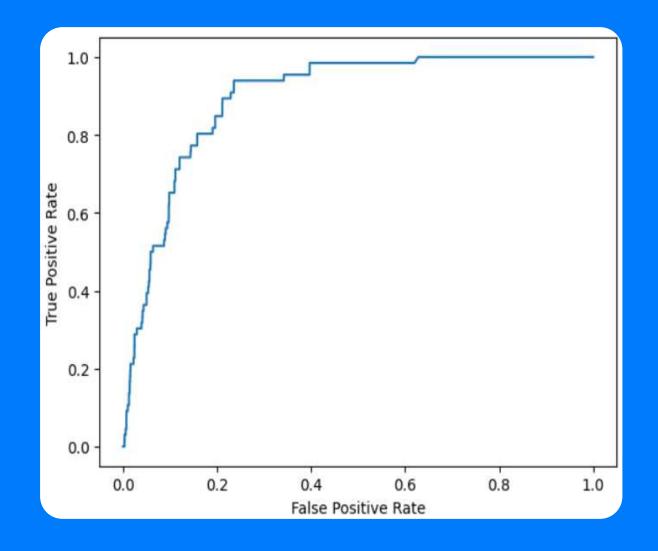
Кейс "Модель склонности клиента к приобретению машиноместа". Постановщик задачи Компания «Самолет»



Цель

Сделать модель которая будет стараться высокоточно искать клиентов которые более склонны к приобретению машиноместа



Подход к решению бизнес задачи

Анализ данных и Технического Задания

Обработка выделенных данных

Обучение регрессивной/классификационной модели

Проверка модели метрикой ROC-AUC

Анализ данных

Изучение структуры данных

Оценка качества данных

Проверка на наличие пропусков, дубликатов и выбросов

Анализ корректности данных

Анализ баланса классов

Range: 14456 entries

Columns: 2666 entries

types: float64(2316),int64(3), object(347)



Анализ технического задания

Цели проекта:

Создание модели для поиска клиентов, наиболее склонных к приобретению машиноместа.

Требования от модели:

Тип задачи(Классификация) Метрики качества (ROC-AUC) Диапазон AUC: от 0.7 до 0.9



Обработка выделенных данных

Предобработка данных

Удаление данных с типом object

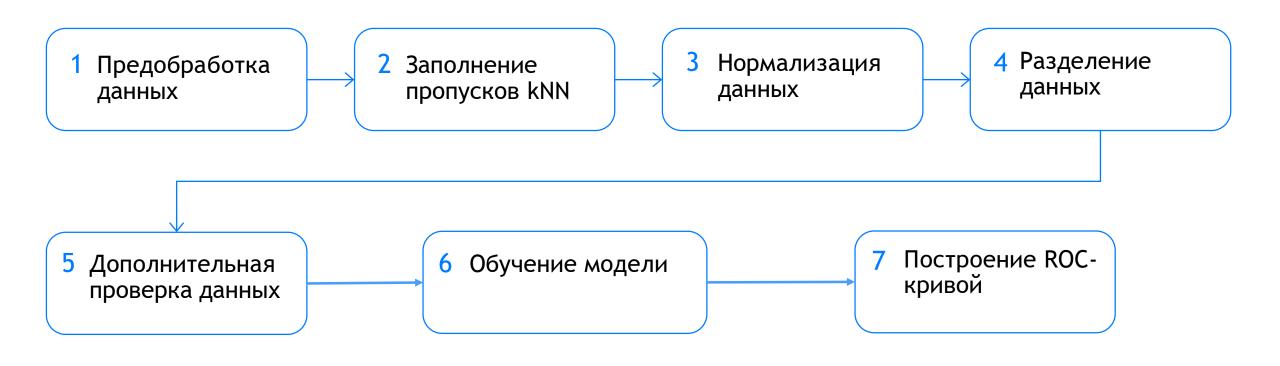
Удаление колонок, содержавших менее 20% Заполнение пропусков с помощью метода KNN

Merge колонок

Разбиение Dataset на две части Объединяем две части Dataset с помощью метода concat

Восстановление порядка колонок

Обучение модели Градиентный бустниг





Обучение LogisticRegression модели

Заполнение данных при помощи kNN

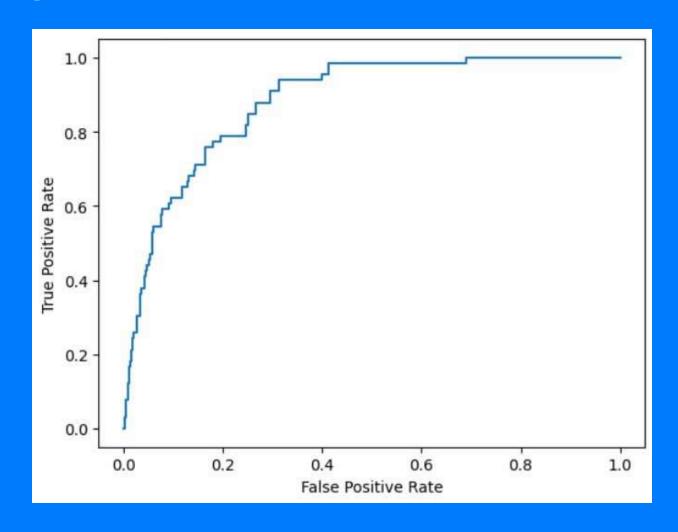
Построение метрики ROC-AUC

Обучение логистическ и- регрессионн ой модели

Проверка модели метрикой ROC-AUC

- Предсказание вероятностей
- Построение ROC кривой

AUC: 0.883

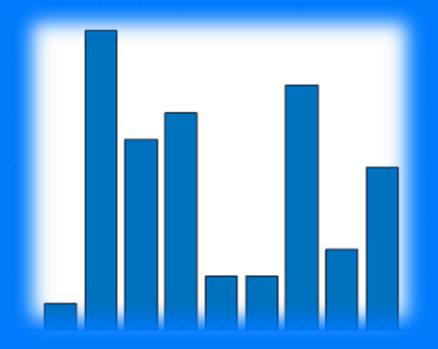


Подробно изложенный процессмоделирования

Сохранен ие модели Оценка модели Обучение модели Создание Подготовк настройка а данных модели

Подготовка данных





Создание и настройка модели





Обучение модели



Оценка модели



Сохранение модели



Мы получили обученную классификационную модель, которая выдаёт высокую точность при угадывании положительного ответа