Руководителю структурного подразделения

«Отдел розничных рисков»

**АНАЛИТИЧЕСКАЯ ЗАПИСКА**

«16» января 2025 г

**О проверки гипотезы о зависимости ПВ и статуса просрочки**

В результате проведения анализа по проверке гипотезы была обнаружена зависимость между первоначальным взносом и вероятностью просрочки кредита. Однако, стоит заметить, что такая зависимость является слабой, и составляет всего -0.3. Это говорит о том, что такая зависимость является обратной.

Вследствие чего нельзя однозначно не брать во внимание варианты, для которых значение ПВ ниже порогового, или оно больше.

Для выбора было взято две модели классификации – логистическая регрессия и классификация случайного леса. Модели сравнивались по метрике roc\_auc и показали значения 0.69 и 0.55 соответственно.

После этого была выбрана логистическая регрессия. Для нее был построен график ROC кривой, который показал хорошее значение – 0.71.

Был построен график кривой Precision – Recall. Наблюдается высокий Precision при низких значениях Recall, что может быть полезно для сценариев, где важно минимизировать количество дефолтных заявок. Однако при увеличении Recall Precision значительно падает.