

ПРИМЕНЕНИЕ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ИССЛЕДОВАНИЯ КРЕДИТНОГО СКОРИНГА

Бигаева Л.А., к.ф-м.н., доцент,

Баймиев В.А., студент,

Бирский филиал УУНиТ, г.Бирск, Россия

Аннотация. В статье рассматривается применение моделей машинного обучения для задачи кредитного скоринга. Проведён сравнительный анализ классических и современных алгоритмов, описаны этапы подготовки данных и метрики оценки качества. Особое внимание уделено требованиям регуляторов к интерпретируемости моделей. Результаты работы могут быть использованы при разработке автоматизированных систем оценки кредитных рисков.

Ключевые слова: кредитный скоринг, машинное обучение, бинарная классификация, оценка кредитного риска, интерпретируемость.

Введение

Кредитный скоринг помогает банку оценить, вернёт ли заёмщик деньги. Раньше использовали простые правила («если зарплата > X, то кредит можно дать»). Они понятны, но не всегда точны.

Машинное обучение анализирует тысячи примеров и находит сложные закономерности, повышая точность прогноза [1].

Цель работы — сравнить классические и современные методы скоринга, выделить проблемы их внедрения.

Какие модели используют?

Задача скоринга — **бинарная классификация**: 0 (вернёт) или 1 (дефолт).

1. **Логистическая регрессия** — простая модель, служит **точкой отсчёта** (бенчмарком). Если сложная модель не работает заметно лучше, её внедрение нецелесообразно.
2. **Ансамбли деревьев** (Random Forest, Gradient Boosting) — точнее, находят неочевидные связи между признаками [2].
3. **Нейронные сети** — для больших и «неаккуратных» данных (тексты, транзакции) [3].

Данные и паттерны

Качество модели зависит от данных. Важный этап — поиск **паттернов** (устойчивых шаблонов поведения):

- «*Резкий рост трат*» перед заявкой → возможное финансовое напряжение;
- «*Нерегулярный доход*» → выше риск для фрилансеров;
- «*Связи с проблемными клиентами*» → возможный мошеннический паттерн.

Также данные нужно: очистить от ошибок, преобразовать текст в числа, уравновесить выборку (дефолтов мало) [1].

Как оценить модель?

- **ROC-AUC / Джини**: насколько хорошо модель разделяет «хороших» и «плохих» клиентов.

- **Precision / Recall:** баланс между ошибочным отказом и выдачей кредита ненадёжному заёмщику.

- **LogLoss:** точность предсказанной вероятности [2].

Почему важно объяснять решения?

Банк обязан объяснить отказ (требования ЦБ РФ). Поэтому «чёрные ящики» используют с осторожностью.

Применяют **объяснимый ИИ (XAI):**

- **SHAP** показывает вклад каждого признака: «Отказ из-за высокой долговой нагрузки (-40 баллов)».

- **Мониторинг дрейфа** отслеживает изменения в поведении клиентов (например, во время кризиса) [3].

Заключение

Машинное обучение повышает точность скоринга, но:

1. Простая модель может быть лучше сложной, если её легче объяснить.
2. Качество данных важнее выбора алгоритма.
3. Прозрачность — обязательное требование в банках.

Будущее скоринга — в балансе точности и объяснимости.

Литература

1. Ульянова О.Д., Золотова Е.А. Исследование применения искусственного интеллекта и моделей машинного обучения в процессе оценки кредитоспособности для повышения точности и эффективности кредитного скоринга // Актуальные научные исследования: от теории к практике : сборник материалов XLIX Международной очно-заочной научно-практической конференции : в 2-х томах. — Москва : Интернаука, 2024. — Т. 2.
2. Черцова И.В., Панкеева А.Д., Муртазаев Г.Л., Аджиева А.Ю. Вероятностные модели оценки кредитных рисков // Финансовая экономика. — 2025. — № 11. — С. 324–327. — ISSN 2072-778X.
3. Кобзенко П.Б. Применение искусственного интеллекта для оценки эффективности кредитного рейтинга в коммерческих банках // Финансовая экономика. — 2025. — № 11. — С. 64–69. — ISSN 2072-778X.

Применение моделей машинного обучения для исследования кредитного скоринга

Автор: Бигаева Л.А., Баймиев В.А.

16.05.2026 10:00 -
