

Описание проекта

Исследование надёжности заёмщиков

Что такое проект

Поздравляем! Вы прошли курс в тренажёре. Самое время проверить полученные знания на практике и выполнить проект. Проект — это настоящий аналитический кейс, решать который вы будете самостоятельно.

Как закончите работу над проектом, отправьте его на проверку наставнику. В течение суток он пришлёт комментарии. Их нужно учесть; доработать проект и вернуть обновлённый вариант наставнику.

Скорее всего, вы опять получите комментарии по обновлённому проекту. Это нормально. Процесс доработки может проходить в несколько циклов.

Проект завершён, когда наставник засчитает все доработки.

Описание проекта

Вам предстоит подготовить проект для кредитного отдела банка. Нужно разобраться, влияет ли семейное положение и количество детей клиента на факт возврата кредита в срок. Входные данные от банка — статистика о платёжеспособности клиентов.

Результаты проекта будут учтены при построении модели **кредитного скоринга** — специальной системы, которая оценивает способность потенциального заёмщика вернуть кредит банку.

Инструкция по выполнению

Задание 1. Проведите исследование

Шаг 1. Откройте файл с данными и изучите общую информацию.

Шаг 2. Предобработайте данные:

- определите и заполните пропущенные значения;
- замените вещественный тип данных на целочисленный;
- удалите дубликаты;
- выделите леммы в значениях столбца с целями получения кредита;
- категоризируйте данные.

Выполняя **Задание 1**, описывайте свои действия в ячейках типа *markdown*.

Поясните:

- какие пропущенные значения вы обнаружили;
- возможные причины появления пропусков в данных;
- по какому принципу заполнены пропуски;
- какой метод вы использовали для поиска и удаления дубликатов и почему;
- возможные причины появления дубликатов в данных;
- какой метод вы использовали для изменения типа данных и почему;
- опишите, как вы проводили лемматизацию целей кредита;
- какие "словари" вы выделили для этого набора данных и почему;

Оформление:

Выполните **Задание 1** в *Jupyter Notebook*. Эта рабочая тетрадь открывается при нажатии кнопки «Далее».

Задание 2. Опишите свои выводы

Шаг 1. Ответьте на вопросы:

- Есть ли зависимость между наличием детей и возвратом кредита в срок?
- Есть ли зависимость между семейным положением и возвратом кредита в срок?
- Есть ли зависимость между уровнем дохода и возвратом кредита в срок?
- Как разные цели кредита влияют на его возврат в срок?

Ответы сопроводите интерпретацией: поясните, о чём именно говорит полученный вами результат.

Шаг 2. Напишите общий вывод.

Оформление:

Задание 2 представьте в виде связного текста в конце вашего *Jupyter Notebook* проекта, в ячейке типа *markdown*. Примените форматирование и заголовки.

Задание 3. Отправьте проект наставнику

Прежде чем отправить проект наставнику, убедитесь, что вы выполнили все задания и написали необходимые пояснения.

Проекты без структуры и выполненных заданий наставники вернут на доработку.

Описание данных

- *children* — количество детей в семье
- *days_employed* — трудовой стаж в днях
- *dob_days* — возраст клиента в годах
- *education* — образование клиента
- *education_id* — идентификатор образования
- *family_status* — семейное положение
- *family_status_id* — идентификатор семейного положения
- *gender* — пол клиента
- *income_type* — тип занятости
- *debt* — имел ли задолженность по возврату кредитов
- *total_income* — доход в месяц
- *purpose* — цель получения кредита

Как будут проверять мой проект?

При проверке работ наставник обращается к критериям оценки. Изучите их прежде, чем приступить к решению. Они станут ориентиром для вас.

На что обращают внимание наставники при проверке проектов:

- Как вы описываете найденные в данных проблемы?
- Какие методы замены типов данных, обработки пропусков и дубликатов применяете?
- Умеете лемматизировать?

- Категоризируете данные? Почему именно таким образом?
- Выводите финальные данные в сводных таблицах?
- Применяете конструкцию *try-except* для обработки потенциальных ошибок?
- Соблюдаете структуру проекта и поддерживаете аккуратность кода?
- Какие выводы делаете?
- Оставляете комментарии к шагам?

Знания и навыки из прошлых уроков пригодятся для успешного выполнения проекта.

Держите под рукой шпаргалки и конспекты.

Успехов!