

Прогнозирование оценки пользователя мобильного приложения

Кейс №2: Пенсионный фонд
Команда №8



Команда проектного практикума



Смирнов Максим Григорьевич

- Полный цикл предобработки данных и EDA
- Формулировка и проверка статистических гипотез
- Архитектура репозитория и версионирование



Никулин Кирилл Андреевич

- Визуализация данных и поддержка EDA
- Разработка и развертывание интерактивного дашборда на Gradio



Ожерельева Ольга Игоревна

- Бизнес-интерпретация результатов
- Гипотезы для продуктовых решений
- Подготовка презентации и представление команды на защита проекта

О проекте



Компания

Пенсионный фонд, предоставляющий клиентам цифровые сервисы для управления пенсионными накоплениями

Продукт и проблема продукта

Мобильное приложение фонда, которое позволяет пользователям получать персональную информацию, оформлять услуги и отслеживать пенсионные выплаты.

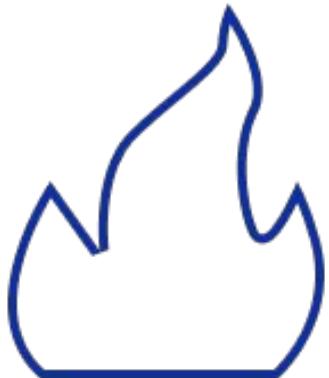
Часть пользователей недовольны приложением и ставят низкие оценки, оставляют жалобы.

Цели проекта

1. Проанализировать факторы, влияющие на удовлетворённость пользователей (CSAT)
2. Подготовить данные для построения прогнозной модели оценки
3. Дать рекомендации для бизнеса на основании анализа

Описание данных

Датасет содержит 42 387 записей с 9 признаками: идентификатор пользователя, даты создания и ответа на опрос, язык, возрастная группа, пол, стаж, доход и целевая переменная - оценка от 1 до 5 (csat_level).



В ходе первичного анализа данных было выявлено
несколько проблем:

1. Пропуски в данных о доходе пользователей, о языке и о возрасте
2. Отрицательное SLA ответа (время ответа раньше, чем время поступления вопроса от пользователя)

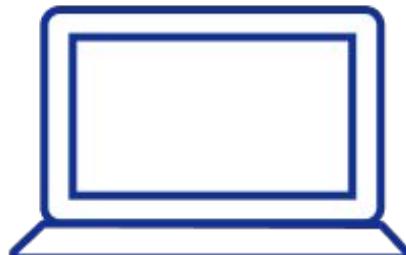
Интерактивный дашборд

Разработан дашборд на Gradio с интерактивными графиками Plotly.

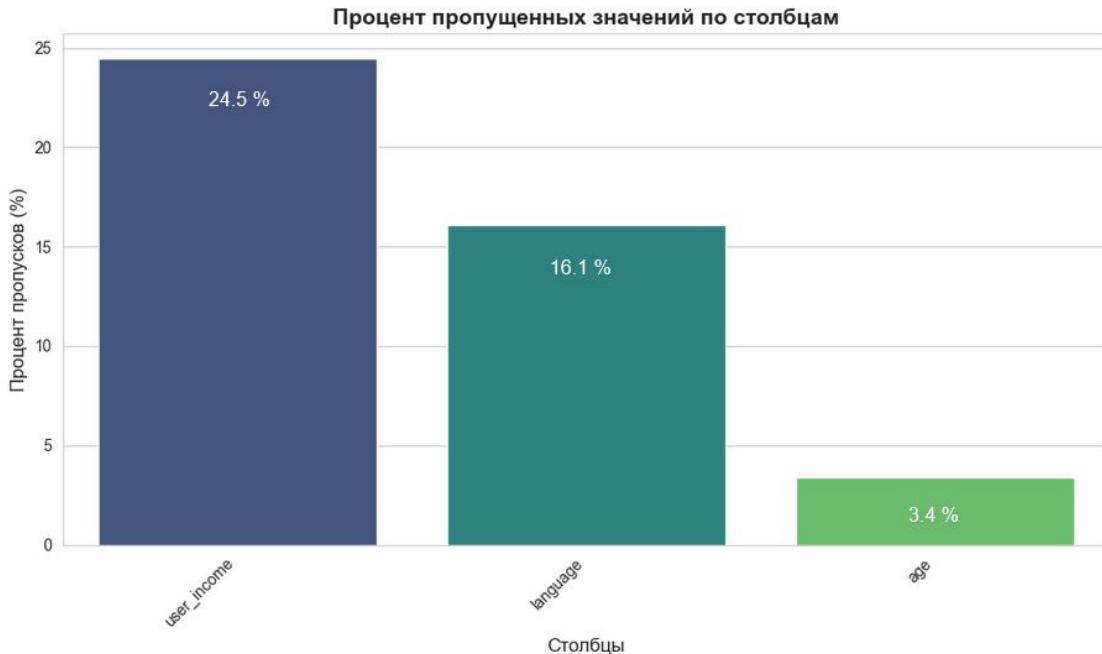
Функции:

- загрузка CSV-файла
- автоматическая очистка данных (исправление дат, заполнение пропусков)
- визуализация распределений по категориям
- анализ зависимости CSAT от возраста, пола и дохода

Дашборд доступен по [ссылке](#)



Обработка пропусков в данных



Пропуски в доходах в количестве ~24.5 % от общего количества данных решили исключить из анализа, так как их заполнение может исказить результаты.

Пропуски в языке заполнили значением RU, так как все остальные значения одинаковы

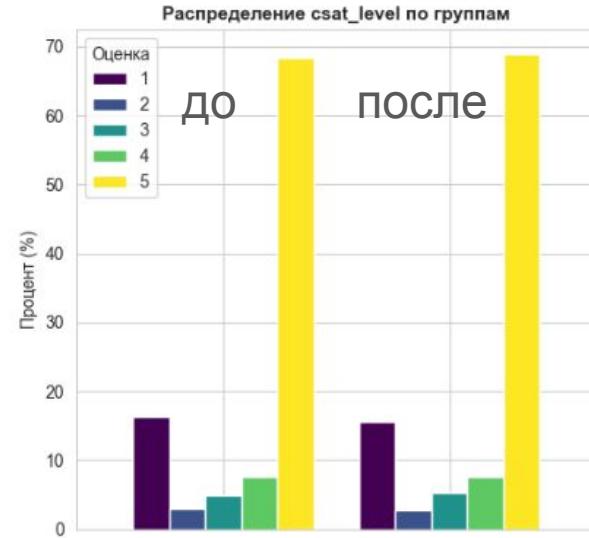
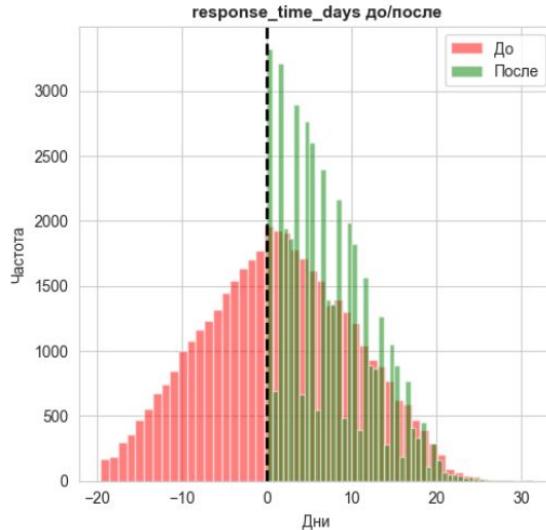
Пропуски в возрасте оставили без изменений

Обработка отрицательного SLA

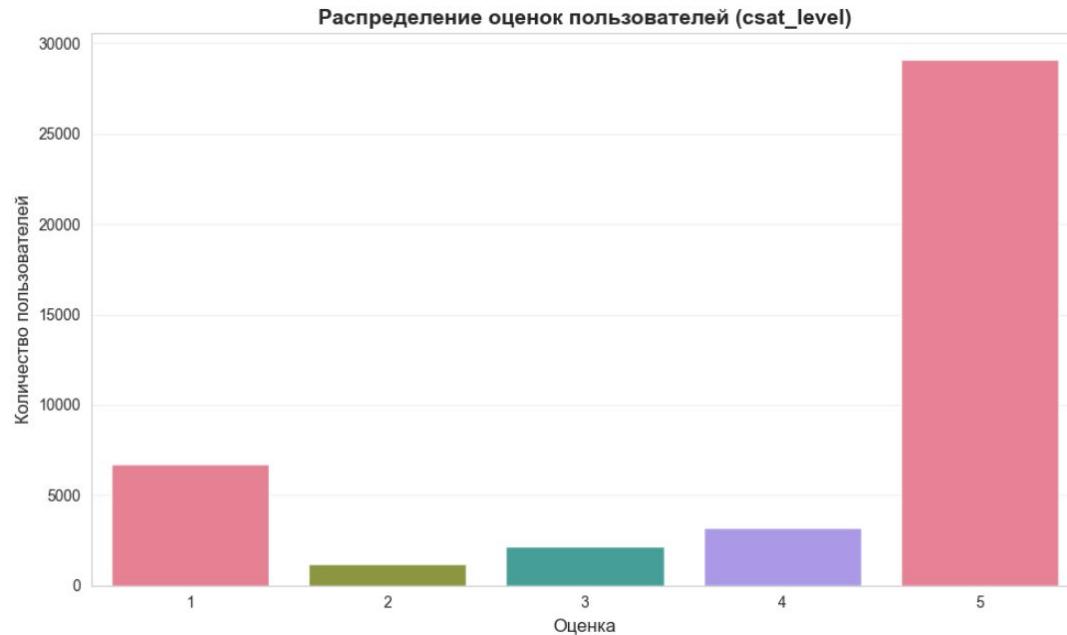
Таких строк оказалось 18400, это 43.4% данных.

Было принято решение поменять местами время ответа и время отправки вопроса.

После манипуляций поменялось общее распределение SLA, однако распределение оценок не изменилось.



Средняя оценка пользователей — 4.1 из 5



Однако самые негативные
оценки ставят чаще
средних



Направляем гипотезы на
поиск драйверов негатива

Гипотезы

1. Пользователи с высоким доходом ставят более низкие оценки
2. Старшие возрастные группы более лояльны
3. Время суток влияет на оценку
4. Старшие возрастные группы с высоким доходом оценивают приложение ниже



Гипотеза №1: Высокий доход → Низкая оценка



- Доля негативных оценок растёт с 13% до 19% с уровнем дохода
- Доля максимальных оценок падает с 72% до 62% с уровнем дохода
- Слабая отрицательная корреляция -0.076

Для подтверждения использовался тест ранговой корреляции Спирмена.

✓ p-value < 0.05, статистически значима

Гипотеза №2: Старше возраст → Выше оценка

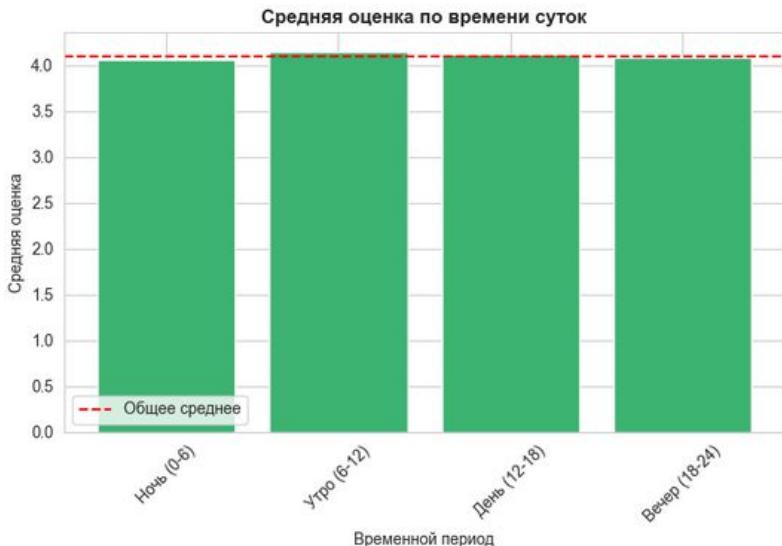
Возраст	Кол-во	Средняя оценка
18-24	3,640	3.914
25-34	9,669	4.028
35-44	11,835	4.145
45-54	9,483	4.148
55-65	6,312	4.172

Слабая положительная корреляция 0.046

Для подтверждения использовался
Критерий Краскела-Уоллиса.

✓ p-value < 0.05, статистически значима

Гипотеза №3: Утром оценки лучше, чем вечером

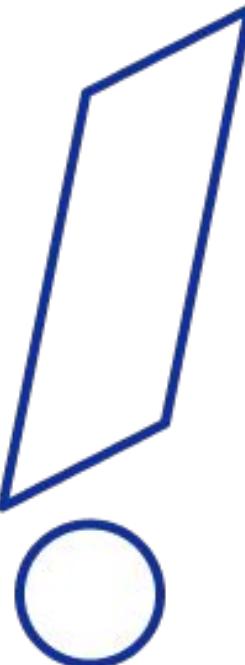


- Утро (6-12): средняя оценка 4.197
- Вечер (18-0): средняя оценка 3.990
- Разница 0.207 - самая большая среди временных паттернов

Для подтверждения использовался Критерий Краскела-Уоллиса.

✓ $p < 0.05$, статистически значима

Гипотеза №4: Эффект взаимодействия возраста и дохода



Гипотеза с самым высоким приоритетом!

Старшие (35-64 года) + высокий доход (3748 человек) - это ключевая целевая аудитория пенсионного фонда

Для подтверждения использовался Критерий Краскела-Уоллиса.

✓ $p < 0.05$, статистически значима

Сегмент	Средняя оценка
Старшие + Низкий/средний доход	4,30
Молодые + Любой доход	4,08
Старшие + Высокий доход	3,88

Ключевые выводы

Сегмент пользователей **35-64 года с высоким доходом** (3748 человек) - это **ключевая целевая аудитория** пенсионного фонда:

- Они активно планируют пенсию
- Имеют деньги для допродаж (НПФ, инвестиции)
- Но они как раз и самые недовольные (3,88 vs 4,30 у самых лояльных)

Разрыв в 0,41 балла - это критическая проблема.

Почему так может быть:

- Высокие ожидания от продукта
- Используют сложные сценарии (расчёт пенсии, прогнозирование)
- Сравнивают с современными банковскими приложениями
- Ценят своё время, а приложение не оптимизировано

Рекомендации для бизнеса

**Для ключевого сегмента
пользователей 35-64 года с
высоким доходом**

- Провести интервью с ключевым сегментом и проанализировать его пользовательский путь - где у пользователей возникают проблемы?
- Приоритизировать разработку новых фичей под их потребности
- Рассмотреть "премиум" опыт (быстрые пути, персональная поддержка) - всё, что может сэкономить их время и улучшить опыт взаимодействия с приложением.

Рекомендации для бизнеса

**Дополнительно, по
результатам EDA и проверки
гипотез**

- Упростить базовые сценарии, не перегружать функциями, чтобы пожилые смогли самостоятельно разобраться в приложении, не привлекая молодёжь.
- В выборке большой дисбаланс женщин и мужчин, возможно, женщины не видят ценности продукта. Провести UX-исследования.
- Отправлять опросы в утренние часы (6-12 часов), в утренние часы оценивают в среднем чуть выше и, возможно, объективнее.

Рекомендации для ML

По результатам
предобработки данных и
EDA

- Бинарная классификация «негатив vs позитив»
- Используем комбинированный признак «возраст × доход»
- Балансируем классы и фокусируемся на негативных оценках (жертвуем общей точностью ради обнаружения редких, но важных случаев)
- Метрики: F1-macro (баланс по всем классам) + Recall негативных оценок (не пропустить недовольного клиента важнее, чем не побеспокоить довольного)
- Рекомендуемые модели: **CatBoost / LightGBM** (работают с категориями)

Спасибо за внимание!

