

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение высшего образования  
«Южно-Уральский государственный университет (национальный исследовательский университет)»  
Высшая школа электроники и компьютерных наук  
Кафедра системного программирования

# **Разработка системы для предсказания успеваемости студентов на основе данных портала «Электронный ЮУрГУ»**

Научный руководитель:

профессор кафедры СП, д.ф.-м.н., доцент

М.Л. Цымблер

Автор работы:

студент группы КЭ-403

Д.В. Старостенок

Челябинск 2024 г.

# АКТУАЛЬНОСТЬ

- Тенденции использования электронных образовательных систем
- Имеется система, хранящая данные об успеваемости студентов

edu@SUSU Личный кабинет Мои курсы Помощь

Установочная информация

ВШЭЭН / СП / 2022/2023 / Весенний семестр / Бакалавриат / Очная

## 2022/2023 Программная инженерия (очная)

Курс Участники

Установочная информация [Свернуть всё](#)

**Аннотация дисциплины**

Дисциплина "Программная инженерия" направлена на ознакомление слушателей с техникой анализа, проектирования и реализации программного обеспечения.

Задачи дисциплины:

1. Познакомить с процессами разработки программного обеспечения;
2. Познакомить с методами спецификации требований и разработке ПО;
3. Познакомить с методами аттестации и развития ПО;
4. Познакомить с основными методологиями управления проектами разработки ПО;
5. Сформировать компетенции применения в практической деятельности различных подходов к управлению проектами разработки ПО

ИАС Университет Помощь English Русский 中文 Здравствуйте, Дмитрий Владимирович! Учетная запись Выход

Личный кабинет студента/аспиранта

Дисциплина: Программная инженерия

Журнал Журнал КР/КП

Имя и фамилия	Макс. балл/вес	07. Представление диаграммы деятельности и последовательности		08. Компиляция отчета и презентации		Текущий контроль всего	Бонус			Промежуточная аттестация	Итого	
		Балл	Рейтинг, %	Балл	Рейтинг, %		Рейтинг, %	Рейтинг, %	Рейтинг, %			
Старостенок Дмитрий Владимирович	0	7	100	6	75	89,62	0	0	0	0	89,62	5 (отлично)



Дисциплина: **Программная инженерия**

	Рейтинг, %	Оценка
Старостенок Дмитрий Владимирович	89,62	5 (отлично)

<b>Факультет:</b>	Высшая школа электроники и компьютерных наук
<b>Кафедра:</b>	Системное программирование
<b>Направление:</b>	Программная инженерия
<b>Период обучения:</b>	4 лет/год(-а), 0 месяцев(-ев/-а)
<b>Квалификация:</b>	бакалавр

## ЦЕЛЬ И ЗАДАЧИ ИССЛЕДОВАНИЯ

---

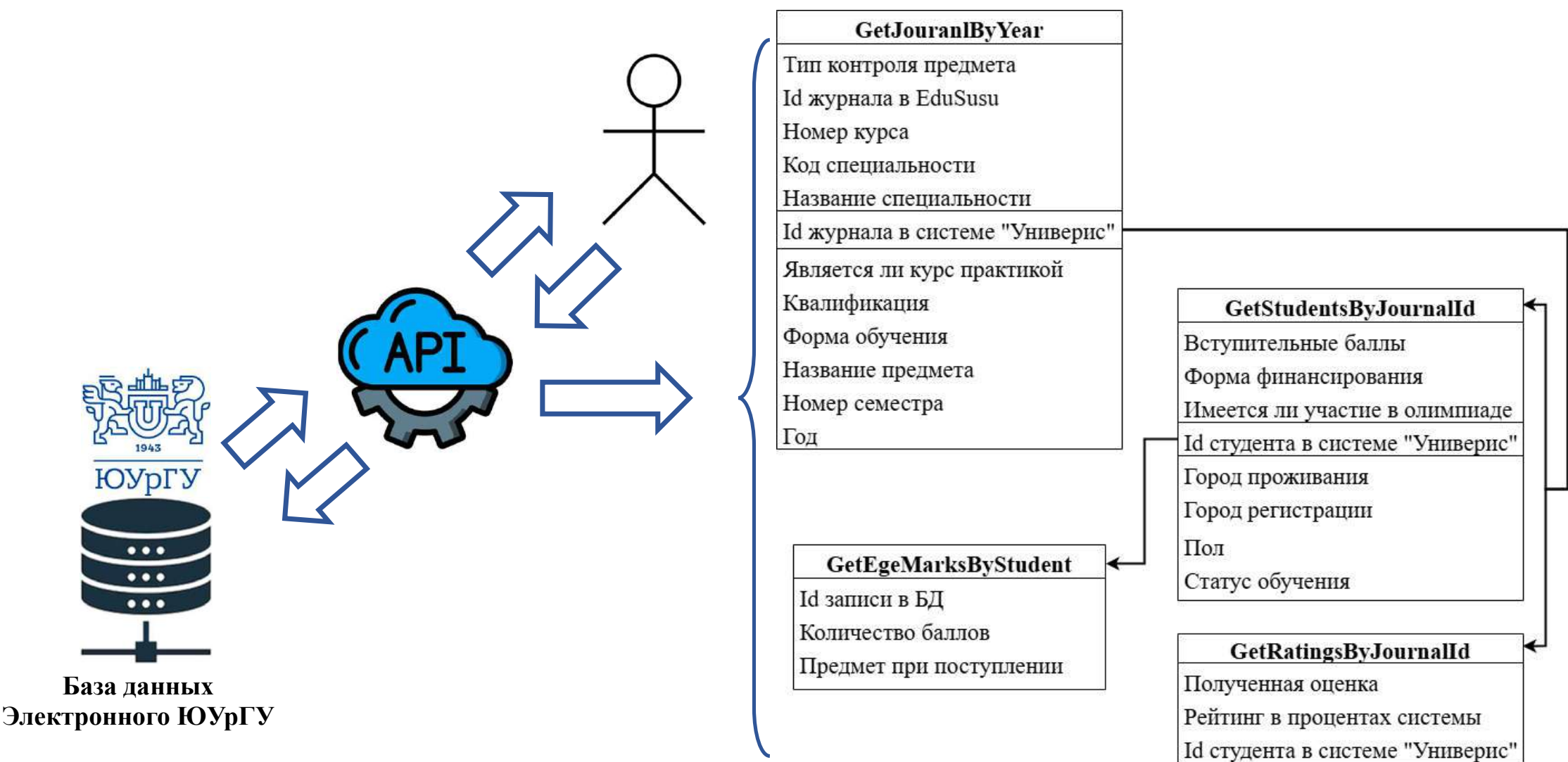
### **Цель:**

Разработка системы для предсказания успеваемости студентов на основе данных Электронного ЮУрГУ

### **Задачи:**

1. Провести анализ предметной области и на его основе разработать модели для предсказания
2. Разработать алгоритм получения данных с применением существующих методов предобработки
3. Разработать приложение, выполняющее предсказание на основе моделей с помощью различных методов анализа данных
4. Провести эксперименты, исследующие точность разработанных моделей данных

# АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ



# ПРОГНОЗНЫЕ МОДЕЛИ

## Предсказание оценки дисциплины за указанный семестр

- Нативные данные студента

Оценка по предмету за N семестр

## Предсказание оценки дисциплины, проходящей в диапазоне семестров

- Нативные данные студента
- Оценка по предмету за N семестр
- Оценка по предмету за N+K семестр

Оценка по предмету за N+2 семестр

## Предсказание оценки по практикам

- Нативные данные студента
- Дециль суммы рейтингов студента по всем дисциплинам N+1 семестра
- Дециль суммы рейтингов студента по всем дисциплинам N+K семестра

Успешность прохождения практики

## Предсказание отчисления студентов по экзаменам/зачетам

- Нативные данные студента
- Дециль суммы рейтингов студента по всем дисциплинам N+K семестра
- Дециль медианного рейтинга студента по дисциплинам экзаменов/зачетов N+K семестра

Отчислен/Учится

## Предсказание успеваемости на основе рейтинга по экзаменам/зачетам

- Нативные данные студента
- Дециль суммы рейтингов студента по всем дисциплинам N+K семестра
- Дециль медианного рейтинга студента по дисциплинам экзаменов/зачетов N+K семестра

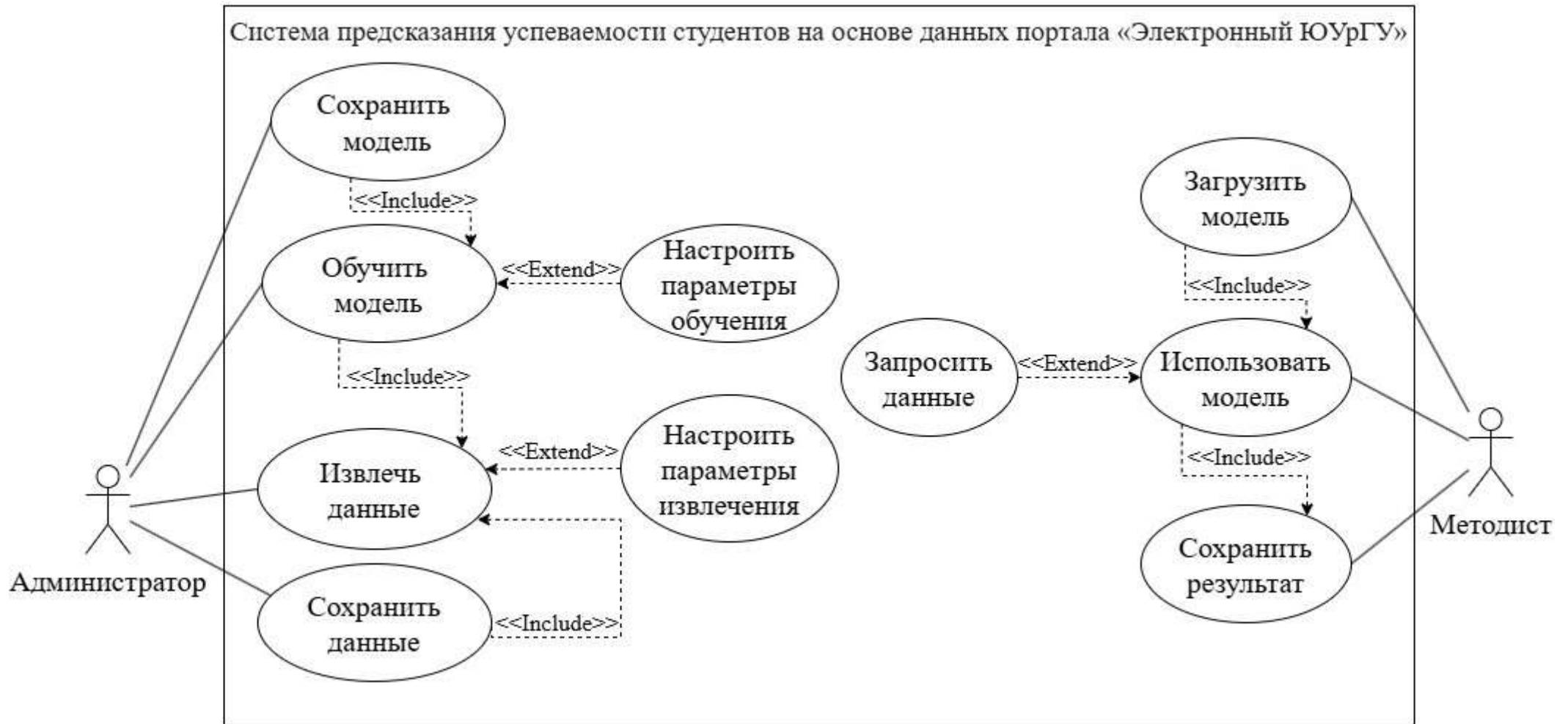
Оценка студента по экзаменам/зачетов следующего семестра

## Предсказание успеваемости при завершении обучения

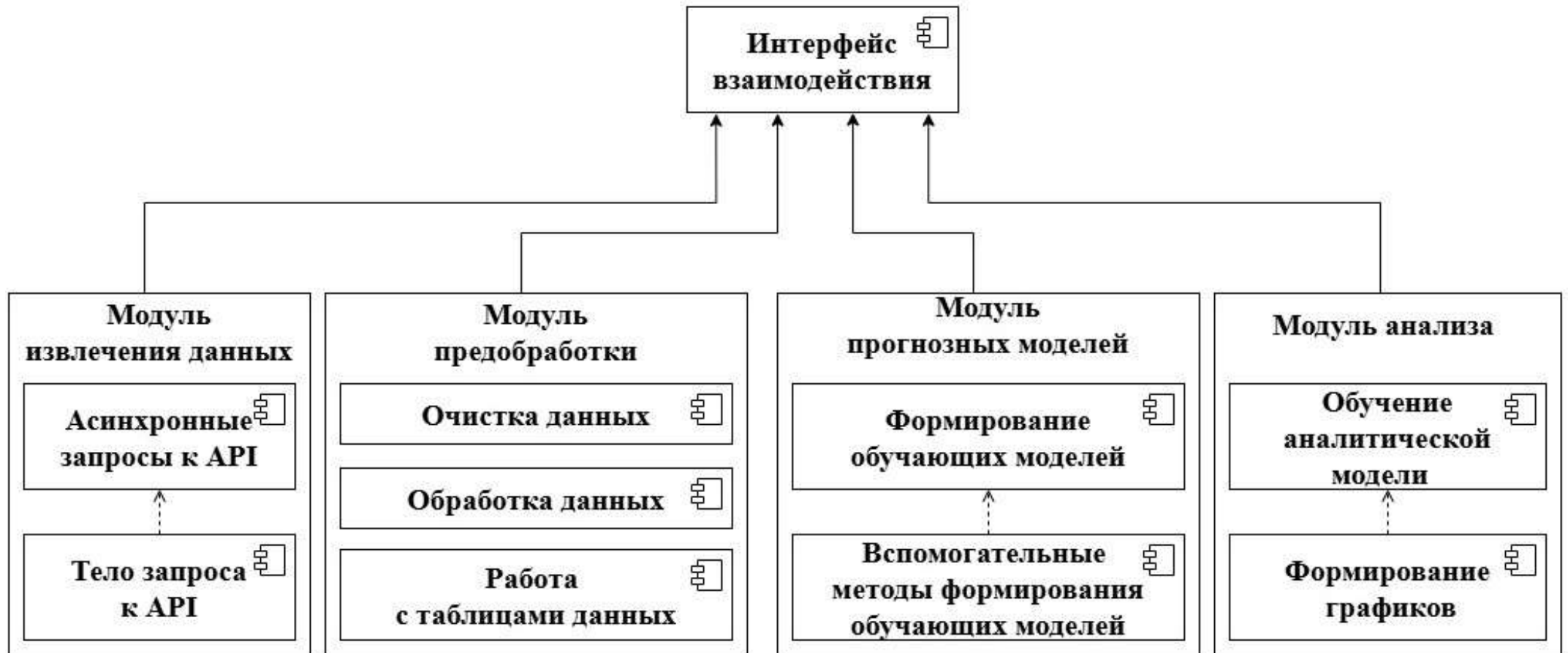
- Нативные данные студента
- Дециль суммы рейтингов студента по всем дисциплинам N+K семестра

Средний балл итогового рейтинга

# МОДЕЛЬ ВАРИАНТОВ ИСПОЛЬЗОВАНИЯ



# МОДУЛЬНАЯ СТРУКТУРА



## СРЕДСТВА РАЗРАБОТКИ

**Язык программирования:** Python 3.12

**Верстка интерфейса:** HTML5, CSS, JavaScript

**Библиотеки:**

Получение данных	<ol style="list-style-type: none"><li>1. Asyncio</li><li>2. Aiohttp</li><li>3. Tenacity</li></ol>
Подготовка данных	<ol style="list-style-type: none"><li>1. Pandas</li><li>2. Numpy</li></ol>
Предсказание	Scikit-Learn – Случайный лес – Градиентный бустинг
Интерфейс пользователя	<ol style="list-style-type: none"><li>1. Django</li><li>2. Matplotlib</li></ol>

**Репозиторий:**

<https://github.com/Hlebnica/PredictingStudentPerformance>





# ИНТЕРФЕЙС ИЗВЛЕЧЕНИЯ ДАННЫХ

DigitalTrace Извлечение данных ▾ Обучение моделей предсказания Предсказание Информация о программе

Извлечение данных из «Электронного ЮУрГУ»

Диапазон годов: От 2019 До 2022  
 Диапазон учебных курсов: От 1 До 4  
 Высшая школа: Высшая школа электроники и компьютерных наук

[Сделать запрос](#)

[Скачать Журнал за 2019-2022 год](#)  
[Скачать Информацию о студентах за 2019-2022 год](#)  
[Скачать Рейтинг студентов по предметам за 2019-2022 год](#)  
[Скачать Результаты ЕГЭ студентов за 2019-2022 год](#)

Журнал за 2019-2022 учебный год (количество строк - 4516)

Тип контроля	Id в EduSusu	Номер курса	Код специальности	Название специальности	Id группы	Id в "Универис"	Является ли практикой	Квалификация	Форма обучения
зачет	0	1	10.03.01	Информационная безопасность	b4831bcd-e3e7-49cd-90e1-d10c593a254c	d6309ece-5b63-424b-9aa9-002fdca50d54	False	бакалавр	очная
зачет	0	1	27.03.04	Управление в технических системах	88628705-4e86-424d-8e08-ce72942285c7	363b48d9-72f6-4a33-aac2-0079f98f0194	False	бакалавр	очная
зачет	0	1	27.03.04	Управление в технических системах	88628705-4e86-424d-	39cc58d5-58ee-4ade-	False	бакалавр	очная

Информация о студентах за 2019-2022 учебный год (количество строк - 79541)

Вступительные баллы	Форма финансирования обучения	Участие в олимпиаде	Id в "Универис"	Город проживания	Город регистрации	Пол	Статус обучения	Id записи	Id группы
227	бюджет	False	174da706-d8a0-4a97-bf83-405bdce504f3	г.Челябинск	г.Челябинск	Мужской	отчислен	d6309ece-5b63-424b-9aa9-002fdca50d54	b4831bcd-e3e7-49cd-90e1-d10c593a254c
159	контракт	False	20da1c94-e28c-48b8-ba89-2703850957b4	г.Челябинск	г.Челябинск	Мужской	отчислен	d6309ece-5b63-424b-9aa9-002fdca50d54	b4831bcd-e3e7-49cd-90e1-d10c593a254c
166	контракт	False	2f425f-4c5c-	г.Челябинск	г.Челябинск	Мужской	отчислен	d6309ece-	b4831bcd-

# ИНТЕРФЕЙС АНАЛИЗА ДАННЫХ

Обучение моделей предсказания

**Модель данных**  
Предсказание оценки дисциплины за указанный семестр

**Диапазон учебных курсов**  
От (семестр)  
1

**Предмет**  
Математический анализ

**Аналитическая модель**  
Случайный лес (классификатор)

► Информация об аналитических моделях

**Гиперпараметры**

Размер тестовой выборки 0,2	Минимальное количество объектов в узле 2	Проводить увеличение выборки? <input checked="" type="checkbox"/>
Количество деревьев 2000	Минимальное количество объектов в листьях деревя 1	Проводить удаление выбросов? <input type="checkbox"/>
Максимальная глубина деревьев 20	Количество признаков при построении деревьев 1	Проводить удаление шумов? <input checked="" type="checkbox"/>

► Информация о гиперпараметрах

**Обучить модель**

# ИНТЕРФЕЙС АНАЛИЗА ДАННЫХ

Составленная модель данных (выборка модели - 1566 строк)

Пол	1 дециль оценки за ЕГЭ	2 дециль оценки за ЕГЭ	3 дециль оценки за ЕГЭ	Город регистрации	Форма финансирования обучения	Оценка за 1 семестр
1	4	3	4	1	1	2
1	2	1	1	1	0	2
1	1	2	1	1	0	2
1	3	3	4	1	1	2
1	3	3	2	0	0	2
1	2	2	1	1	0	2
1	2	2	5	0	0	2
1	5	3	3	1	1	4

Результат тестирования модели

Таблица предсказанных результатов

Пол	1 дециль оценки за ЕГЭ	2 дециль оценки за ЕГЭ	3 дециль оценки за ЕГЭ	Город регистрации	Форма финансирования обучения	Оценка за 1 семестр	Предсказанный результат
1	2	2	2	1	1	3	3
1	1	1	1	1	0	2	2
1	2	2	2	1	1	3	3
1	1	1	1	0	1	3	2
1	4	4	3	0	1	3	3
1	3	3	2	1	1	3	3
1	5	5	4	1	1	4	4
1	4	3	5	0	1	3	3

Метрики:  
 Accuracy оценка: 0.73  
 Precision оценка: 0.73  
 Recall оценка: 0.75  
 F1 оценка: 0.72

Сохранить результат анализа данных

Сохранить обученную модель

# ИНТЕРФЕЙС АНАЛИЗА ДАННЫХ

Диаграмма важности признаков

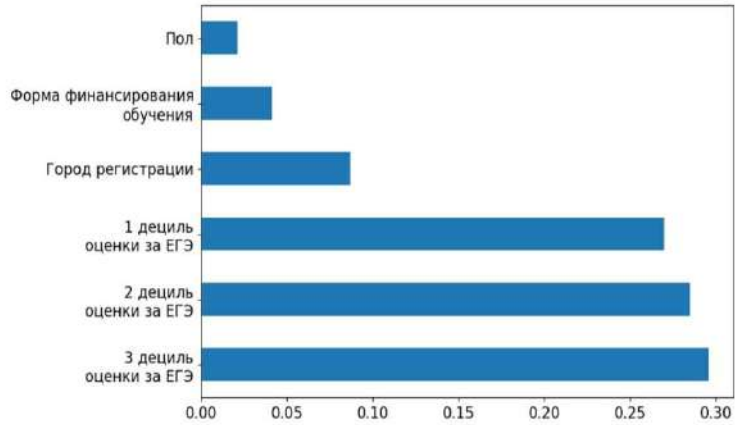
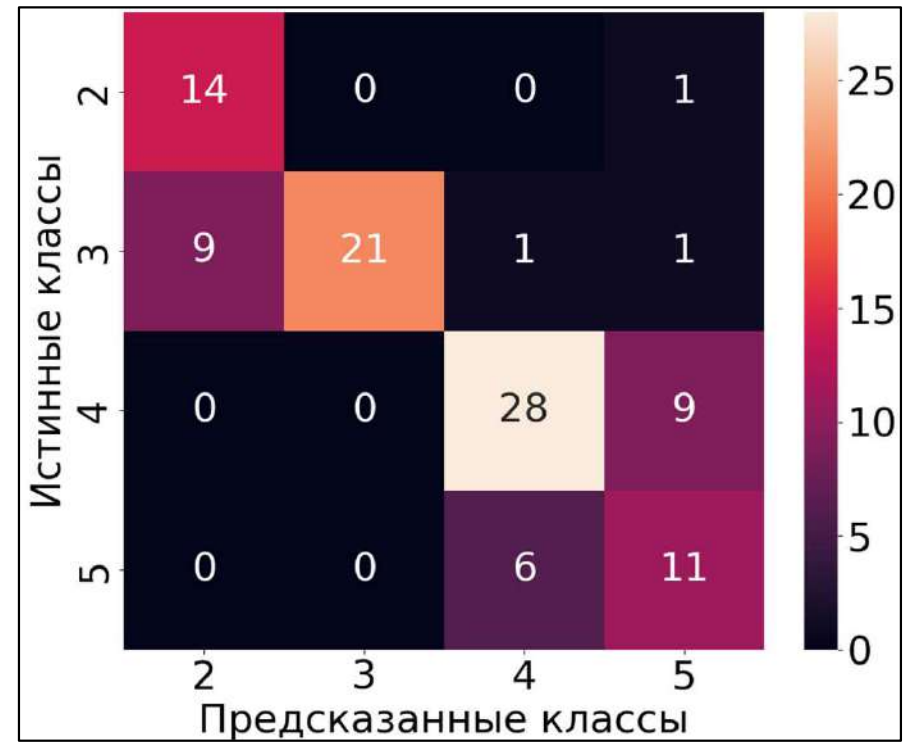
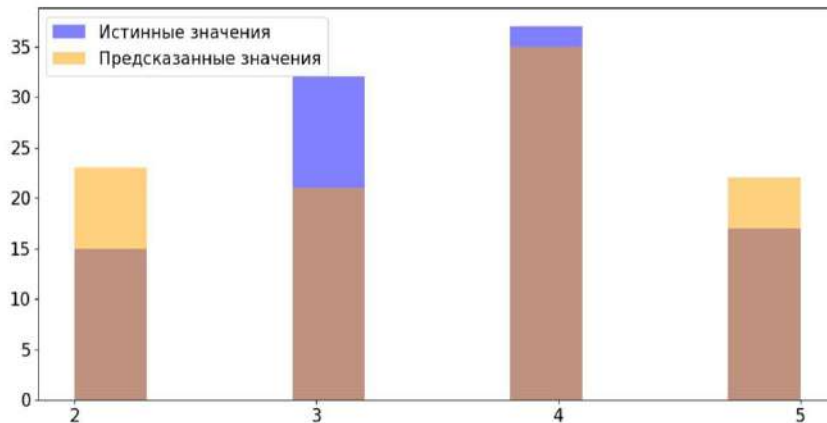


Диаграмма предсказаний и реальных значений



# ИНТЕРФЕЙС ПРЕДСКАЗАНИЯ

Предсказание

Файл с обученной моделью

Выбор файла  Не выбран ни один файл

Диапазон годов, по которым будет идти предсказание

От  До

Информации о загруженной модели, по которой проводился анализ

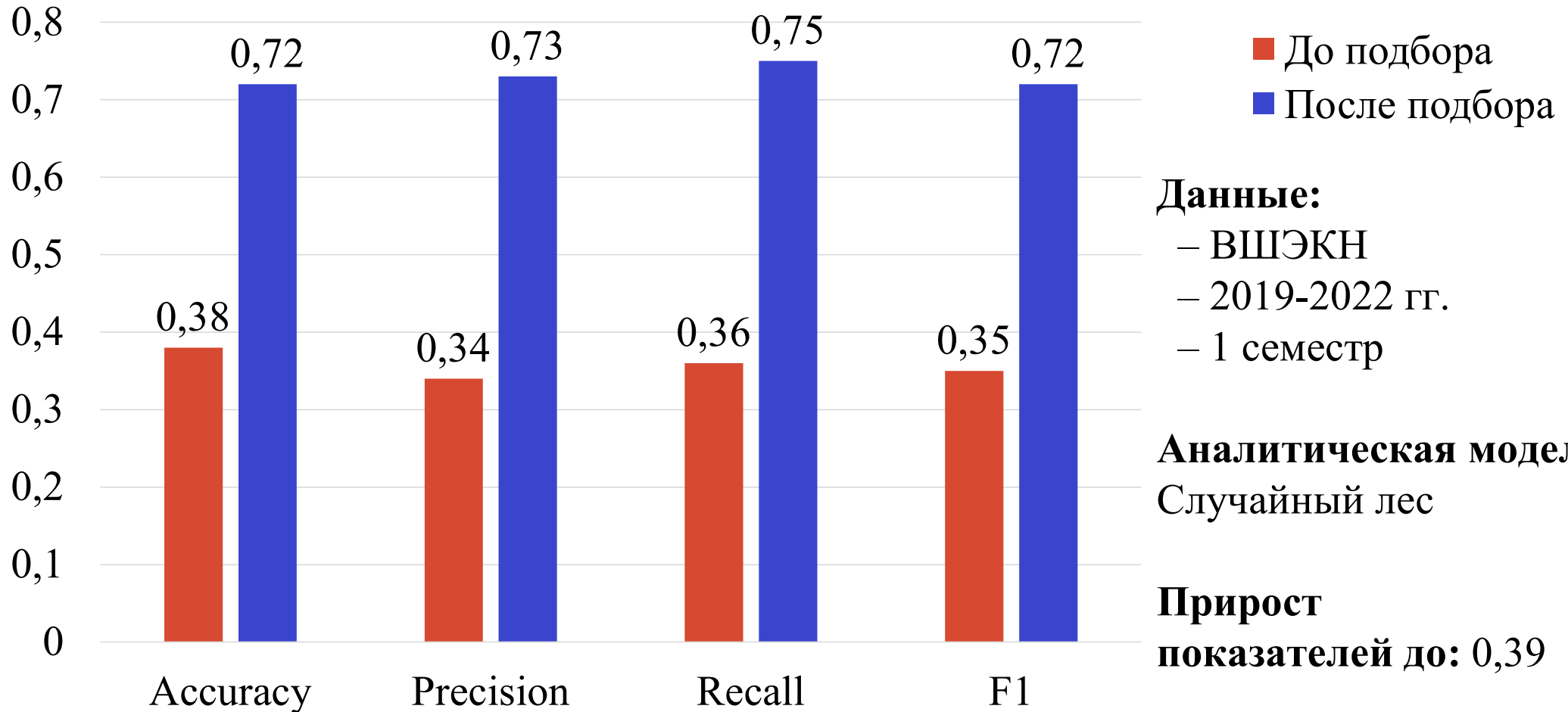
**Высшая школа:** Высшая школа электроники и компьютерных наук  
**Диапазон годов:** от 2019 до 2022  
**Семестр:** 1  
**Предмет:** Математический анализ  
**Метрики обученной модели:**  
Accuracy оценка: 0.73  
Precision оценка: 0.73  
Recall оценка: 0.75  
F1 оценка: 0.72

Результаты предсказания

Id студента	Пол	1 дециль оценки за ЕГЭ	2 дециль оценки за ЕГЭ	3 дециль оценки за ЕГЭ	Город регистрации	Форма финансирования обучения	Предсказанный результат
0b22c5e8-dcaf-4ab7-a11a-35f88004e09a	Мужской	1	1	1	Иногородний	Бюджет	2
107dff39-cc02-45cc-baaa-798da48ac31b	Мужской	5	5	5	Иногородний	Бюджет	5
1e85052e-d031-45b5-bb89-bea778599759	Мужской	2	2	1	Иногородний	Бюджет	3
2d32a7ab-8793-4e2a-bd34-774b693a24dc	Женский	2	1	1	Иногородний	Бюджет	3
4831659c-56f2-49fd-9d24-e151f3cb88db	Мужской	2	4	2	Иногородний	Бюджет	3

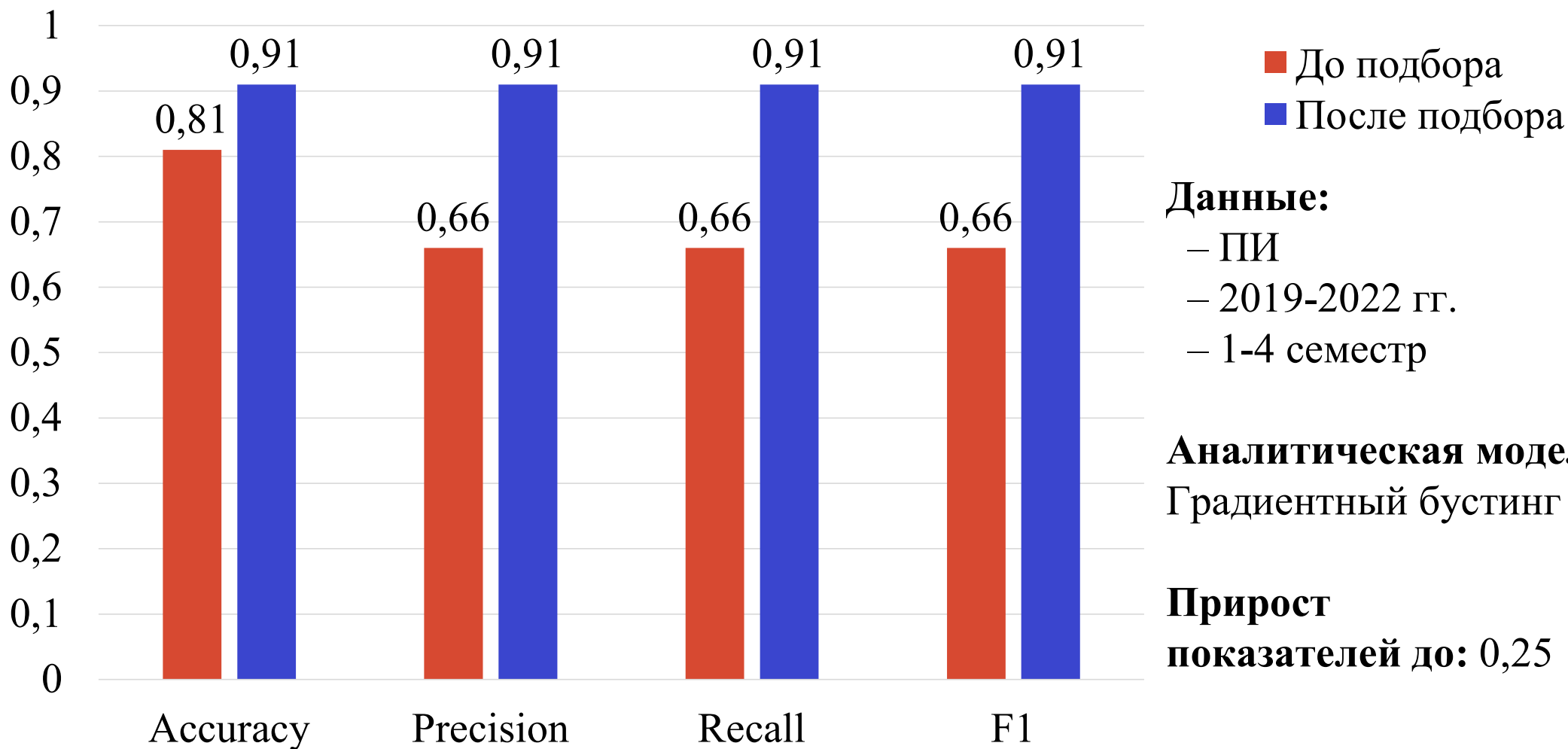
## ЭКСПЕРИМЕНТЫ

### Предсказание оценки дисциплины за указанный семестр



# ЭКСПЕРИМЕНТЫ

## Предсказание отчисления по экзаменам



# ПРЕДСКАЗАТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

## Оценки за 1 семестр по математическому анализу

Id студента	Пол	1 дециль оценки за ЕГЭ	2 дециль оценки за ЕГЭ	3 дециль оценки за ЕГЭ	Город регистрации	Форма финансирования обучения	Предсказанный результат
0b22c5e8-dcaf-4ab7-a11a-35f88004e09a	Мужской	1	1	1	Иногородный	Бюджет	2
107dff39-cc02-45cc-baaa-798da48ac31b	Мужской	5	5	5	Иногородный	Бюджет	5
1e85052e-d031-45b5-bb89-bea778599759	Мужской	2	2	1	Иногородный	Бюджет	3
2d32a7ab-8793-4e2a-bd34-774b693a24dc	Женский	2	2	1	Иногородный	Бюджет	3
4831659c-56f2-49fd-9d24-e151f3cb88db	Мужской	3	4	2	Иногородный	Бюджет	3
915c0f4c-3740-4547-a984-b66836a2439a	Женский	2	2	5	Иногородный	Бюджет	3
98478dfd-31c8-458f-8c47-0cbc74d06bf4	Мужской	2	4	5	Иногородный	Бюджет	3
abf5c2d5-b52e-4f75-9327-b4a66d95081a	Мужской	4	3	4	Иногородный	Бюджет	5
25b2f8c2-fd30-489b-ab40-cb4226f12c19	Мужской	1	1	1	Иногородный	Бюджет	2
3837d3ad-4842-4cd6-8e9c-e6fca609bd26	Мужской	4	3	5	Иногородный	Бюджет	3
5048d615-377f-4152-a9b4-fdfe3a71d4b1	Мужской	3	3	2	Иногородный	Бюджет	3



# ПРЕДСКАЗАТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

## Отчисления студентов за 1–2 семестр по зачетам

Id студента	Пол	1 дециль оценки за ЕГЭ	2 дециль оценки за ЕГЭ	3 дециль оценки за ЕГЭ	Город регистрации	Форма финансирования обучения	Дециль суммарного рейтинга за 1 семестр	Дециль медианного рейтинга по зачетам за 1 семестр	Дециль суммарного рейтинга за 2 семестр	Дециль медианного рейтинга по зачетам за 2 семестр	Предсказанный результат
0b22c5e8-dcaf-4ab7-a11a-35f88004e09a	Мужской	1	1	1	Иногородний	Бюджет	3	4	3	3	Будет отчислен
107dff39-cc02-45cc-baaa-798da48ac31b	Мужской	5	5	5	Иногородний	Бюджет	3	5	3	3	Не будет отчислен
1e85052e-d031-45b5-bb89-bea778599759	Мужской	2	2	1	Иногородний	Бюджет	3	4	3	3	Не будет отчислен
2d32a7ab-8793-4e2a-bd34-774b693a24dc	Женский	2	2	1	Иногородний	Бюджет	2	3	3	2	Будет отчислен
4831659c-56f2-49fd-9d24-e151f3cb88db	Мужской	3	4	2	Иногородний	Бюджет	3	5	3	3	Не будет отчислен
915c0f4c-3740-4547-a984-b66836a2439a	Женский	2	2	5	Иногородний	Бюджет	3	5	3	3	Не будет отчислен
98478dfd-31c8-458f-8c47-0cbc74d06bf4	Мужской	2	4	5	Иногородний	Бюджет	3	5	3	3	Не будет отчислен
abf5c2d5-b52e-4f75-9327-	Мужской	4	3	4	Иногородний	Бюджет	2	3	3	2	Не будет отчислен

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ

---

1. Проведен анализ предметной области и на его основе разработаны модели предсказания
2. Разработан алгоритм получения данных с применением существующих методов предобработки
3. Разработано приложение, выполняющее предсказание на основе моделей с помощью различных методов анализа данных
4. Проведены эксперименты, исследующие точность разработанных моделей данных