

ОЦЕНКА ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ МЕТРИК КАЧЕСТВА АЛГОРИТМОВ БИНАРНОЙ КЛАССИФИКАЦИИ ПРИ ВАРЬИРОВАНИИ ПАРАМЕТРОВ НЕЙРОСЕТИ И ТИПОВ ПРОПУСКОВ В ДАННЫХ

Ю. Ю. Феста

ПАО «Сбербанк», Москва, Россия
E-mail: festa.y.yura@gmail.com

Исследование доверительных интервалов метрик алгоритмов классификации в низко-дефолтных портфелях имеет ключевое значение для оценки надежности и стабильности прогнозируемых показателей. Такой тип портфелей характеризуются малым количеством дефолтов, и малым размером выборки, что делает задачу прогнозирования особенно сложной. Доверительный интервал показывает диапазон возможных значений метрики качества алгоритмов классификации, в котором истинное значение находится с определенной вероятностью. В условиях ограниченного количества данных, использование доверительных интервалов становится критическим. Оно помогает избежать выбора модели исходя только из точечной оценки, что может привести в дальнейшем к реализации модельного риска. В данной работе исследуются доверительные интервалы метрики классификации с использованием нейросетевых моделей с учетом природы пропусков в данных. В данной работе представлен способ моделирования типов пропусков в данных. Показано, что снижение случайности пропусков приводит к нелинейному изменению зависимости точечной оценки метрики и доверительного интервала заметно резкое снижение ширины доверительных интервалов при увеличении числа нейронов на скрытом слое, при этом увеличивается среднее значение метрики. Чем меньше обучающая выборка, тем больше выражены области сужения ширины доверительного интервала.

ESTIMATION OF CONFIDENCE INTERVALS OF ACURACY METRICS OF BINARY CLASSIFICATION ALGORITHMS WHEN VARYING NEURAL NETWORK PARAMETERS AND TYPES OF MISSING DATA

Y. Y. Festa

The study of confidence intervals of metrics of classification algorithms in low-debt portfolios is of key importance for assessing the reliability and stability of forecasted indicators. This type of portfolio is characterized by a small number of defaults and a small sample size, which makes the forecasting task especially difficult. The confidence interval shows the range of possible values of the quality metric of classification algorithms, in which the true value is found with a certain probability. In conditions of a limited amount of data, the use of confidence intervals becomes critical. It helps to avoid choosing a model based only on a point estimate, which can lead to the realization of model risk in the future. In this paper, we study the confidence intervals of the classification metric using neural network models, taking into account the nature of data gaps. This paper presents a way to model the types of gaps in the data. It is shown that a decrease in the randomness of omissions leads to a nonlinear change in the dependence of the point estimate of the metric and the confidence interval, a noticeable sharp decrease in the width of the confidence intervals with an increase in the number of neurons on the hidden layer, while the average value of the metric increases. The smaller the training sample, the more pronounced the areas of narrowing of the confidence interval width.

Введение

В 2021 году Центральный банк Российской Федерации представил доклад о применении больших данных и методов машинного обучения в финансовом секторе [1]. В данном документе рассматриваются потенциальные преимущества использования технологий машинного обучения (ML), включая повышение точности дифференцирования рисков, улучшение количественных оценок рисков, а также совершенствование процессов хранения и обработки данных. Вместе с тем анализируются и возможные ограничения их применения. К основным рискам, связанным с использованием ML, специалисты относят методологические и модельные риски, такие как проблемы, обусловленные качеством и полнотой данных, а также недостаточной квалификацией персонала, работающего с моделями. В том же году Европейское банковское управление выпустило аналогичный дискуссионный документ, ориентированный на углубленное изучение вопросов и перспектив, связанных с применением алгоритмов машинного обучения для расчета вероятности дефолта (PD) в рамках подходов на основе внутренних рейтинговых систем [2]. Отдельно подчеркивается необходимость обеспечения прозрачности и интерпретируемости используемых моделей. Сравнительный обзор методов машинного обучения и искусственного интеллекта, используемых для анализа PD, приведен в работе [3]. Подробный обзор методологий и требований к моделям PD в соответствии с требованиями МСФО 9 представлен в работе [4].

В этой статье продемонстрировано как типы пропусков данных влияют на доверительные интервалы метрик классификации. Показано, что в зависимости от объема обучающей выборки при фиксированном балансе классов эффекты становятся менее выраженными. Продемонстрировано, что снижение случайных пропусков приводит к нелинейным изменениям зависимости точечных оценок метрики и их доверительных интервалов. При увеличении количества нейронов в скрытых слоях наблюдается значительное уменьшение ширины доверительных интервалов, а также увеличение среднего значения метрики. Чем меньше размер обучающей выборки, тем более выраженными становятся зоны сужения ширины доверительного интервала.

Источники

Использование ML в PD

Модели машинного обучения и искусственного интеллекта при расчёте вероятности дефолта заемщика являются наиболее востребованными в настоящее время. Это подтверждает как пристальное внимание со стороны регуляторов, так и количеству исследований, посвященных этой теме. В статье [5] представлен подход к расчету вероятности дефолта как к временному ряду на основе всей последовательности финансовых отчетов компаний. Авторы утверждают, что предложенная модель, в основе которой лежит рекуррентная нейросеть, превосходит стандартные методы машинного обучения на тестовой выборке. Использование методов машинного обучения для определения вероятности дефолта для корпоративных портфелей рассматривается в статье [6]. Авторы предлагают в качестве выбора наилучшей модели более «классические» модели машинного

обучения такие как как градиентный бустинг и деревья решений. Модель оценивалась по точечным оценкам метрик. В статье [7] авторы показывают, что их PD модель с использованием алгоритма градиентного бустинга позволяет на исторических данных повысить экономию от 12,4% до 17% с точки зрения нормативных требований к достаточности капитала. С помощью алгоритмов машинного обучения и искусственного интеллекта также можно определить наиболее значимые признаки дефолта. В статье [8] авторы используют модели логистической регрессии и нейросети для определения наиболее значимых признаков дефолта по микрокредитам. Проблеме различия в подходах к низкодефолтным портфеля посвящена статья [9]. В ней авторы предлагают реализацию метода максимального правдоподобия основанного на Гауссовских процессах чтобы избежать проблемы недостатка данных в обучающей выборке.

Типы пропусков в данных

Дональд Рубин [10] в 1976 году предложил классификацию пропусков данных на три основных типа: полностью случайные (MCAR), случайные в зависимости от наблюдаемых данных (MAR) и неслучайные (MNAR). Данная классификация послужила отправной точкой для последующего развития методов обработки недостающих данных, включая методы «внедрения» и исключения наблюдений с пропусками. В работе акцентировано внимание на важности учета причин возникновения пропусков, что способствовало лучшему пониманию потенциальных последствий их игнорирования или ошибочной интерпретации. Классификационная схема Рубина оказала существенное влияние на статистику и машинное обучение и используется до сих пор [11] при проведении анализа данных и создании моделей. Осознание типов пропусков стало критичным фактором обеспечения достоверности и надежности выводов, полученных на основе неполных данных, особенно при использовании моделей машинного обучения [12]. В нашем исследовании мы интерпретировали MCAR, MAR, MNAR следующим образом:

- MCAR - пропуски данных случайны, они не зависят ни от объясняющих ни от целевых переменных;
- MAR - пропуски имеют некую статистическую зависимость, не зависят ни от объясняющих ни от целевых переменных;
- MNAR - пропуски не случайны, они зависят от значения целевой переменной.

В [12, 13] авторы упоминают, что в случае MNAR данные восстановить невозможно, однако, по всей видимости, никто не исследовал, как именно MNAR влияет на качество моделей.

Доверительные интервалы

В большинство работ по использованию моделей машинного обучения в задачах определения вероятности дефолта, в том числе и представленных выше, используют точечные оценки метрик классификации. Однако простые точечные оценки качества алгоритмов классификации не являются достаточной основой для выбора наилучшей модели. Арифметически разные значения могут быть статистически равными. Именно поэтому используются доверительные интервалы

для показателей точности классификации. Их можно вывести, используя нормальное приближение или с использованием процедур Агрести-Коули, Уилкоксона, Джеффреса (Jeffreys interval) [14]. Нормальное приближение предполагает, что выборочное распределение нормально распределено с известной дисперсией. Оно основано на центральной предельной теореме, согласно которой сумма независимых случайных величин стремится к нормальному распределению при большом числе наблюдений. В контексте нашего исследования это предполагает, что при небольшом объеме выборки и/или большом дисбалансе классов данное приближение наименее консервативно и доверительные интервалы будут довольно широки. Интервал Джеффреса основан на предположении, что выборочное распределение принадлежит семейству бета-распределений. Это делает его более универсальным, чем нормальное приближение, поскольку оно может использоваться даже для малых выборок. Данное приближение полезно для анализа биномиальных данных и данных, ограниченных диапазоном от нуля до единицы (вероятность). Он позволяет строить доверительные интервалы для малых выборок, когда нормальное приближение становится менее надежным [15]. В исследовании приводятся два типа доверительных интервалов для показателей точности модели: нормальное приближение и интервал Джеффреса.

Методология

Источники данных

В данной работе сравнивались результаты работы модели на двух наборах данных: наборе данных по корпоративным дефолтам, данные получены из открытых источников. Исходный набор данных состоит из 4043 строк, 26 объясняющих переменных и долей дефолта в 0.04. Названия переменных обезличены и стандартизированы. В данной работе не ставится цель получить наилучшие метрики качества алгоритма классификации, поэтому примем допущение, что все представленные переменные значимые. Разбиение на обучающую и контрольную выборку уже представлено в данных, поэтому сохраним готовые метки подвыборки. Также исследовался синтетический набор данных, сгенерированный самостоятельно реализованным white-box алгоритмом. Мы генерируем синтетический набор данных с идентичным количеством наблюдений и долей дефолта, при этом также генерируя типы пропусков. Это позволит нам экспериментально проверять влияние каждого типа пропуска на результаты модели, в том числе на доверительные интервалы метрик. В работе рассматривается, как меняются доверительные интервалы для показателей точности классификации при различных размерах выборки и типов пропусков в данных при фиксированном балансе классов.

Алгоритм генерации синтетических данных

Для генерации данных использовался следующий порядок действий:

- 1) Сгенерировано n столбцов из нормального распределения $\sim N(0,1)$ и 1 столбец из равномерного распределения $\sim U(0,1)$
- 2) Для каждой строки рассчитана сумма по всем переменным из нормального распределения и равномерного, взятого с весом w , назовем это значение *Score*

3) Для полученных значений $Score$ рассчитана медиана

4) Для строк, значения которых меньше или равны медиане из предыдущего шага присваивается метка 0 («не дефолт»), выше – 1 («дефолт»)

Математически вышеописанные шаги можно записать следующим образом:

$$Score_j = \sum_{i=1}^n X_{ij} + w * U_i \quad (1)$$

где n – количество объясняющих переменных (в нашем случае равно 26); j – количество наблюдений, $X_{ij} \sim N(0,1)$ – нормально распределенная случайная величина, $U_i \sim U(0,1)$ – равномерно распределенная случайная величина (шум), $w = 0.3$ – весовой коэффициент.

Далее определим метку класса Y_j :

$$Y_j = 0 \mid Score_j \leq Median(Score_j), \text{ else } Y_j = 1 \quad (2)$$

где 1 - дефолт, 0 - не дефолт.

Таким образом мы получаем набор данных со сбалансированными метками классов. В дальнейшем мы увеличим количество строк для генерации в 2, 3, 4 и 5 раз при этом сохраняя долю дефолтов в выборке путем моделирования пропусков.

Моделирование типов пропусков

Далее, в зависимости от смоделированного типа пропусков в данных, мы исключаем наблюдения, пока не получим желаемую долю наблюдений целевого класса. Отмечается, что в исследовании мы стараемся избегать случаев «черного ящика» при генерации данных, поэтому для моделирования MCAR не используем генератор случайных чисел из математических пакетов, для этого использовался многократный вызов испытания Бернулли. Предполагаем, что данный подход менее детерминирован, чем алгоритм генерации случайных с заданным изначальным состоянием.

- Для моделирования ситуации MCAR использовалась рекурсивная функция вызова испытания Бернулли с фиксированной вероятностью успеха 0.005. С такой вероятностью в одной итерации удалялась случайная строка. Цикл повторялся пока не достигалось требуемое значение доли целевого класса в выборке.

- Для моделирования ситуации MAR в цикле фиксировался индекс строки и удалялся каждый i -й элемент массива, где i зависит от текущего размера массива. Цикл также повторялся пока не достигалось требуемое значение доли целевого класса в выборке.

- Для моделирования ситуации MNAR в зависимости от требуемого значения доли целевого класса фиксировался перцентиль по $score$, рассчитанным на этапе генерации данных, и удалялись все строки со значением $score$ ниже зафиксированного перцентиля.

Несмотря на рекомендации [16] использовать всю выборку, не разбивая ее на обучающую и тестовую, в данном исследовании все же производится разбиение

ние. Автор аргументирует свою позицию, что любое разделение ретроспективных данных не может заменить эксперимент, чего всегда не хватает экономической науке (за исключением сферы поведенческой и экспериментальной экономики). В данном исследовании в качестве модели классификации выбрана нейросеть, и даже при сильной регуляризации она может просто «запомнить» метки классов при небольшом объеме данных.

Расчет доверительных интервалов

В исследовании рассматриваются два типа доверительных интервалов: нормальная аппроксимация и интервал Джеффреса. Формула для расчета нижней и верхней границы для нормальной аппроксимации:

$$CI_l^N = \left(\frac{S}{n}\right) - \gamma_{\frac{\alpha}{2}} * \sqrt{\frac{\left(\frac{S}{n}\right) * \left[1 - \left(\frac{S}{n}\right)\right]}{n}}, \quad (3)$$

$$CI_u^N = \left(\frac{S}{n}\right) + \gamma_{\frac{\alpha}{2}} * \sqrt{\frac{\left(\frac{S}{n}\right) * \left[1 - \left(\frac{S}{n}\right)\right]}{n}} \quad (4)$$

где n – число испытаний, S – число успехов, $\gamma_{\alpha/2}$ – квантиль нормального распределения с уровнем значимости $\alpha/2$.

Формула для расчета нижней и верхней границы интервала Джеффреса:

$$CI_l^J = Beta\left(\frac{\alpha}{2}; S + 0.5; n - S + 0.5\right), \quad (5)$$

$$CI_u^J = Beta\left(1 - \frac{\alpha}{2}; S + 0.5; n - S + 0.5\right) \quad (6)$$

где n – число испытаний, S – число успехов, $Beta(\alpha, a_1, a_2)$ квантиль Бета распределения с параметрами a_1 и a_2 .

Для построение доверительных интервалов для биномиальной переменной необходимо определить число «успехов» и «неудач» для серии экспериментов.

Так как мы имеем дело с задачей классификации то в качестве метрик будем использовать стандартную матрицу ошибок [17].

Матрица ошибок классификации

| | | |
|------------------|----------------------|-------------------------|
| | Дефолт (предсказано) | Не дефолт (предсказано) |
| Дефолт (факт) | TP | FP |
| Не дефолт (факт) | FN | TN |

Accuracy — это соотношение правильно предсказанных случаев дефолта и не дефолта для всех наблюдений:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

Precision — это доля правильных предсказаний класса от общего числа случаев истинных значений дефолта:

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

Recall — это способность модели находить все дефолты в выборке:

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

F1 — гармоническое среднее между *Precision* и *Recall*:

$$F_1 = 2 * \frac{Precision \cdot Recall}{Precision + Recall} \quad (10)$$

Расчета числа «успехов» и «неудач» для каждой из метрик следует из формул, указанных выше и представлен в табл. 2.

Таблица 2

Исходы испытания Бернулли для метрик классификации

| Метрика | Число успехов | Число неудач |
|------------------|---------------|--------------|
| <i>Accuracy</i> | $TP + TN$ | $FP + FN$ |
| <i>Precision</i> | TP | FP |
| <i>Recall</i> | TP | FN |
| <i>F1</i> | $2*TP$ | $FN+FP$ |

Далее для каждой из метрик по числу «успехов» и «неудач» считаем доверительные интервалы для нормального приближения и аппроксимации Джеффреса используя метод *proportion_confit* пакета *statsmodels* [18].

Параметры модели полносвязной нейросети прямого распространения

В данном исследовании использована архитектура полносвязной сети прямого распространения. Данная архитектура достаточно интерпретируема и такие модели не требуют больших вычислительных затрат на обучение. Несмотря на появление нейронных сетей со сложными архитектурами лежащие в основе больших языковых моделей, сети данной архитектуры до сих пор актуальны. Как отмечают исследователи, пик популярности использования данной архитектуры уже прошел [19], однако данная архитектура остается востребованной у исследователей в таких областях как инженерные науки [20] и медицина [21]. На рисунке 1 представлен пример архитектуры такой сети с одним скрытым слоем.

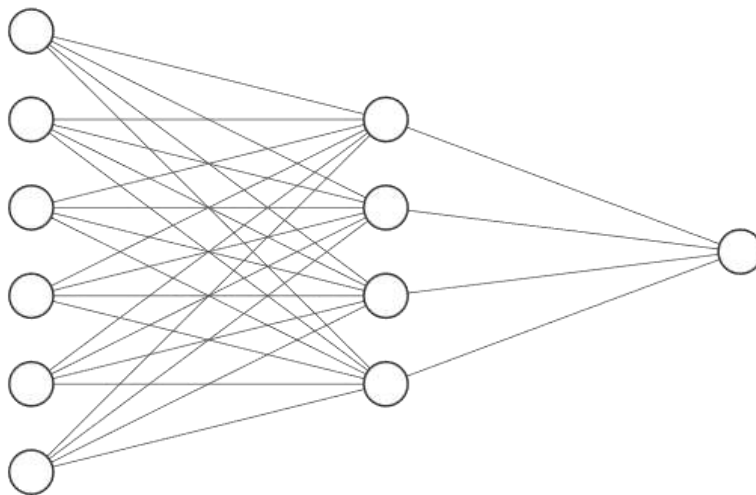


Рис. 1. Полносвязная нейросеть с одним скрытым слоем

Для построения модели нейросети использовался пакет *keras*. Количество нейронов на входном слое рассчитывалось как n , где n – количество объясняющих переменных. Количество нейронов на скрытом слое принималось от $0.5*n$ до $1.5*n$. Функция активации на скрытом слое – *relu*. На выходном слое один нейрон с функцией активации *sigmoid*. *Loss function* - '*binary_crossentropy*'.

Основные полученные результаты

При моделировании типов пропусков и параметров нейросети обнаружены следующие закономерности:

1) Отмечается, что чем меньше случайности в пропусках, тем выше точечные оценки метрик, при этом на малой выборке зависимость ширины доверительного интервала от среднего значения метрики от линейной переходит в степенную. Чем выше значение метрики, тем меньше ширина доверительного интервала. Далее приложены графики для метрики *F1*. График зависимости для реальных данных приложен на рисунке 2. Зависимость для MCAR, MAR и MNAR представлены на рисунках 3, 4 и 5 соответственно.

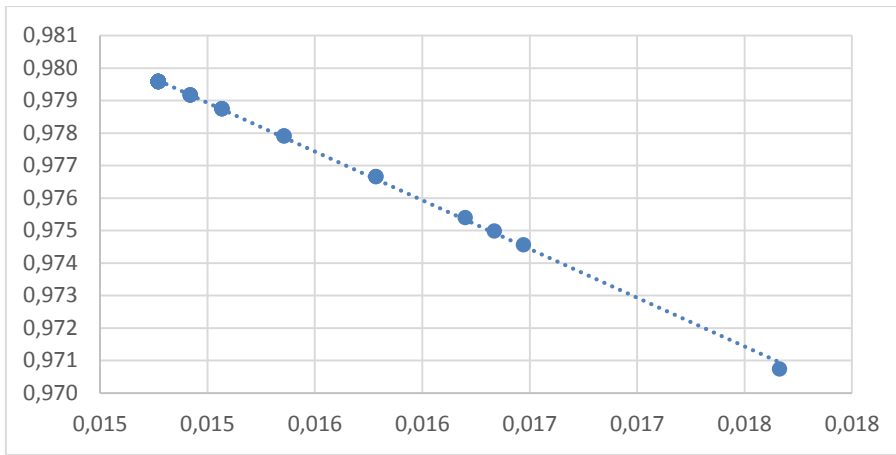


Рис. 2. Реальный набор данных

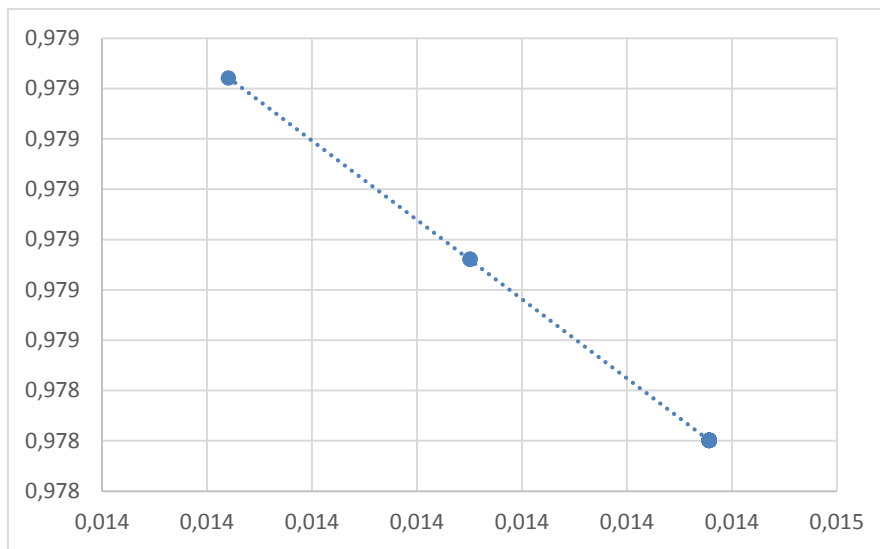


Рис. 3. Набор данных с пропусками MCAR

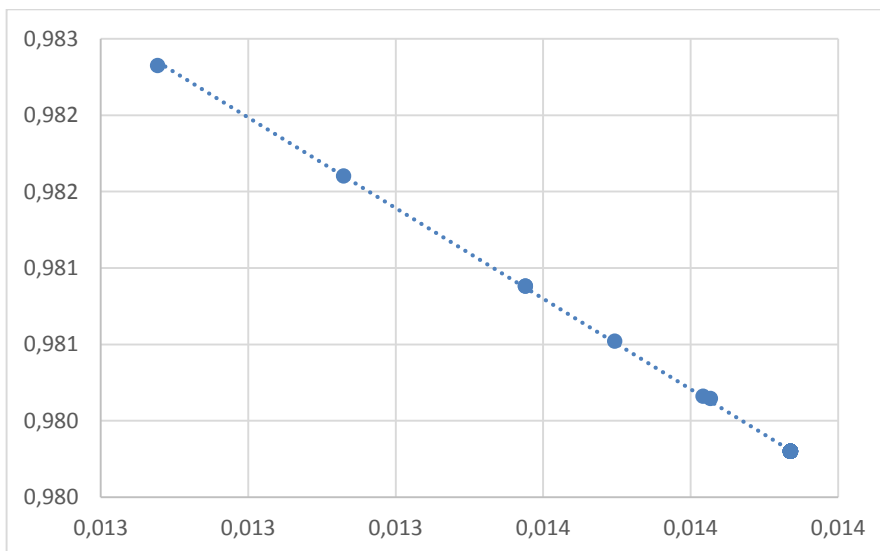


Рис. 4. Набор данных с пропусками MAR

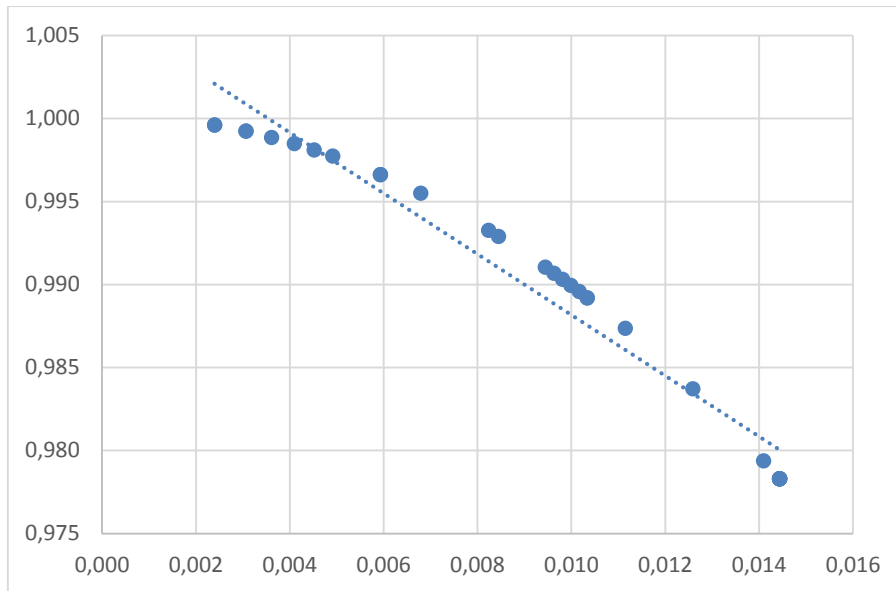


Рис. 5. Набор данных с пропусками MNAR

2) В случае MNAR заметно резкое снижение ширины доверительных интервалов при увеличении числа нейронов на скрытом слое, при этом увеличивается среднее значение метрики. Чем меньше обучающая выборка, тем больше выражены области сужения ширины доверительного интервала. Предполагается, что для выбора наилучшей модели необходимо найти параметры модели, где при наибольшей. В случае MCAR и MAR ширина интервалов практически не изменяется. Зависимость для MCAR, MAR и MNAR представлены на рисунках 6, 7 и 8 соответственно.

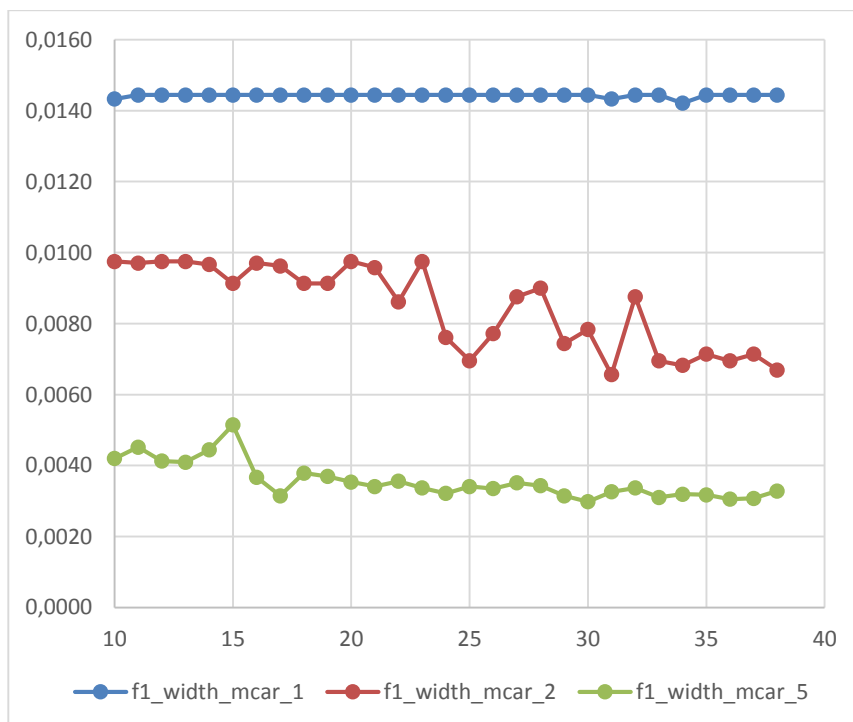


Рис. 6. Набор данных с пропусками MCAR

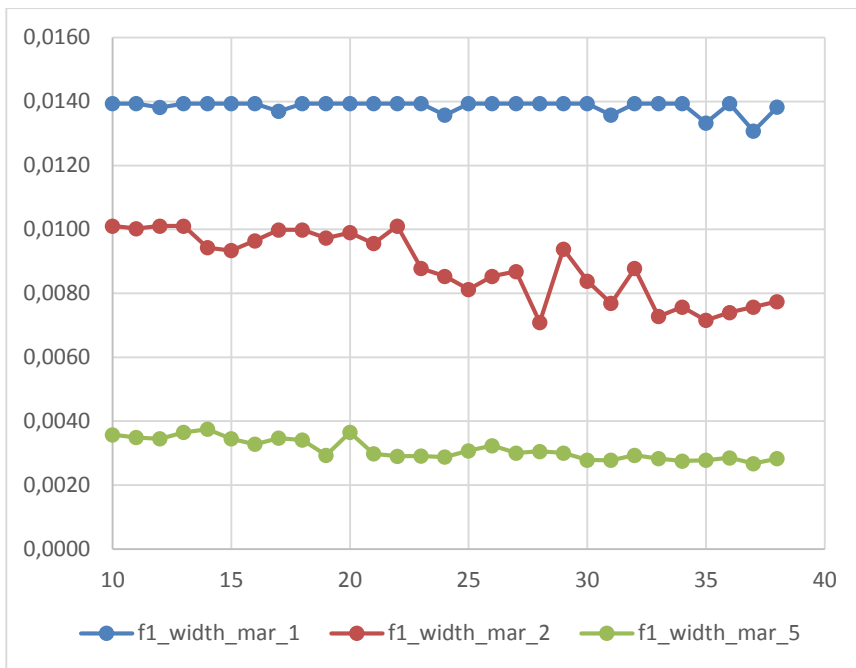


Рис. 7. Набор данных с пропусками MAR

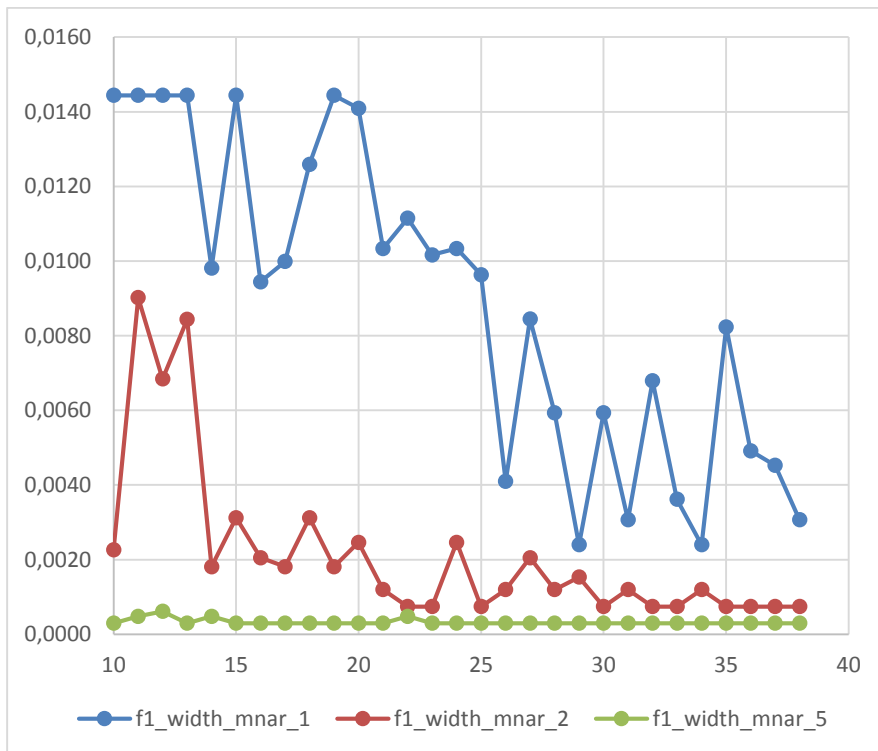


Рис. 8. Набор данных с пропусками MCAR

3) В случае MNAR зависимость ширины метрики $F1$ от $Accuracy$ отклоняется от линейного тренда.

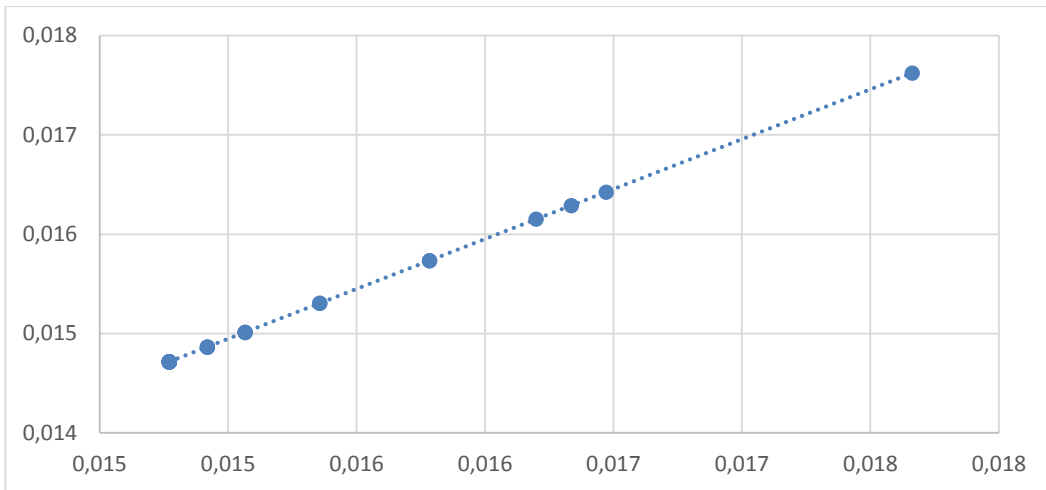


Рис. 9. Реальный набор данных

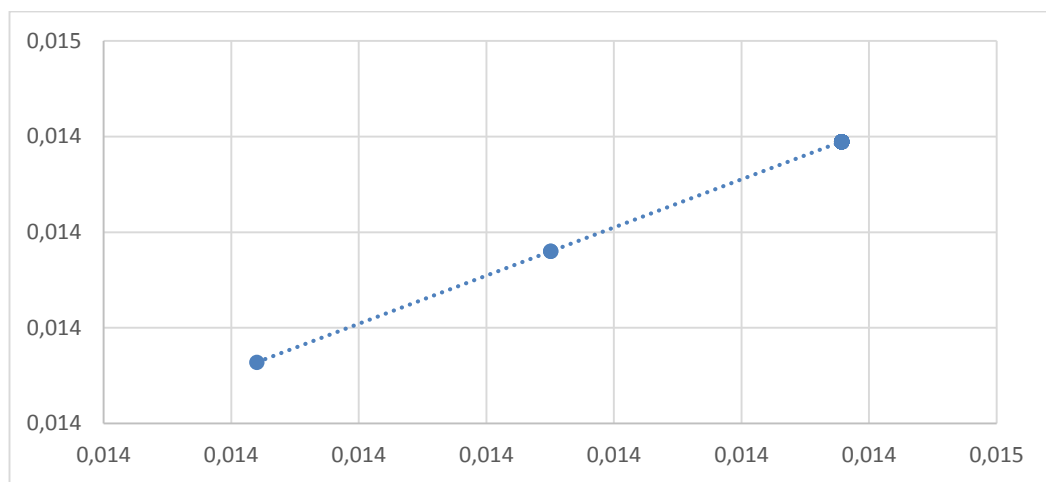


Рис. 10. Набор данных с пропусками MCAR

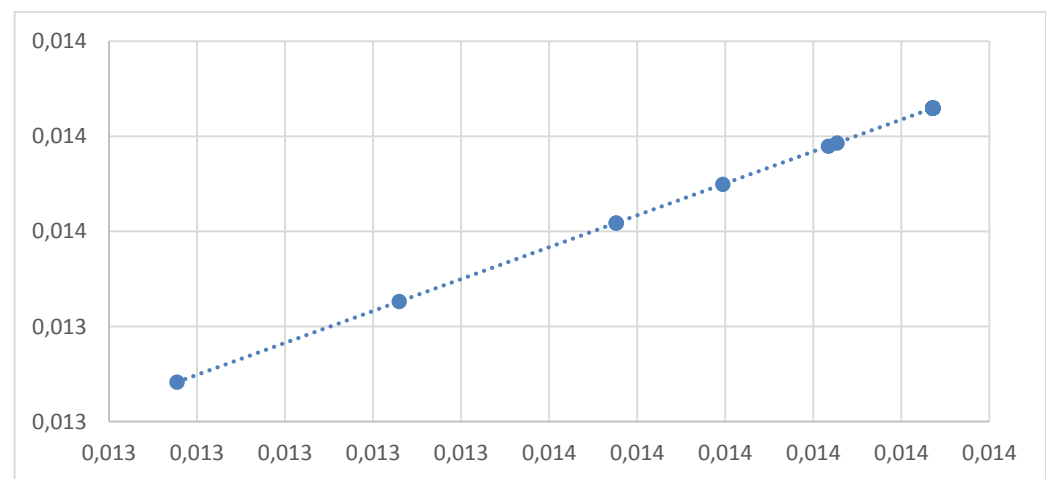


Рис. 11. Набор данных с пропусками MAR

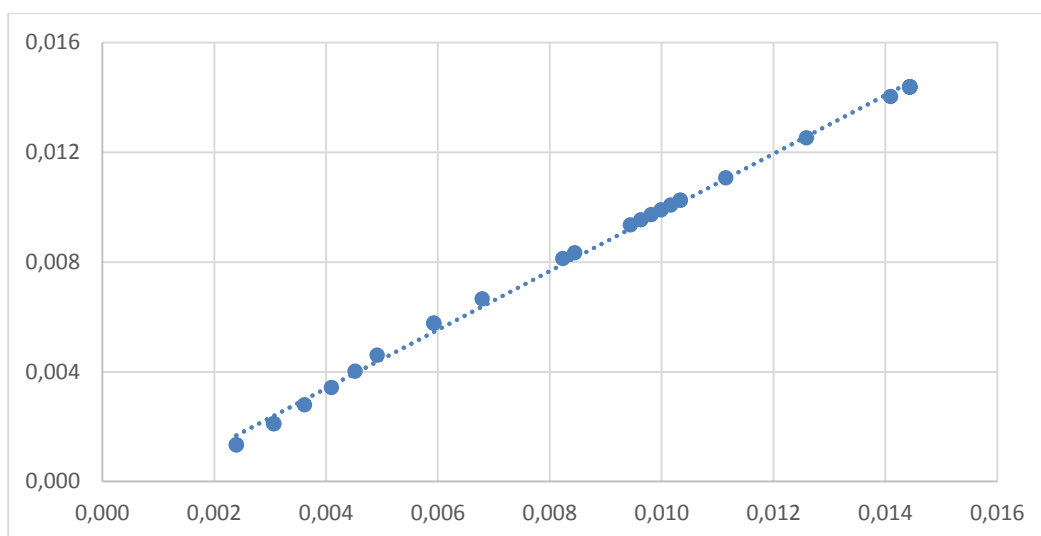


Рис. 12. Набор данных с пропусками MNAR

Выводы

В работе предложен алгоритм моделирования пропусков в данных для типов MCAR, MAR, MNAR. Данный алгоритм может быть полезен при использовании метода Undersampling в случае необходимости получить нужный баланс классов, что может быть полезно при моделировании портфелей заемщиков в задаче расчёта вероятности дефолта. Методы уменьшения целевого класса могут быть особенно полезны при моделировании низкодефолтных портфелей, где всегда ярко выражен дисбаланс классов и природа отсутствия наблюдений нам неизвестна. Причины наличия таких пропусков выходят за пределы исследования данной статьи, но, зачастую, исследователи данных уже поставлены перед фактом наличия неполной выборки и дополнительный сбор данных не представляется возможным. Найденные закономерности могут позволить определить тип пропусков в данных и в дальнейшем выбрать подход к заполнению этих пропусков, либо с использованием black-box моделей статистических пакетов с заданными параметрами, либо вручную, отбирая данные основываясь на экспертизе аналитика. В дальнейших исследованиях планируется параметризовать полученные результаты, чтобы по оценке доверительных интервалов и точечной оценки метрики можно было определить тип пропусков в исходном наборе данных и предложить наиболее оптимальный метод тиражирования данных для минимизации эффектов, присущим низкодефолтным портфелям.

СПИСОК ЛИТЕРАТУРЫ

1. Доклады для общественных консультаций. Банк России. [Электронный ресурс]. URL: https://www.cbr.ru/analytics/d_ok/#y2021 (дата обращения: 15.11.2024).
2. EBA consults on machine learning for internal ratings-based models. European Banking Authority. [Электронный ресурс]. URL: <https://www.eba.europa.eu/publications-and-media/press-releases/eba-consults-machine-learning-internal-ratings-based-models> (дата обращения: 15.11.2024).
3. Bonini S., Caivano G. Probability of Default Modeling: A Machine Learning Approach // *Mathematical and Statistical Methods for Actuarial Sciences and Finance: MAF 2018*. 2018. С. 173-177.

4. Martinelli F. и др. Predicting Probability of Default Under IFRS 9 Through Data Mining Techniques // Web, Artificial Intelligence and Network Applications Advances in Intelligent Systems and Computing. 2020. С. 959-969.
5. Gated recurrent unit network: A promising approach to corporate default prediction - Thor Journal of Forecasting Wiley Online Library. 2024. [Электронный ресурс]. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/for.3057> (дата обращения: 13.11.2024).
6. Moscatelli M. [и др.]. Corporate default forecasting with machine learning // Expert Systems with Applications. 2020. Vol. 161. С. 113567.
7. Alonso-Robisco A., Carbó J. M. Can machine learning models save capital for banks? Evidence from a Spanish credit portfolio // International Review of Financial Analysis. 2022. Vol. 84. С. 102372.
8. Analysis of default risk in microfinance institutions under the Basel III framework - Durango-Gutiérrez // International Journal of Finance & Economics - Wiley Online Library 2023 [Электронный ресурс]. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/ijfe.2475> (дата обращения: 13.11.2024).
9. He J., Khedher A., Spreij P. Calibration of the rating transition model for high and low default portfolios. 2024
10. Inference and missing data. Biometrika. Oxford Academic. [Электронный ресурс]. URL: <https://academic.oup.com/biomet/article-abstract/63/3/581/270932> (дата обращения: 13.11.2024).
11. Morimoto J. The Effect of Sample Size and Missingness on Inference with Missing Data. 2022.
12. Emmanuel T. [и др.]. A survey on missing data in machine learning // Journal of Big Data. 2021. № 1 (8). С. 140.
13. Journal of Statistical Software 2011. Vol. 45.
14. Engelmann B., Hayden E., Tasche D. Testing Rating Accuracy. Risk, 16, 82-86. - References - Scientific Research Publishing. 2003. [Электронный ресурс]. URL: <https://www.scirp.org/reference/referencespapers?referenceid=2081392> (дата обращения: 13.11.2024).
15. Brown L. D., Cai T. T., DasGupta A. Interval Estimation for a Binomial Proportion // Statistical Science. 2001. № 2 (16). С. 101-117.
16. Diebold F. X. Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests // Journal of Business & Economic Statistics. 2015. № 1 (33). С. 1-9.
17. Raschka S., Mirjalili V. Python Machine Learning Machine Learning and Deep Learning with Python, Scikit-Learn and TensorFlow 2. 3rd Edition, Packt Publishing, Birmingham. // Scientific Research Publishing. 2019. [Электронный ресурс]. URL: <https://www.scirp.org/reference/referencespapers?referenceid=3293920> (дата обращения: 13.11.2024).
18. Perktold J. [и др.]. statsmodels/statsmodels: Release 0.14.2. 2024.
19. Sadler I., Reimann N., Sambell K. Feedforward practices: a systematic review of the literature // Assessment & Evaluation in Higher Education. 2023. № 3 (48). С. 305-320.
20. Chen Z. [и др.]. Residual strength prediction of corroded pipelines using multilayer perceptron and modified feedforward neural network // Reliability Engineering & System Safety. 2023. (231). С. 108980.
21. Avinash S. [и др.]. Early Detection of Malignant Tumor in Lungs Using Feed-Forward Neural Network and K-Nearest Neighbor Classifier // SN Computer Science. 2023. № 2 (4). С. 195.