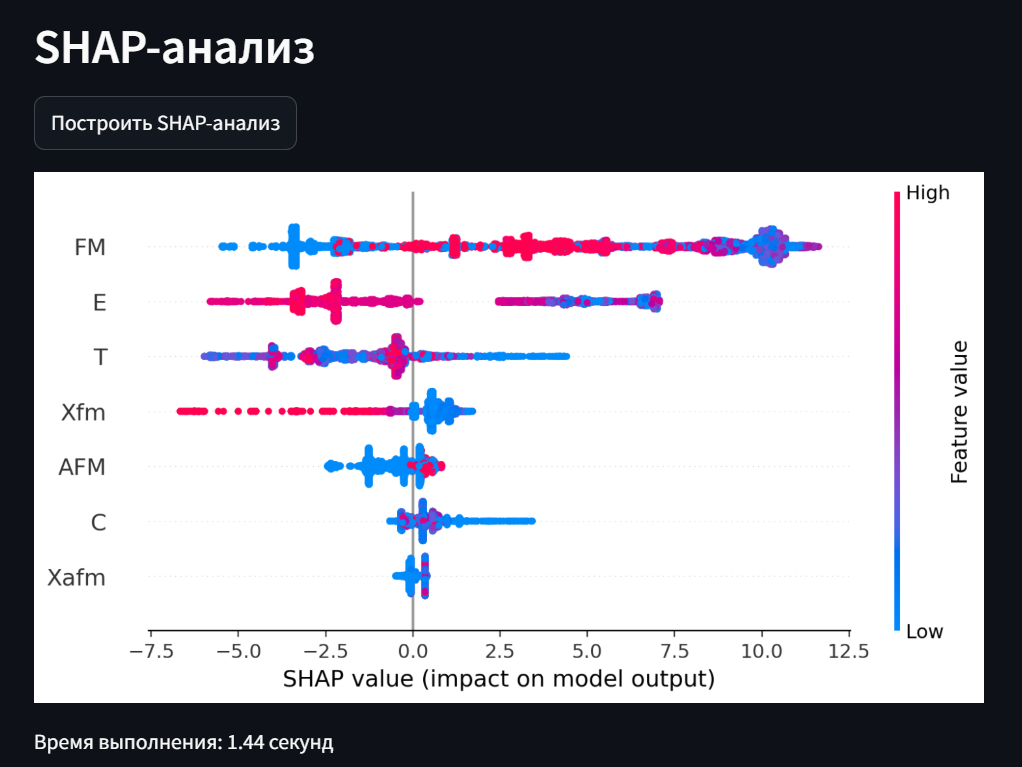


Рисунок 7.8 – Демонстрация анализа модели с помощью SHAP в сервисе Streamlit

7. /predict/ — принимает CSV-файл с данными, возвращает предсказания модели, вероятности и выполняет группировку по метке sample (рисунок 7.9).

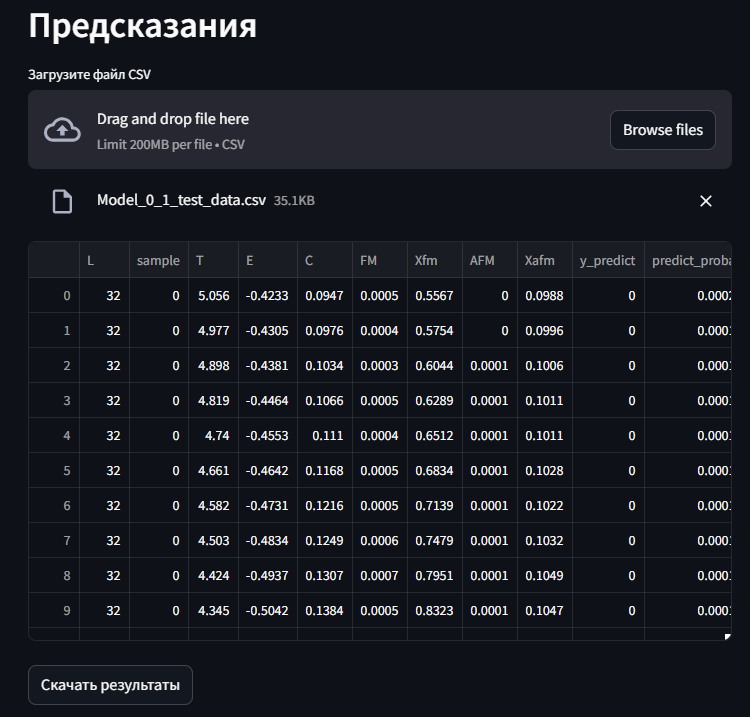


Рисунок 7.9 – Демонстрация вычисления предсказаний и вероятности классов для данных пользователя в сервисе Streamlit

Для обеспечения стабильности и работоспособности приложения была интегрирована система автотестирования с использованием GitHub Actions –инструмента для автоматизации CI/CD процессов, который запускает тесты и другие задачи при обновлении кода. Это обеспечивает оперативное выявление ошибок и поддержание высокого качества приложения. В рамках тестирования реализованы модульные, интеграционные, регрессионные, приемочные, нагрузочные и параметризованные тесты, что позволяет проверить корректность работы всех ключевых компонентов приложения.

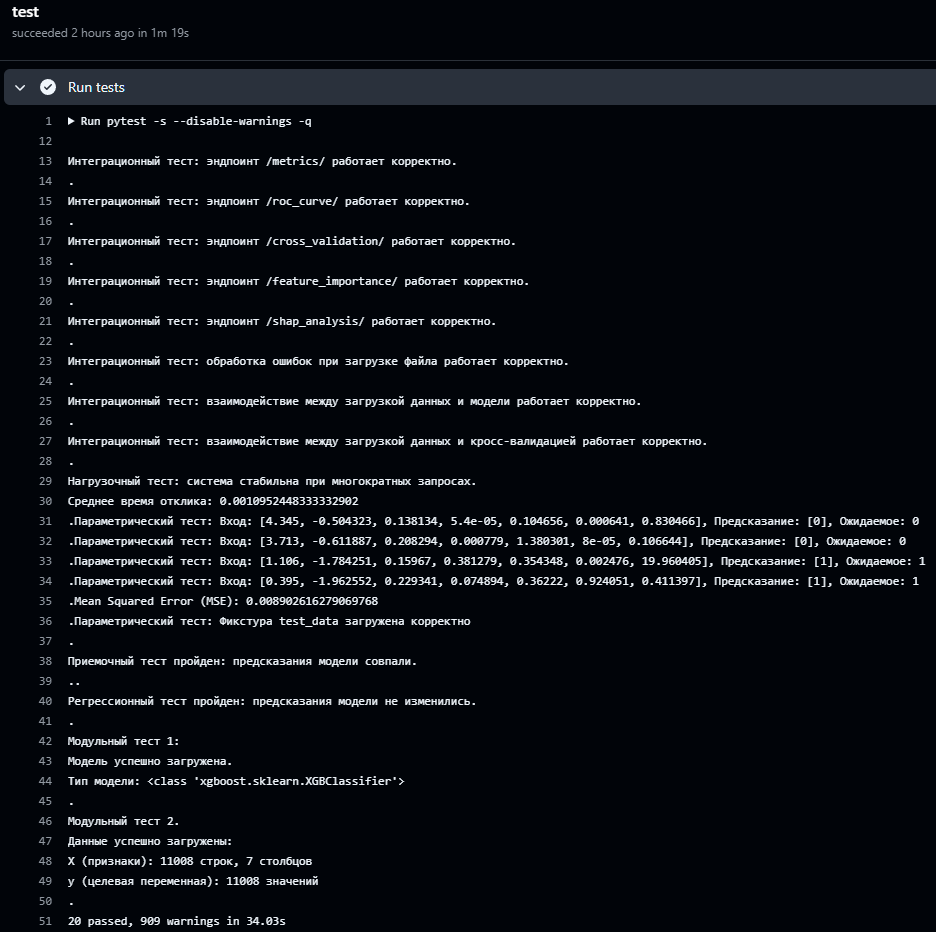


Рисунок 7.10 – Результат автотестирования API приложения через GitHub Actions

### ЗАКЛЮЧЕНИЕ

Подводя итог проделанной работы в рамках проектного практикума, нами была разработана модель машинного обучения на языке программирования Python для трех версий модели Изинга, «0» без внешнего магнитного поля, «1» с внешним магнитным полем и «2» с взаимодействием первых и вторых соседей. Одними из главных задач, стоящих перед нами, были автоматизация анализа данных и определение версии модели Изинга, которая наилучшим образом описывает наблюдаемые свойства системы. Проанализировав различные алгоритмы машинного обучения для решения задачи бинарной классификации, наилучший результат предсказания был получен при использовании алгоритма XGBoost. Этот алгоритм показал высокую производительность и надежность в условиях поставленной задачи.

На основе алгоритма XGBoost были разработаны модели машинного обучения для классификации версий моделей сверхпроводников 0–1 и 0–2. Каждая из моделей машинного обучения (0–1 и 0–2) показала высокую производительность и точность при определении классов объектов. Так значения площади под ROC-кривой для моделей версии 0–1 и версии 0–2 составили 97,35 %, и 99,72 % соответственно.

Стабильность модели была проверена с помощью перекрестной проверки путем разбиения набора данных на 5 частей (фолдов). Метрики по всем фолдам показали значения выше 0,97 для обоих моделей, что позволяет сделать выводы о высоком качестве классификации моделями на разных данных.

Оценка важности признаков, которые оказывают наибольшее влияние на результат классификации, показала, что при определении версии 0-1 наибольшее влияние на результат классификации оказывают ферромагнитная намагниченность (FM) («importance score» более 0,45), энергия системы (E) со значением 0,22 и антиферромагнитная намагниченность (AFM), а при определении версии 0-2 антиферромагнитная намагниченность (AFM), ферромагнитная намагниченность (FM) и энергия системы (E) со значениями примерно в 0,37, 0,28 и 0,27 соответственно.

Визуализация дерева решений подтвердила, что модели как для версии 0–1, так и для версии 0–2 для классификации используют наиболее важные признаки: FM, E и AFM.

Проведены исследования возможности определения версии модели по минимальному набору признаков (C и FM или C и Xfm) по требованию заказчика. Для версии 0–1 наилучшие результаты были достигнуты при использовании параметров C и FM, где средняя точность по результатам кросс-валидации составила 95.64 %, а среднее значение ROC-AUC 0.99. Для классификации версий 0–2 наилучшие результаты были достигнуты при использовании параметров C и Xfm. Средняя точность по результатам кросс-валидации составила 95.64 %, среднее значение ROC-AUC при кросс-валидации составило 0.99.

В рамках данной работы также был разработано приложение, который предоставляет возможность получать результаты предсказаний модели и оценки производительности (Accuracy, Precision, Recall, F1, ROC-AUC) на данных, загружаемых пользователем. Полный исходный код, документация, примеры использования и инструкции по развертыванию приложения размещены на платформе GitHub по ссылке <https://github.com/Maddenss/AI_phase_diagram_API>.

Дальнейшие исследования и разработки могут быть направлены на оптимизацию модели и расширение её функциональности, например, на определение других версий моделей сверхпроводников. В дальнейшей работе возможно провести улучшение функциональности приложения, в частности предоставить возможность предсказания версии 0–2 и выбор отдельных признаков пользователем.