

## **АНАЛИЗ ЧУВСТВИТЕЛЬНОСТИ РАЗЛИЧНЫХ СЕГМЕНТОВ ЗАЕМЩИКОВ К ПРИЗНАКАМ ДЛЯ КРЕДИТНОГО СКОРИНГА: ПОВЕДЕНЧЕСКОЙ БИОМЕТРИИ И ГЕОЛОКАЦИОННОЙ СТАБИЛЬНОСТИ**

**Деркач С. А.<sup>1</sup>**

**Научный руководитель – канд. экон. наук, доцент Белинская И. В.<sup>1</sup>**

<sup>1</sup>Университет ИТМО

E-mail: derkach\_sofya@mail.ru

### **Введение**

Цифровизация финансового сектора и рост дистанционных каналов обслуживания повышают требования к точности оценки кредитного риска. Традиционные скоринговые системы преимущественно опираются на кредитную историю, финансово-демографические характеристики и транзакционную активность, однако для новых заемщиков и заемщиц со «скудной» кредитной историей информативность этих данных ограничена. В ответ на проблему банки и исследовательское сообщество рассматривают альтернативные данные, в том числе поведенческую биометрию и геолокационные параметры, как дополнительные факторы для улучшения ранжирования кредитных заявок и снижения доли ошибочных решений [1].

Поведенческая биометрия фиксирует динамические паттерны взаимодействия пользователя с цифровым устройством: тайминги ввода, устойчивость ритма нажатий, вариативность пауз, которые характеризуют «цифровую стабильность». Геолокационная стабильность отражает предсказуемость перемещений человека, в том числе радиус активности и типичное удаление от базовой точки, выступая косвенными маркерами устойчивости. Применение таких данных сопряжено с рисками, связанными с влиянием внешних факторов, техническим шумом, возможными перекосами по сегментам и требованиями правового регулирования обработки персональных данных [2].

Цель исследования – оценить чувствительность различных сегментов заемщиков к включению агрегированных признаков поведенческой биометрии и геолокационной стабильности в рамках кредитного скоринга. Под чувствительностью понимаются изменение качества модели при добавлении альтернативных факторов к базовому набору признаков и изменение разделяющей способности модели по сегментам.

### **Основная часть**

Эмпирическая часть выполнена как вычислительный эксперимент на открытых данных. В качестве базового использован датасет на 30 тыс. наблюдений с бинарной целевой переменной «дефолт/не дефолт» и традиционными скоринговыми признаками [3]. Для формирования альтернативных факторов рассчитаны нормированные индексы из открытых наборов: характеристики динамики ввода для построения агрегатов устойчивости поведения и GPS-траектории для расчета индексов пространственного разброса и регулярности перемещений [4; 5]. Для выявления неоднородности эффекта выполнена сегментация по возрасту: 18–24, 25–34 и 35+ лет. Для сопоставимости результатов использована логистическая регрессия со стандартизацией признаков. Построены три модели: M0 – базовая, M1 – расширенная с биометрическими агрегатами, M2 – расширенная с агрегатами геолокации. Качество оценивалось по ROC-AUC и критерию Колмогорова–Смирнова на тестовой выборке по возрастным сегментам.

Результаты показывают доминирующий вклад традиционных скоринговых факторов. Базовая модель демонстрирует устойчивое качество: ROC-AUC в диапазоне 0,708–0,742 по сегментам. Добавление агрегатов поведенческой биометрии не приводит к улучшению: по всей выборке  $\Delta$ ROC-AUC близка к нулю. Геолокационные агрегаты в среднем дают минимальный эффект:  $\Delta$ ROC-AUC составляет 0,001, что соответствует сохранению качества на уровне M0. Для проверки вклада групп факторов применен перестановочный анализ, оценивающий падение качества при разрушении информации в группах признаков [6]. Перемешивание показывает, что вклад биометрической группы практически отсутствует, а геолокационная группа дает слабый, но более заметный вклад (0,006).

### **Выводы**

Полученные результаты свидетельствуют, что в текущей реализации альтернативные признаки не являются заменителем кредитной истории и не обеспечивают значимого прироста метрик качества ранжирования. Практическая ценность таких данных может проявляться в иной роли – как дополнительные сигналы для принятия решений. Поведенческая биометрия может повышать надежность взаимодействия и поддерживать выявление аномалий, снижая вероятность мошенничества, а геолокационные индексы могут применяться как контекстные маркеры для маршрутизации заявок. Использование альтернативных данных требует соблюдения правовых требований [7] и пилотного тестирования с оценкой не только ROC-AUC, но и прикладных эффектов, например доли одобрений и доли ручных проверок.

### **Литература**

1. Махмадов О. С., Шарипов Б. М. Методика оценки кредитоспособности заемщика в современных условиях // Вестник российского экономического университета имени Г. В. Плеханова. 2018. № 3. С. 72-81.
2. Enhancing credit scoring accuracy with a comprehensive evaluation of alternative data // PLoS ONE. 2024. Vol. 19, no. 5. P. 1-14. doi:10.1371/journal.pone.0303566
3. Default of Credit Card Clients [Электронный ресурс]. Режим доступа: <https://archive.ics.uci.edu/dataset/350/default+of+credit+card+clients> (дата обращения: 25.12.2025).
4. Keystroke Dynamics [Электронный ресурс]. Режим доступа: <https://www.kaggle.com/datasets/carnegiecyllab/keystroke-dynamics-benchmark-data-set> (дата обращения: 25.12.2025).
5. GPS Trajectories [Электронный ресурс]. Режим доступа: <https://archive.ics.uci.edu/dataset/354/gps+trajectories> (дата обращения: 25.12.2025).
6. Credit Scoring Approaches Guidelines [Электронный ресурс]. Режим доступа: <https://thedocs.worldbank.org/en/doc/935891585869698451> (дата обращения: 25.12.2025).
7. Федеральный закон от 27 июля 2006 г. № 152-ФЗ «О персональных данных» [Электронный ресурс]. Режим доступа: [https://www.consultant.ru/document/cons\\_doc\\_law\\_61801/](https://www.consultant.ru/document/cons_doc_law_61801/) (дата обращения: 25.12.2025).