


# Новый подход к оценке чувствительности модели кредитного скоринга

Аракелян Гарник Г.

Лектор кафедры информационных систем и информационных технологий бизнеса  
Армянский государственный экономический университет (Ереван РА)

 <https://orcid.org/0000-0001-6217-9681>  
[g.arakelyan0393@gmail.com](mailto:g.arakelyan0393@gmail.com)

УДК 336.7; EDN: TBETZH; JEL: E5, E59, H80

DOI: 10.58587/18292437-2025.4-223

**Ключевые слова и словосочетания:** кредитный скоринг, машинное обучение, энтропия, бинарная классификация, кредитный риск, кредитоспособность, метод SHAP

## Վարկային սքորինգի մոդելի զգայունության գնահատման նոր մոտեցում Առաքելյան Գարնիկ Հ.

Տեղեկատվական համակարգերի և տեղեկատվական տեխնոլոգիաների բիզնեսի ամբիոնի դասախոս,  
ՀՊՏՀ (Երևան, ՀՀ)

**Ամփոփագիր.** Բանկային համակարգը հանդիսանում է տնտեսության կարևոր հատված և այն հաճախ բախվում է տարբեր տեսակի ռիսկերի հետ:

Այս հոդվածում քննարկվում է վարկային կազմակերպությունների, այդ թվում բանկերի ֆինանսական կորուստների հիմնական աղբյուր հանդիսացող վարկային ռիսկերի մասին: Սույն հոդվածում իրական տվյալների հիման վրա վարկային ռիսկի գնահատման և օպտիմալ կառավարման նպատակով կառուցվել է սքորինգային մոդել՝ կիրառելով մեքենայական ուսուցման XGBoost ալգորիթմը: Այն վերադարձնում է սքոր միավորներ 300-850 միջակայքում: Աշխատանքում ստացված մոդելի մասով գնահատվել են ROC AUC, PR AUC, Precision, Recall, F1 score որակական ցուցանիշները:

Հետազոտության շրջանակում մշակվել է նոր մոտեցում այս կամ այն փոփոխականների արժեքների փոփոխման դեպքում մեքենայական ուսուցման մոդելի զգայունությունը գնահատելու համար: Մոտեցման հիմքում ընկած է էնտրոպիայի հասկացությունը:

Հետազոտության արդյունքները ցույց են տալիս, որ մեքենայական ուսուցման մոդելների կիրառումը թույլ է տալիս գնահատել և օպտիմալ կառավարել վարկային ռիսկը, արդյունքում նվազեցնելով առաջացող կորուստները: Իսկ առաջարկվող նոր մոտեցման կիրառումը հայտնի SHAP մեթոդի հետ թույլ է տալիս առավել խորը գնահատել մոդելը և նրա զգայունությունը փոփոխականների արժեքի փոփոխման դեպքում:

**Հանգուցաբառեր և բառակապակցություններ՝** վարկային սքորինգ, մեքենայական ուսուցում, էնտրոպիա, բինար դասակարգում, վարկային ռիսկ, վարկունակություն, SHAP մեթոդ

## A New Approach to Assessing the Sensitivity of a Credit Scoring Model

Arakelyan Garnik H.

Lecturer, Chair of Information Systems and Information Technologies in Business,  
Armenian State University of Economics (Yerevan, RA)

**Abstract.** The banking system is a key sector of the economy and often faces various types of risks. This article focuses on credit risks, which are the main source of financial losses for credit institutions, including banks. Based on real data, a scoring model was developed in this study to assess and optimally manage credit risk, using the XGBoost machine learning algorithm. The model returns score values in the range of 300–850. Quality metrics such as ROC AUC, PR AUC, Precision, Recall, and F1 score were evaluated for the developed model.

Within the scope of the research, a new approach was proposed to assess the sensitivity of a machine learning model to changes in variable values. The approach is based on the concept of entropy.

The results show that the use of machine learning models makes it possible to assess and optimally manage credit risk, thereby reducing potential losses. Moreover, the proposed new approach, when applied together with the well-known SHAP method, allows for a deeper evaluation of the model and its sensitivity to changes in variable values.

**Keywords & phrases:** credit scoring, machine learning, entropy, binary classification, credit risk, creditworthiness, SHAP method

### Введение

Кредитные учреждения осуществляют свою деятельность в нестабильной среде и не имея

полной информации о ней могут понести значительные финансовые потери. Одним из основных источников подобных потерь являются

кредитные риски. В 1997 г. Базельский комитет по банковскому надзору назвал кредитный риск основным видом финансового риска, с которым сталкиваются финансовые институты в своей деятельности. Этот факт отражает прокатившуюся по всему миру в 1980-1990-х годах волну корпоративных банкротств, ставших результатом кредитного риска [2, с. 323].

Для оценки и управления кредитными рисками кредитные организации часто применяют различные готовые математические модели, например, скоринговая модель FICO Score, которая разработана компанией Fair Isaac Services Limited [4]. В последнее время наряду с развитием технологий машинного обучения для оптимизации данного процесса кредитные учреждения создают свои собственные модели на основе имеющихся у них данных. Зачастую данные модели представляют собой “черный ящик”, и кредитные учреждения часто не могут точно объяснить результат прогноза модели. Для решения этой проблемы используются различные методы. Одним из наиболее известных методов является SHAP, который объясняет вклад каждой переменной в результат модели. Однако данный метод не позволяет оценивать чувствительность модели к изменению значения признаков.

Целью данного исследования является исследование поведения заемщиков на основе реальных данных, создание модели XGBoost для оценки кредитного риска, а также реализация нового алгоритма для оценки чувствительности модели к изменению значения той или иной переменной.

### Основная часть

В данном исследовании использовались реальные данные о кредитах «Юнибанк» ООО. В табличных данных имеются 13 переменных. Выборка содержит 4603 наблюдений. Рассмотрим независимые переменные.

1. **Семейное положение.** Категориальная переменная, которая может иметь «Женат (замужем)», «Вдовец (вдова)», «Разведен (разведена)» и «Одинокий (одинокая)» значения;

2. **Образование.** Категориальная переменная, которая может иметь «Ученая степень», «Высшее образование», «Среднее профессиональное» и «Средняя школа» значения;

3. **Наличие имущества.** Категориальная переменная, которая может иметь «Наличие

недвижимости», «Наличие движимого имущества», «Наличие недвижимости и движимого имущества» и «Отсутствие» значения;

4. **Пол.** Категориальная переменная, которая может иметь «Женский» и «Мужской» значения;

5. **Возраст заемщика.** Числовая переменная, которая содержит значения от 20 до 65;

6. **Количество дней просрочки за последние 12 месяцев.** Числовая переменная, которая содержит значения от 0 до 1600;

7. **Количество просрочек.** Числовая переменная, которая содержит значения от 1 до 67. Здесь имеются пропущенные данные. На основании полученной от банка информации стало ясно, что у данных клиентов нет просроченной задолженности. Поэтому пропущенные значения были изменены на 0;

8. **Количество изменений в классах риска.** Числовая переменная, которая содержит значения от 0 до 28;

9. **Кредитная нагрузка.** Числовая переменная, которая содержит значения от 0 до 2 984 303. Данные представлены в драмах;

10. **Длительность кредитной истории в банке.** Числовая переменная, которая содержит значения от 0 до 488. Данные представлены в днях;

11. **Максимально погашенные кредиты.** Числовая переменная, которая содержит значения от 0 до 25 131 048. Данные представлены в драмах;

12. **Сумма контракта.** Числовая переменная, которая содержит значения от 30,000 до 2,300,000. Данные представлены в драмах РА. Здесь имеются выбросы. Это клиенты, которым банк выдал кредит на сумму более 1,000,000 драмов РА. Таких наблюдений 7. Было подсчитано, что удаление данных наблюдений не оказывает существенного влияния на качество моделей. Поэтому при создании моделей эти наблюдения не удалялись из набора данных;

13. **Дефолт.** Бинарная переменная (1 - дефолт, 0 - нет дефолта). На основании полученной от банка информации стало ясно, что дефолтом считается наличие хотя бы одного раза просроченных обязательств на 90 дней и более за последний год.

В таблице 1 представлены сгруппированные реальные данные и уровень дефолта для каждой группы.

Таблица 1. Группы имеющихся данных в зависимости от наличия дефолта

Группа	Сумма (тыс. драм)	Дефолт (%)
<b>Пол</b>		
Женский	612,678	41,2%
Мужской	557,540	49,9%
<b>Образование</b>		
Среднее образование	1,339	29,4%
Ученая степень	155,561	43,6%
Высшее образование	291,127	44,0%
Среднее профессиональное	722,190	46,2%
<b>Семейное положение</b>		
Разведен (разведена)	13,085	23,5%
Женат (замужем)	913,540	43,0%
Одинокий (одинокая)	17,931	48,8%
Вдовец (вдова)	225,660	55,6%
<b>Количество просрочек</b>		
0	540,573	40,6%
1+	629,644	49,4%
<b>Кредитная нагрузка</b>		
0	317,685	28,6%
1-300,000	207,519	45,3%
300,001+	645,013	53,6%
<b>Сумма контракта</b>		
0-200,000	520,209	40,9%
200,001-400,000	418,168	42,2%
400,001+	231,840	60,8%
<b>Количество дней просрочки за последние 12 месяцев</b>		
0-30	1,155,316	44,7%
31+	14,902	91,1%
<b>Наличие имущества</b>		
Наличие недвижимого и движимого имущества	59,682	41,8%
Наличие движимого имущества	280,834	44,0%
Наличие недвижимого имущества	70,760	44,1%
Отсутствие имущества	758,940	46,2%
<b>Возраст</b>		
20-25	210,717	53,7%
26-35	350,813	47,6%
36-50	370,738	41,3%
51+	237,949	40,8%
<b>Количество изменений в классах риска</b>		
0	910,921	47,9%
1+	259,297	36,2%
<b>Длительность кредитной истории в банке (день)</b>		
0-270	1,156,563	45,4%
271-365	10,143	41,2%
366+	3,512	39,1%
<b>Максимально погашенные кредиты</b>		
0-350,000	810,315	45,1%
350,001+	359,903	45,7%

В рамках данного исследования использовались различные методы анализа, например, для выявления закономерностей и связей в данных использовался корреляционный анализ.

Корреляция показывает статистическую взаимосвязь двух или нескольких случайных величин [1].

Исследование данных и моделирование осуществлялось с применением программы

Excel, языка программирования Python и его соответствующих библиотек.

В таблице 2 представлено корреляционная матрица для числовых переменных, откуда видно, что имеется умеренная связь между переменными «Количество изменений в классах риска» и «Количество просрочек», а также «Длительность кредитной истории в банке» и «Кредитная нагрузка».

Таблица 2. Корреляционная матрица

	Сумма контракта	Возраст	Кол-во дней просрочки за последние 12 месяцев	Кол-во просрочек	Кол-во изменений в классах риска	Кред. нагрузка	Длительность кредит. истории в банке	Макс. погаш. кредиты	Дефолт
Сумма контракта	1,00	0,07	0,04	0,01	0,01	0,07	0,10	0,13	0,10
Возраст	0,07	1,00	0,01	0,05	0,08	0,14	0,21	0,12	-0,08
Кол-во дней просрочки за последние 12 месяцев	0,04	0,01	1,00	0,34	0,38	-0,01	0,09	0,14	0,08
Кол-во просрочек	0,01	0,05	0,34	1,00	0,57	0,05	0,35	0,18	0,04
Кол-во изменений в классах риска	0,01	0,08	0,38	0,57	1,00	0,00	0,22	0,12	-0,05
Кред. нагрузка	0,07	0,14	-0,01	0,05	0,00	1,00	0,50	0,17	0,15
Длительность кред. истории в банке	0,10	0,21	0,09	0,35	0,22	0,50	1,00	0,26	-0,05
Макс. погаш. кредиты	0,13	0,12	0,14	0,18	0,12	0,17	0,26	1,00	0,00
Дефолт	0,10	-0,08	0,08	0,04	-0,05	0,15	-0,05	0,00	1,00

После анализа данных была создана модель кредитного скоринга. По сути задача кредитного скоринга сводится к решению задачи бинарной классификации, то есть зависимая переменная является «Дефолт».

В качестве алгоритма машинного обучения был выбран XGBoost, на основе которого была построена модель. Данный мощный алгоритм, обладает высокой производительностью за счет регуляризации, обработки пропусков, параллельных вычислений, устойчивости к переобучению и к дисбалансу классов.

Были оценены качественные показатели полученной модели, а полученные вероятности были приведены в скор баллы в диапазоне от 300 до 850.

В целях моделирования и оценки качества построенной модели имеющиеся данные были разделены на обучающую (80% от общего объема данных) и тестовую (20% от общего объема данных) выборки. Это стандартная практика, позволяющая объективно оценить способность модели обобщать информацию на новых, ранее не виденных данных. Числовые

переменные (кроме «Количество изменений в классах риска») были стандартизованы стандартизация с применением инструмента StandardScaler из библиотеки Sklearn [6] языка программирования Python, который позволяет преобразовать данные таким образом, что они приобретают нулевое среднее значение и единичное стандартное отклонение. С применением инструмента get\_dummies из библиотеки Pandas [5] языка программирования Python все категориальные переменные были преобразованы в числовой формат, что позволяет учитывать данные переменные в модель без нарушения ее математических допущений. В этом случае для каждого уникального значения категориального признака создается отдельный бинарный столбец.

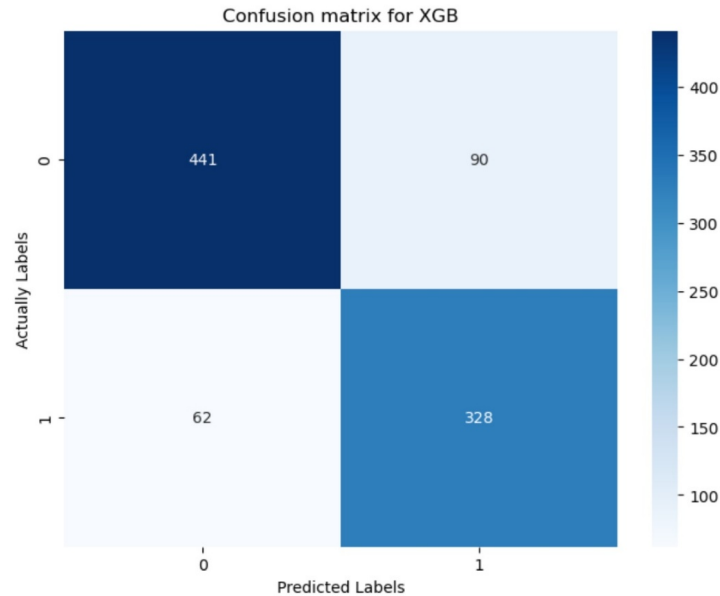
Для нахождения лучшего решения при создании модели была осуществлена оптимизация гиперпараметров. Так оптимальным значением параметра `n_estimators` является 260 (диапазон поиска: 100-300), `max_depth` является 7 (диапазон поиска: 1-7), `learning_rate` является 0.04379 (диапазон поиска:  $10^{-4}$ -1), `reg_lambda`

является 0.027651(диапазон поиска:  $10^{-4}$ - $10^4$ ), min\_child\_weight является 0.1905517 (диапазон поиска:  $10^{-4}$ - $10^4$ ), subsample является 0.75158 (диапазон поиска: 0,3-1), colsample\_bytree является 0.75167 (диапазон поиска: 0,3-1), scale\_pos\_weight является 1.15011 (диапазон поиска:  $10^{-3}$ - $10^3$ ).

В качестве наилучшего порога классификации (Threshold) было выбранно 0.41339466 значение, при котором Accuracy составляет

0,835, Precision составляет 0,785, Recall составляет 0,841, а F1-мера [3] составляет 0,812. На рисунке 1 представлена матрица ошибок [8], согласно которой модель идентифицирует 84,1% дефолтных клиентов (328/390), однако при этом модель ошибочно классифицирует 16,9% недефолтных клиентов как дефолтных (90/531). Выше указанные показатели рассчитаны для тестовых данных.

Рисунок 1. Матрица ошибок для тестовых данных



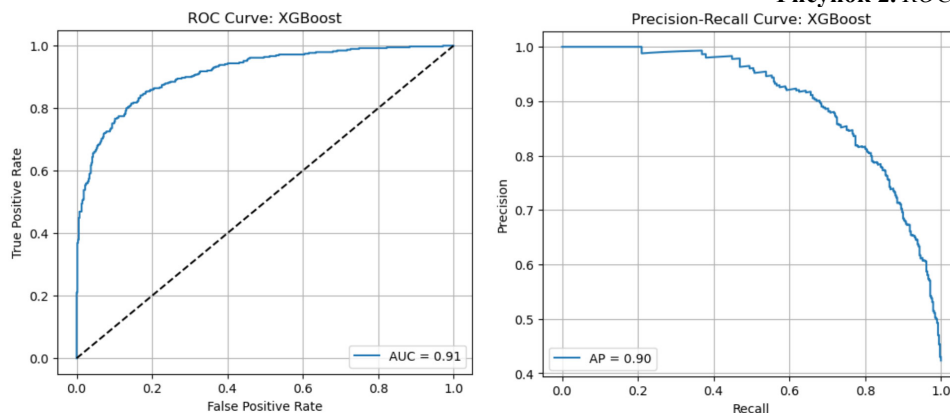
Полученные вероятности тестовых данных для класса 0 были приведены в скор баллы согласно ниже указанной формуле

$$Score = \left( \ln \left( \frac{P(y = 0)}{1 - P(y = 0)} \right) * \frac{850 - 300}{|Percentile_1| + |Percentile_{99}|} \right) + \left( 850 - \frac{850 - 300}{|Percentile_1| + |Percentile_{99}|} * Percentile_1 \right),$$

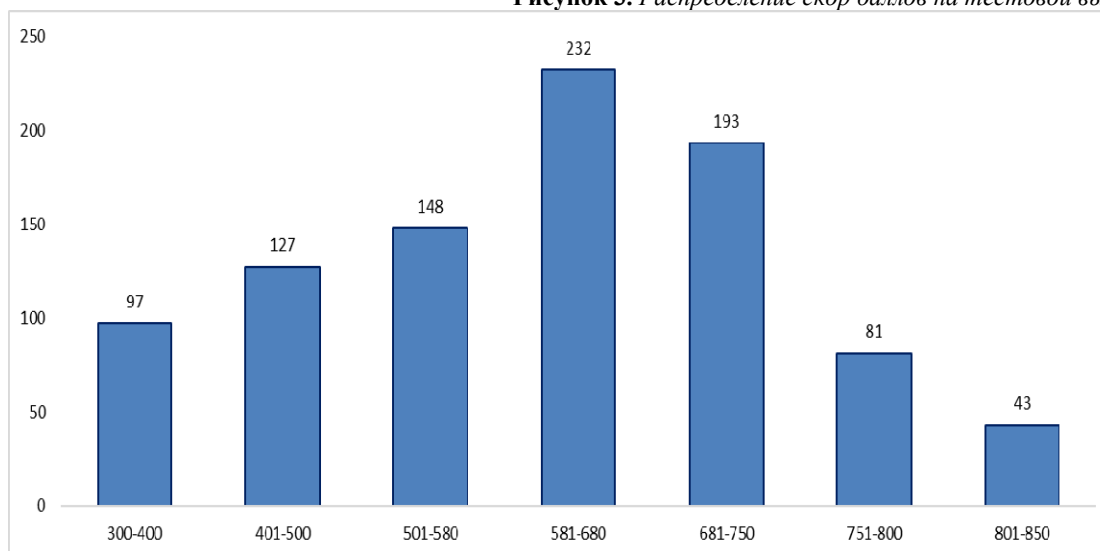
где Score – скор балл,  $P(y = 0)$  – вероятность принадлежности наблюдения к классу 0,  $Percentile_1$  и  $Percentile_{99}$  – 1-ый и 99-ый перцентиль для прологарифмированного ряда.

На рисунке 2 изображены кривые ROC и PR [7] для полученной модели, а на рисунке 3 изображено распределение полученных скор баллов для тестовой выборки.

Рисунок 2. ROC и PR кривые



**Рисунок 3.** Распределение скор баллов на тестовой выборке



Выше указанный порог классификации соответствует 608 скор баллу. В таблице 3 представлены данные тестовой выборки в зависимости от скор балла. Согласно данной таблице финансовая организация при внедрении данной модели могла бы не выдать кредитов на сумму 113,021 тыс. драм, из которых в состоянии дефолта находятся 90,395 тыс. драм.

Использование данной модели поможет снизить потери при дефолте на 85,1% (90,395 / 106,238). Это в свою очередь может привести к снижению цены риска, которая участвует в формировании процента по кредиту. Соответственно, финансовая организация при внедрении модели может снизить процент по кредитам, что делает ее конкурентоспособной на рынке.

**Таблица 3.** Разбивка портфеля в зависимости от скор балла и наличия дефолта (тестовая выборка)

Группа по скор баллу	Дефолт		Всего		Дефолт (%)	
	Объем портфеля (тыс. драм)	Количество	Объем портфеля (тыс. драм)	Количество	Объем портфеля	Количество
300-608	90,395	328	113,021	419	80,0%	78,3%
609-850	15,843	62	119,001	502	13,3%	12,4%
<b>Всего</b>	<b>106,238</b>	<b>390</b>	<b>232,022</b>	<b>921</b>	<b>45,8%</b>	<b>42,3%</b>

Выше указанная модель как уже отмечалась является “черным ящиком”, и кредитные учреждения часто не могут точно объяснить результат прогноза модели. Для решения этой проблемы используются различные методы. Одним из наиболее известных методов является SHAP, который объясняет вклад каждой переменной в результат модели. Однако данный метод не позволяет оценивать чувствительность модели к изменению значения признаков.

В результате исследования, для оценки чувствительности моделей бинарной классификации к изменению значений переменных, участвовавших в модели, был разработан соответствующий метод. Рассмотрим шаги для реализации данного метода для числовых переменных:

1. Рассматриваемая переменная меняется, например, на 1 или 1%. Изменение нужно выполнить как в направлении увеличения, так и в направлении уменьшения значения признака. Все остальные признаки неизменяются.

2. Пусть в модели зависимая переменная является вероятностью принадлежности к классу 1. Рассчитываем изменение вероятности в двух направлениях: в увеличении (+), и в уменьшении (-) значения признаков, согласно ниже указанным формулам.

$$\Delta p_+ = p_1 - p_0,$$

$$\Delta p_- = p_1 - p_0,$$

где,  $p_0$  является вероятностью принадлежности к классу 1 до осуществления изменений согласно пункту 1, а  $p_1$  является вероятностью

принадлежности к классу 1 после осуществления соответствующих изменений.

Если  $\Delta p_+ > 0$  и  $\Delta p_- > 0$ , то любое изменение в переменной приводит к увеличению вероятности принадлежности к классу 1. В данном случае переменная действует монотонно и модель чувствительна к любым изменениям.

Если  $\Delta p_+ < 0$  и  $\Delta p_- < 0$ , то любое изменение в переменной приводит к уменьшению вероятности принадлежности к классу 1. В данном случае переменная действует монотонно и модель чувствительна к любым изменениям.

Если  $\Delta p_+ > 0$  и  $\Delta p_- < 0$ , то увеличение значения переменной приводит к увеличению вероятности принадлежности к классу 1, а уменьшение – ее снижению. В данном случае модель реагирует в направлении изменения переменной.

Если  $\Delta p_+ < 0$  и  $\Delta p_- > 0$ , то увеличение значения переменной приводит к снижению вероятности принадлежности к классу 1, а уменьшение – ее увеличению. В данном случае модель реагирует в противоположном направлении изменения переменной.

Если  $\Delta p_+ > 0$  и  $\Delta p_- = 0$  или  $\Delta p_+ = 0$  и  $\Delta p_- > 0$ , то изменение переменной воздействует только в направлении увеличения вероятностей. Модель чувствительна когда значение переменной либо увеличивается, либо уменьшается.

Если  $\Delta p_+ < 0$  и  $\Delta p_- = 0$  или  $\Delta p_+ = 0$  и  $\Delta p_- < 0$ , то изменение переменной воздействует только в направлении уменьшения вероятностей. Модель чувствительна когда значение переменной либо увеличивается, либо уменьшается.

Если  $\Delta p_+ = 0$  и  $\Delta p_- = 0$ , то изменение значения признака не влияет на модель.

3. Рассчитываем изменение неопределенности в двух направлениях: в увеличении (+), и в уменьшении (-) значения признаков, согласно ниже указанным формулам.

$$H_0 = -p_0 * \ln(p_0) - (1 - p_0) * \ln(1 - p_0),$$

$$H_1 = -p_1 * \ln(p_1) - (1 - p_1) * \ln(1 - p_1),$$

$$\Delta H_+ = H_1 - H_0,$$

$$\Delta H_- = H_1 - H_0.$$

Где  $H_0$  является энтропией до осуществления изменений согласно пункту 1, а  $H_1$  является энтропией после осуществлении соответствующих изменений.

Если  $\Delta H_+ > 0$  и  $\Delta H_- > 0$ , то при любом изменении в переменной приводит к снижению уверенности модели в прогнозе.

Если  $\Delta H_+ < 0$  и  $\Delta H_- < 0$ , то при любом изменении в переменной приводит к увеличению уверенности модели в прогнозе.

Если  $\Delta H_+ > 0$  и  $\Delta H_- < 0$ , то увеличение значения переменной приводит к снижению уверенности модели в прогнозе, а уменьшение – ее увеличению.

Если  $\Delta H_+ < 0$  и  $\Delta H_- > 0$ , то увеличение значения переменной приводит к увеличению уверенности модели в прогнозе, а уменьшение – ее уменьшению.

Если  $\Delta H_+ = 0$  и  $\Delta H_- > 0$  или  $\Delta H_- < 0$ , то увеличение значения переменной не влияет на уверенность модели в прогнозе, а уменьшение значения переменной приводит либо к снижению, либо к увеличению уверенности модели в прогнозе.

Если  $\Delta H_- = 0$  и  $\Delta H_+ > 0$  или  $\Delta H_+ < 0$ , то уменьшение значения переменной не влияет на уверенность модели в прогнозе, а увеличение значения переменной приводит либо к снижению, либо к увеличению уверенности модели в прогнозе.

Если  $\Delta H_+ = 0$  и  $\Delta H_- = 0$ , то изменение значения признака не влияет на уверенность модели в прогнозе.

4. Объединяем результаты 2-го и 3-го пунктов согласно нижеуказанным формулам

$$Inf_- = \Delta p_- * (-\Delta H_-),$$

$$Inf_+ = \Delta p_+ * (-\Delta H_+),$$

$$Inf = \frac{|Inf_+| + |Inf_-|}{2},$$

Где  $Inf$  – итоговая оценка, которая является положительной мерой влияния признаков и чем больше, тем значительнее эффект. Для одного признака ее интерпретация будет правильной в сравнении с аналогичным показателем других признаков.

Рассмотрим предложенный алгоритм для категориальных переменных.

1. Категориальные переменные закодируются числами.

2. Пусть в модели зависимая переменная является вероятностью принадлежности к классу 1. Рассчитываем изменение вероятности в случае принятия всех возможных значений данной категориальной переменной (за исключением уже имеющегося значения) в соответствии с приведенной ниже формулой.

$$\Delta p_i = p_1 - p_0,$$

Где  $p_0$  - это вероятность принадлежности к классу 1 до преобразования указанной

переменной, а  $p_i$  - вероятность в случае  $i$ -го возможного значения данной переменной.

Если  $\Delta p_i > 0$ , то изменение значения переменной приводит к увеличению вероятности принадлежности к классу 1.

Если  $\Delta p_i < 0$ , то изменение значения переменной приводит к снижению вероятности принадлежности к классу 1.

Если  $\Delta p_i = 0$ , то изменение значения переменной не влияет на изменение вероятностей принадлежности к той или иной группе.

3. Вычисляем изменение неопределённости модели в случае принятия всех возможных значений данной категориальной переменной (за исключением уже имеющегося значения) в соответствии с приведёнными ниже формулами

$$H_0 = -p_0 * \ln(p_0) - (1 - p_0) * \ln(1 - p_0),$$

$$H_i = -p_i * \ln(p_i) - (1 - p_i) * \ln(1 - p_i),$$

$$\Delta H_i = H_i - H_0,$$

где  $H_0$  представляет собой энтропию до преобразования переменной, а  $H_i$  - энтропию в случае  $i$ -го возможного значения данной переменной.

Если  $\Delta H_i > 0$ , то изменение переменной приводит к увеличению неопределённости.

Если  $\Delta H_i < 0$ , то изменение переменной приводит к уменьшению неопределённости.

Если  $\Delta H_i = 0$ , то изменение переменной не влияет на неопределённость.

4. Объединяем результаты 2-го и 3-го пунктов согласно нижеуказанным формулам

$$Inf_i = \Delta p_i * (-\Delta H_i),$$

$$Inf = \frac{1}{n} \sum_{i=1}^n |Inf_i|,$$

Где  $Inf_i$  представляет собой оценка влияние переменной при  $i$ -ом значении,  $Inf$  – итоговая оценка, которая является положительной мерой влияния признаков и чем больше, тем значительнее эффект. Для одного признака ее интерпретация будет правильной в сравнении с аналогичным показателем других признаков.

По выше указанной методологии можно оценить влияние изменения признаков для отдельного наблюдения. На уровне выборки можно осуществить расчет по всем наблюдениям и затем усреднить оценку и получить оценку на уровне модели.

В рамках данного исследования на основе выше указанного метода были рассчитаны показатели чувствительности модели к изменению уже стандартизированных числовых переменных на 1 значение. В таблице 4 для числовых переменных представлены рассчитанные усредненные показатели на уровне выборки по указанному методу и среднее абсолютное значение SHAP по каждому признаку.

Таблица 4. Чувствительность модели к изменению числовых переменных в сравнении с методом SHAP

Переменная	$\Delta p_+$	$\Delta H_+$	$Inf_+$	$\Delta p_-$	$\Delta H_-$	$Inf_-$	$Inf$	SHAP
Возраст	0,002335	0,011536	0,011415	0,010437	0,013862	0,013081	0,012248	0,225703
Количество дней просрочки за последние 12 месяцев	-0,105338	-0,037047	0,085747	0,022189	0,024548	0,024818	0,055283	0,451755
Количество просрочек	0,115637	0,055751	0,032288	-0,015243	0,019940	0,008769	0,020528	0,407233
Количество изменений в классах риска	-0,017670	0,000957	0,037618	0,031453	0,017346	0,011449	0,024534	0,409482
Кредитная нагрузка	0,146369	0,053286	0,044006	-0,009588	0,036175	0,025269	0,034637	0,682995
Длительность кредитной истории в банке (день)	-0,070502	0,008194	0,030372	-0,111887	-0,063505	0,073839	0,052106	0,818025
Максимально погашенные кредиты	0,083158	0,005766	0,027394	-0,000238	0,067402	0,026488	0,026941	0,409332

Из таблицы 4 можно заметить, что наибольшую ценность имеют

1. Увеличение значения признака “Количество дней просрочки за последние 12 месяцев” приводит к уменьшению вероятности принадлежности к классу 1 и увеличению уверенности модели в прогнозе, а уменьшение значения признака приводит к увеличению вероятности принадлежности к классу 1 и уменьшению уверенности модели в прогнозе. Данное явление может быть связано с низким уровнем корреляции между рассматриваемым признаком и дефолтом, а также может быть следствием нелинейных зависимостей в модели. Рассматривая расчетные значения важности признака можно сделать вывод, что как для SHAP, так и для разработанного метода признак считается важным. Заметим, что SHAP относит его к более средне-высоким по важности, а разработанный метод чуть выше, чем важность у других признаков.

2. Увеличение значения признака “Кредитная нагрузка” приводит к увеличению веро-

ятности принадлежности к классу 1, а уменьшение – к уменьшению вероятности принадлежности к классу 1. Любое изменение приводит к уменьшению уверенности модели в прогнозе. Как для разработанного метода, так и для SHAP показатель является важным.

3. Изменение значения признака “Длительность кредитной истории в банке (день)” приводит к уменьшению вероятности принадлежности к классу 1. Увеличение снижает уверенность модели в прогнозе, а уменьшение значения признака приводит к повышению уверенности в прогнозе. Для SHAP показатель является первым по важности среди числовых переменных, а для разработанного метода признак по важности находится на втором месте.

В таблице 5 для категориальных переменных представлены рассчитанные урденные показатели на уровне выборки по указанному методу и среднее абсолютное значение SHAP по каждому признаку.

**Таблица 5.** Чувствительность модели к изменению категориальных переменных в сравнении с методом SHAP

Переменная	$\Delta p$	$\Delta H$	Inf mean	SHAP
Семейное положение разведен (разведена)	-0,076070	-0,041910	0,016082	0,021532
Семейное положение женат (замужем)	0,040075	0,001930	0,012182	0,075783
Семейное положение одинокий (одинокая)	-0,024256	-0,014826	0,001959	0,009537
Семейное положение вдовец (вдова)	0,112948	0,028077	0,026650	0,203603
Образование ученая степень	0,034172	0,028252	0,011342	0,038519
Образование среднее образование	0,000000	0,000000	0,000000	0,000000
Образование высшее образование	-0,005723	0,000656	0,001391	0,039997
Образование среднее профессиональное	0,020949	0,020345	0,008273	0,114345
Наличие имущества - отсутствие	0,022313	0,004034	0,006316	0,066077
Наличие имущества - наличие движимого имущества	0,007932	0,004852	0,001586	0,021650
Наличие имущества - наличие недвижимого и движимого имущества	0,015692	0,016682	0,003053	0,026464
Наличие имущества - наличие недвижимого имущества	-0,025017	-0,006510	0,002566	0,013647
Пол - женский	0,002416	0,004047	0,002817	0,082787
Пол - мужской	0,000763	0,001424	0,000232	0,025264

Из таблицы 5 можно заметить, что наибольшую ценность имеют

1. Изменение значения признака “Семейное положение разведен (разведена)” приводит к уменьшению вероятности принадлежности к классу 1 и увеличению уверенности модели в прогнозе. Как разработанный метод, так и метод SHAP видят влияние, первый чуть выше, второй – более сдержано.

2. Изменение значения признака “Семейное положение женат (замужем)” приводит к

увеличению вероятности принадлежности к классу 1 и уменьшению уверенности модели в прогнозе. Метод SHAP считает данный показатель более значимым чем разработанный метод.

3. Изменение значения признака “Семейное положение вдовец (вдова)” приводит к увеличению вероятности принадлежности к классу 1 и уменьшению уверенности модели в прогнозе. Как для разработанного метода, так и для SHAP показатель сильно важен.

4. Изменение значения признака “Образование ученым” и “Образование среднего профессионального” приводит к увеличению вероятности принадлежности к классу 1 и уменьшению уверенности модели в прогнозе. Метод SHAP чуть сильнее подчеркивает значимость признака по сравнению с разработанным методом.

5. Изменение значения признака “Пол - женский” и “Пол - мужской” приводит к увеличению вероятности принадлежности к классу 1 и уменьшению уверенности модели в прогнозе. Метод SHAP считает первый показатель значимым, а второй имеет значимость небольшую. Разработанный метод оба признака признает небольшой по степени важности.

Исходя из вышеуказанного можно сделать вывод, что предложенный метод уникален и помогает оценивать чувствительность модели к изменениям в признаках. Данный метод не заменяет метод SHAP. Разработанный метод и SHAP меряют разные аспекты важности признаков, так SHAP даёт картину глобального среднего вклада признака в предсказание (учитывает как величину, так и направление влияния), а разработанный метод показывает среднюю чувствительность модели к изменению признака в выборке (то есть показывает стабильность модели).

Учитывая это когда значение SHAP высокий, а значение разработанного метода низкий, то признак важен для небольшой подгруппы объектов (локальный «триггер»), а когда SHAP низкий, а значение разработанного метода высокий, то признак даёт небольшой, но постоянный вклад по всей выборке.

На основе всего выше сказанного можно утверждать, что оба метода вместе дают двухмерную карту важности признаков и на их пересечении можно находить самые интересные переменные.

#### **Заключение**

Любая кредитная организация действует в нестабильной среде и, соответственно, для нее важно своевременное выявление и управление рисками. Основным типом рисков является кредитный риск.

Целью данного исследования была создание модели кредитного скоринга на основе алго-

ритма машинного обучения (XGBoost), а также разработка метода оценки чувствительности модели к изменению значения признаков, участвовавших в прогнозировании. На основе полученных результатов можно сделать вывод, что на основе энтропии можно создавать метод оценки чувствительности модели. Данный метод дополняет уже существующий SHAP метод и их совместное использование помогает оценивать модель и на пересечении их результатов найти интересные закономерности.

Помимо выше указанного можно сделать вывод, что использование технологий машинного обучения позволяет оценивать и оптимально сдерживать кредитный риск, и, как следствие, снизить убытки по возможным потерям по ссудам.

#### **Перечень использованной литературы**

1. Корреляция (Correlation), <https://wiki.loginom.ru/articles/correlation.html> (Дата обращения: 08.08.2025)
2. Лобанов А. А., Чугунов А. В. Энциклопедия финансового риск-менеджмента. – М.: Альпина Паблишер, 2003. – 323 с.
3. Evaluation Metrics for Classification, May 17, 2023, <https://medium.com/@mlmind/evaluation-metrics-for-classification-fc770511052d> (Date of access: 06.08.2025)
4. FICO score, <https://acra.am/upnp/?lang=en> (Date of access: 08.08.2025)
5. Pandas User Guide, [https://pandas.pydata.org/docs/user\\_guide/index.html](https://pandas.pydata.org/docs/user_guide/index.html) (Date of access: 06.08.2025)
6. Sklearn User Guide, [https://scikit-learn.org/stable/user\\_guide.html](https://scikit-learn.org/stable/user_guide.html) (Date of access: 06.08.2025)
7. Understanding AUC — ROC and Precision-Recall Curves, Sep 19, 2022, <https://medium.com/@data.science.enthusiast/auc-roc-curve-ae9180eaf4f7> (Date of access: 06.08.2025)
8. What is A Confusion Matrix in Machine Learning? The Model Evaluation Tool Explained, Nov 10, 2024, <https://www.datacamp.com/tutorial/what-is-a-confusion-matrix-in-machine-learning> (Date of access: 06.08.2025)

Сдана/Հանձնվել է 06.08.2025

Рецензирована/Գրախոսվել է 13.08.2025

Принята/Ընդունվել է 19.08.2025