

Итоговая аттестация

- Система кредитного scoring
- Вариант В
- Датасет: Give Me Some Credit
- Калашников Алексей Сергеевич 2025

Актуальность и цель

- Высокие риски при кредитовании
- Необходимость автоматизации решений
- Цель: оценка вероятности дефолта заемщика

Постановка задачи

- Задача двоичной классификации
- Целевая переменная: `SeriousDlqin2yrs`
- Вход: финансовые и анкетные данные

Требования к системе

- ROC-AUC ≥ 0.75
- Интерпретируемость
- Недискриминационность
- Время расчета ≤ 200 мс

Архитектура решения

- Хранение данных
- Подготовка данных
- Модель прогнозирования
- Сервис принятия решений

Организация данных

- Исходные данные
- Очищенные данные
- Витрина признаков

Подготовка данных

- Обработка пропусков
- Ограничение выбросов
- Производные признаки
- Разбиение на выборки

Используемые модели

- Random Forest, Gradient Boosting, SVM

Оценка качества и недискриминационности

- ROC-AUC
- Выбор порога
- Анализ по возрастным группам

Выводы

- Спроектирована система скоринга
- Учтены требования качества и интерпретируемости
- Возможность дальнейшего развития