**Обзор проделанной работы**

**Цель работы**: исследовать различные методы отбора признаков. Выбрать признаки, которые потенциально могут улучшить существующую модель. Добавить отобранные признаки в существующую модель и сравнить качество с исходной моделью.

**Методы отбора признаков**

Были рассмотрены следующие методы:

1. Univariate selection
2. Backward elimination
3. Forward selection
4. Lasso classifier
5. PCA
6. Feature importance используя XgBoost
7. Feature importance используя Catboost

Из этих методов были выбраны лучшие по соотношению метрика roc\_auc / кол-во признаков.

Более подробно ознакомиться с результатами исследований вы можете в файле **comparison\_methods.xlsx.** В результате, лучшими оказались два метода — **Backward elimination** и **Forward selection.**

**Добавление признаков в существующую модель**

В работе рассматривалось две модели — предсказание дефолта клиента (**pd**) и грейсовости (**grace**). Оба выбранных метода проявили себя схожем образом. В случае pd добавление признаков только ухудшило качество модели. В случае grace добавление признаков наоборот улучшило модель. **Logloss** на первоначальной модели – **0.58**, на обновленной – **0.52**. Более подробно ознакомиться с результатами вы можете в таблице **comparison\_tinkoff.xlsx**.

**Польза для бизнеса**

Благодаря улучшению качества модели на грейсовость клиентов, банк сможет лучше выявлять таких клиентов и предлагать им другие условия, чтобы обезопасить себя от чрезмерного использования клиентами программ лояльности, акциями, скидками и т.д.

**Заключение**

Были протестированы различные методы отбора признаков. Лучше всего себя проявили **stepwise** методы – **Backward elimination** и **Forward selection**. Они основаны на последовательном добавление/удалении признаков. Логичным продолжение было бы применение метода **Recursive Feature Elimination**, который схож с использованными, но требует больших вычислительных возможностей. Протестировать его в полной мере на персональном компьютере, к сожалению, не удалось.

.