KBTU RISK MANAGEMENT CASE COMPETITION 2024

Задание кейса

* Разработка модели оценки кредитоспособности физических лиц с использованием алгоритмов машинного обучения на основе предоставленных данных
* Аргументация подходов к разработке модели: какие факторы оказывают наиболее существенное влияние на оценку кредитоспособности и прогнозы модели, точность прогнозов и стабильность модели
* Разработка интерфейса для использования модели, позволяющего вводить данные и получать прогнозы

Набор данных содержит 24 различных параметра по 1 300 000 заемщикам, в том числе по следующим направлениям:

* сумма задолженности
* история использования кредитов
* возраст, пол
* количество дней просрочки
* количество закрытых кредитов
* сумма и иные характеристики платежей
* количество финансовых организаций, предоставивших кредиты

train.csv – 1 000 000 заемщиков

test.csv – 300 000 заемщиков

Ваша задача — спрогнозировать для каждого заемщика (customer\_ID) вероятность дефолта. Дефолт в наборе данных обозначен как target, где значение 1 присвоено дефолтным заемщикам.

Модели для оценки вероятности дефолта разрабатываются на обучающейся выборки (train.csv), а точность организаторами будет оцениваться на тестовой выборке (test.csv) метрикой AUC-PR (Precision-Recall Area Under the Curve).

Результаты необходимо сдавать в виде submission.csv с полями customer\_id и PD (спрогнозированная вероятность дефолта). На первом этапе необходимо сдать submission.csv и файл в формате PDF с аргументацией подходов к разработке модели (какие факторы оказывают наиболее существенное влияние на оценку кредитоспособности и прогнозы модели, точность прогнозов и стабильность модели).

По итогам первого этапа будет сформирован список команд, которые проходят во второй этап. На втором этапе командам необходимо разработать интерфейс для использования модели, позволяющего вводить данные и получать оценку вероятности дефолта.

Предоставленные данные не должны передаваться третьим сторонам и не должны обсуждаться между командами.

Ссылка на Данные:

https://drive.google.com/drive/folders/1650YKxQTeMa4fo85lMTfrS\_tCn4XD9gD?usp=sharing