

Обзор литературы по теме «Применение машинного обучения в управлении кредитными рисками»

Столярова Анастасия, БЭАД223

10 декабря 2023 г.

Аннотация

Кредитный риск - это риск финансовых потерь, которая возникает, когда заемщик не выполняет своих финансовых обязательств перед кредитором. Для уменьшения этого риска банки и финансовые учреждения должны проводить оценку кредитоспособности заемщиков, контролировать их платежи и использовать различные модели поведения. В последние годы наблюдается рост проблемных активов (NPA) и мошенничества, поэтому важно использовать надежные методы для прогнозирования эффективности займов. Исследователи активно изучают различные методы машинного обучения для повышения точности оценки кредитного риска, помимо рейтинговых агентств и компаний.

В представленном обзоре рассматривается 12 статей, опубликованных в рейтинговых профильных научных журналах в период с 2019 по 2023 год. На основании выбранного шорт-листа статей производился анализ методов, улучшающих решение задачи предсказания дефолта, и кредитного скоринга в частности, в том числе в соответствии с запросами индустрии. Несмотря на сложности в интерпретации методов машинного обучения, наблюдаются тренд к более частому использованию методов машинного обучения кредитными организациями для анализа рисков, современные исследования направлены на улучшение интерпретируемости моделей. Еще один тренд связан с использованием ансамблевого обучения, которое позволяет уменьшить тяготеющие результаты скоринга к домену.

1 Введение

Кредитный риск стал одной из ключевых областей финансового исследования в последние годы, особенно в свете последних финансовых кризисов. Банки и другие кредитные учреждения прилагают значительные усилия для оценки и управления кредитными рисками, используя различные методы и подходы, включая традиционные статистические методы, а также более сложные и современные методы машинного обучения.

В этом обзоре литературы мы рассмотрим существующие подходы и модели, используемые для оценки кредитного риска, выявим ключевые проблемы и ограничения, а также

обозначим потенциальные направления для дальнейшего развития и улучшения этих методов. Мы также обратим внимание на примеры использования этих методов на конкретных выборках данных и рассмотрим, как применение методов машинного обучения может помочь в решении нестандартных задач оценки кредитного риска.

При анализе кредитных рисков основное внимание уделяется улучшению точности прогнозирования дефолта заемщиков, однако существуют и другие интересные и важные аспекты, которые могут быть рассмотрены с помощью методов машинного обучения. Это включает в себя выявление скрытых закономерностей и зависимостей в данных, а также прогнозирование поведения заемщиков в условиях изменяющегося экономического климата.

2 Методология работы

Первым этапом нашего исследования было определение критериев отбора статей. Мы сосредоточились на профильных международных научных журналах с высоким уровнем цитирования и актуальности научных работ. Включены статьи, опубликованные в течение последних трех лет, что позволяет увидеть самые свежие тенденции в области оценки кредитного риска с использованием методов машинного обучения.

Далее, мы провели систематический обзор выбранных исследований. Начинали мы с изучения ранее опубликованных обзоров литературы, чтобы определить ключевые вопросы и ограничения в использовании методов машинного обучения в данной области. Затем мы перешли к анализу более свежих научных статей и исследований, чтобы выявить новые подходы и решения, предложенные учеными для преодоления выявленных проблем.

Следующим этапом был анализ результатов применения методов машинного обучения на реальных данных. Это помогло нам понять, как эти методы работают на практике и какие результаты они могут дать.

Целью нашего исследования было представить общий обзор современных подходов и тенденций в области оценки кредитного риска с использованием методов машинного обучения. Мы стремились выявить потенциальные направления для дальнейшего развития и улучшения этих методов, а также обозначить возможные проблемы и ограничения при их использовании.

3 Практика применения методов машинного обучения при решении задачи кредитного скоринга

Для того, чтобы определить спектр проблем, возникающих при применении методов машинного обучения для кредитного скоринга, мы изучили наиболее полные обзоры литературы, охватывающие длительные промежутки времени, чтобы получить наиболее полное пред-

ставление об использовании методов машинного обучения для решения подобных задач. Сиддхарт Бхаторе и группа исследователей [2] провели систематический обзор существующих исследовательских методов и методов машинного обучения для оценки кредитного риска. Было рассмотрено в общей сложности 136 документов по оценке кредитных рисков, опубликованных в период с 1993 года по март 2019 года. Авторы изучили влияние гиперпараметров на методы ML, используемые для оценки кредитного риска, проанализировали недостатки нынешних исследований и выявили тенденций в области оценки кредитных рисков. Было замечено, что ансамблевые и гибридные модели с нейронными сетями и SVM (метод опорных векторов) все больше используются для кредитного скоринга, прогнозирования непроизводительных активов и обнаружения мошенничества. В статье выделяются текущие проблемы кредитного скоринга. Каждая модель имеет свои риски и сложности, и нельзя полностью полагаться на одну модель для оценки кредитоспособности или обнаружения мошенничества, согласно "теореме об отсутствии бесплатного обеда". Это связано с проблемой переносимости модели на новые данные или области применения. Банки, находящиеся в разных странах, имеют разные правила и регуляции, поэтому их данные могут существенно отличаться. Если модель обучается на данных из одной области и тестируется на данных из другой, это может привести к потере стабильности. Исследователи изучают эту проблему с помощью ансамблевых методов, которые показали себя лучше, чем отдельные классификаторы. Однако основной проблемой ансамблевого обучения является сложность интерпретации результатов, поэтому улучшение интерпретируемости ансамблевых моделей - это важная область исследований, требующая дальнейшего исследования. По мнению авторов, исследования из обзорной коллекции указывают на необходимость разработки более конкретных инструментов, которые могут решить проблему изменения доменов наборов данных и также обеспечить гибкость в добавлении любого типа модели для оценки кредитного риска. Авторы предполагают, что возможные будущие исследования могли бы заключаться в объединении статистических моделей и моделей машинного обучения в единый инструмент, который помог бы оценить кредитный риск в соответствии с требованиями финансового органа. Поскольку большинство сотрудников банков не являются технологически подкованными, создание интерфейсов, которые не требуют технического понимания, но обеспечивают параллельную обработку, самоадаптацию, самообучение, устойчивость и гибкость оценщикам, повысит применение методов машинного обучения.

В результате анализа данного источника мы выделили основные проблемы: сложность интерпретации результатов, препятствующая повсеместному внедрению методов машинного обучения для анализа кредитных рисков в финансовых организациях из-за законодательных ограничений, отсутствие агрегированной модели, применимой к любому набору данных без потери устойчивости. Дальнейший отбор статей мы проводили в соответствии с тем, как они позволяют решать выделенные проблемы, а также включили в рассмотрение новейшие исследования, позволяющие оценить кредитные риски в нетривиальных моделях, усложненных дополнительными стимулами, присущими участникам кредитных отношений.

4 Изучение предлагаемых улучшений моделей машинного обучения для оценки кредитных рисков

Мы выделили 5 кейсов, в которых предлагаются наиболее эффективные способы разрешения проблем, возникающих при применении машинного обучения при анализе кредитных рисков. Отдельно был изучен вопрос оправданности применения методов машинного обучения для анализа рисков в целом. Фабио Сигрист и Никола Лойенбергер [9] исследовали применение методов машинного обучения для моделирования вероятности дефолта компаний. Они создали новую гибридную модель, сочетающую метод бустинг деревьев с латентным компонентом чувствительности, позволяющим моделировать корреляции, которые не могут быть объяснены наблюдаемыми предикторами. В исследовании использовались следующие модели также рассматривают ансамлевые модели модель, которая комбинирует прогнозы четырех независимых методов: бустинг дерева, нейронные сети, случайный лес и линейную модель риска. Авторы обнаружили, что модели на основе машинного обучения более точны в прогнозировании, чем линейные модели, особенно на долгосрочных горизонтах. Вероятной причиной этого является наличие более сильных эффектов взаимодействия для более длинных горизонтов прогнозирования по сравнению с короткими горизонтами. Среди всех применяемых методов наиболее точными оказались деревья решений. Эта модель обеспечивает наивысшую точность прогнозирования для одно- и двухлетних горизонтов прогнозирования.

Одна из проблем, возникающая при анализе кредитных рисков связана с сложностями возникающими при попытке учесть в модели факторы среды, к которым относится, в частности, экономическая и политическая обстановка в стране. Саба Моради и Фарима Мохатаб Рафией [8] в своем исследовании описали разработку динамической модели для оценки кредитного риска, которая может учитывать колебания политико-экономических факторов. Авторы отметили, что традиционные статические модели, используемые банками, часто не могут точно предсказать невыплаты клиентов, особенно в условиях экономического кризиса. В своей работе они предложили модель, основанную на нечеткой логике и адаптивной сети на основе нечеткого вывода (ANFIS). Эта модель обучается на ежемесячных данных профилей клиентов и затем использует определенные факторы и их базовые правила для второго раунда оценки в системе нечеткого вывода. Авторы утверждают, что их модель является хорошей заменой текущим статическим моделям, так как она может превзойти традиционные модели, особенно в условиях экономического кризиса. Стоит отметить, что в исследовании использовался абстрактный математический аппарат, который затрудняет решение проблемы интерпретируемости данных.

Одним из способов решить проблему плохого обобщения модели для работы с произвольным набором данных является применения специально обученной нейронной сети. Ти Ли, И Пенга и Ганг Коу [6] описали новый подход к обучению нейронных сетей для решения задачи кредитного скоринга, называемый NyströmNet. Этот подход использует метод

Нистрёма, который помогает улучшить оценку кредитоспособности и анализ подшаблонов, устраняя искажения расстояний в функциях ядра и настройку параметров. NyströmNet состоит из двух основных модулей - модуля обучения метрики расстояния и модуля Nyström, которые взаимодействуют друг с другом для достижения оптимальных результатов. Эксперименты с использованием этого подхода показали улучшение в классификации на 2-11% и 2-14% с использованием новых сгенерированных распределений. Подход также имеет преимущества в практическом применении, поскольку он не требует сложной настройки параметров, использует меньше памяти и легко интегрируется с использованием специальных программных инструментов, которые позволяют автоматически вычислять градиенты функций. PyTorch - это один из таких фреймворков, который предоставляет удобные инструменты для работы с нейронными сетями и градиентным спуском.

Возможное решение проблемы невозможности интерпретации предложила группа исследователей в составе Сю Чжу, Цинъюэн Чу, Синьчан Сун, Пин Ху и Лу Пэн. Статья [12] посвящена применению машинного обучения и интерпретируемых моделей для прогнозирования дефолтов по займам. Авторы сравнивают эффективность моделей логистической регрессии, дерева решений, XGBoost и LightGBM, используя большую выборку данных. В статье подчеркивается, что, несмотря на отличные результаты прогнозирования с помощью машинного обучения, эти модели не всегда предоставляют достаточно информации для принятия решений. Даже при высокой точности прогноза, результат анализа решений может показаться нелогичным. Поэтому важно, чтобы те, кто непосредственно принимают решения о выдаче кредита, понимали модель и правила прогнозирования. Для решения этой проблемы авторы используют интерпретируемые модели, которые помогают понять, как работает модель машинного обучения. Они используют метод LIME (Local Interpretable Model-Agnostic Explanations) для объяснения результатов прогнозирования. Авторы выявляют важные характеристики, влияющие на вероятность дефолта, с точки зрения интерпретируемого машинного обучения. Они приходят к выводу, что срок займа, класс займа, статус владения домом, указанный заемщиком при регистрации, сумма займа, размер платежа, отношение долга к доходу и кредитный рейтинг заемщика являются значимыми факторами, влияющими на дефолт по личному займу. В исследовании также подчеркивается, что объяснение модели и явное указание правил прогнозирования с помощью LIME может укрепить доверие пользователей к модели.

В контексте кредитного скоринга, ансамблевые методы, основанные на деревьях решений, такие как метод случайного леса, обеспечивают лучшую классификационную производительность, чем стандартные модели логистической регрессии. Однако логистическая регрессия остается эталоном в индустрии кредитного риска, в основном потому, что отсутствие интерпретируемости ансамблевых методов несовместимо с требованиями финансовых регуляторов. Елена Думитреску, Салливан Ху, Кристоф Хурлин и Сесси Токпави [4] предложили высокопроизводительный и интерпретируемый метод кредитного скоринга под названием "штрафная логистическая регрессия дерева"(PLTR), который использует информацию из деревьев решений для улучшения производительности логистической регрес-

сии. Формально, правила, извлеченные из различных короткоглубинных деревьев решений, построенных с оригинальными прогностическими переменными, используются в качестве предикторов в модели штрафной логистической регрессии. PLTR позволяет улавливать нелинейные эффекты, которые могут возникнуть в данных кредитного скоринга, сохраняя при этом внутреннюю интерпретируемость модели логистической регрессии. Метод Монте-Карло и эмпирические приложения, использующие четыре реальных набора данных о кредитных просрочках, показывают, что PLTR прогнозирует кредитный риск значительно точнее, чем логистическая регрессия, и сопоставимо с методом случайного леса.

5 Сферы использования машинного обучения для решения задачи кредитного скоринга

В рамках обзора литературы были отдельно рассмотрены нетривиальные кейсы, использования машинного обучения при определении кредитных рисков. Выделенные методы позволяют адаптировать модель кредитного скоринга для работы в условиях большей неопределенности и ассиметрии информации.

5.1 Применение моделей машинного обучения для оценки кредитных рисков в сельских районах на примере Перу

Гуина Сотомайор Альзамора, Мигель Ромилио Асейтуньо-Рохо и Генри Иван Кондори-Алехо [1] рассматривали применение моделей машинного обучения для оценки кредитных рисков в сельских районах Перу. Это связано с тем, что сельское население Перу имеет ограниченный доступ к финансовой системе из-за высокой стоимости кредитов и высокого уровня риска (неплатежей), вызванных «серым» сектором экономики. Традиционный процесс оценки для предоставления микрокредитов сопряжен с человеческим риском, так как оценку кредитоспособности агента давали местные кредитные консультанты, которые определяли платежеспособность клиентов с использованием методологий управления рисками, с целью снижения непогашенных кредитов. В статье описывается кейс применения машинного обучения для оценки кредитного риска в сельских районах Перу. Авторы использовали различные модели машинного обучения, чтобы определить наиболее точный уровень утверждения для процесса предоставления кредита и последующего снижения кредитного риска. Эти модели учитывали значимые переменные процесса оценки микрокредита в сельских районах, применяли техники, такие как SMOTE и K-fold, и оценивали модели с помощью некоторых метрик, таких как точность, полнота, F1 Score, AUC, ROC. Модель LightGBM, основанная на деревьях решений, показала отличный уровень утверждения, с коэффициентом успешности займа 96,20%. Результаты по снижению коэффициента просрочки доказывают, что оптимально использовать технологические инструменты, такие как модели машинно-

го обучения, для поддержки принятия решений экспертами по оценке кредитного риска в сельских районах. Таким образом, модель машинного обучения справилась с решением задачи кредитного скоринга условиях сильной асимметрии информации, что демонстрирует возможность успешно адаптировать подобные модели для работы с специфичными выборками.

5.2 Методы оценки кредитных рисков для малых и средних предприятий (МСП)

Алессандро Битетто, Паола Черкиелло, Стефано Филомени, Алессандра Танда и Барбара Тарантино [3] исследовали применение машинного обучения в оценке кредитного риска для малых и средних предприятий. Они сравнивают два подхода к оценке кредитных рейтингов подобных компаний: классический параметрический подход, использующий упорядоченную пробит-модель, и непараметрический подход, использующий модель случайного леса на основе исторических данных (HRF). Для исследования авторы использовали уникальный и запатентованный набор данных, включающий на уровне отдельных фирм квартальные данные, собранные Европейским инвестиционным банком и международной страховой компании по выборке из 464 итальянских МСП за период 2015-2017 годов. Набор данных включает информацию о 464 итальянских МСП-компаниях за период с 2015 по 2017 год. Результаты показали, что подход HRF превосходит традиционную упорядоченную пробит-модель, подчеркивая, что продвинутые методы оценки с использованием техник машинного обучения могут быть успешно применены для прогнозирования кредитного риска МСП, особенно при высоких асимметриях информации. Кроме того, авторы использовали значения Шэпли (согласно модели, предложенной Лойдом Шепли в 1953 году, это априорная мера ожиданий участника предстоящего розыгрыша, выраженная в числовом виде) для оценки важности каждой переменной в прогнозировании кредитного риска МСП. Таким образом, они смогли не только улучшить точность оценки кредитного риска, но и объяснить, какие факторы наиболее важны в этом процессе.

5.3 Управление кредитными рисками в условиях пандемии COVID-19

Марта Рамос Гонсалес, Антонио Партал Урена и Пилар Гомес Фернандез- Агуадо [5] изучали влияние пандемии COVID-19 на риск управления финансовыми учреждениями, особенно в контексте дефолта по ипотечным кредитам. Они используют методы машинного обучения для прогнозирования ожидаемых потерь от дефолта (Expected Loss Best Estimate, ELBE) в 2022 году, с учетом и без учета влияния пандемии. Авторы обнаружили, что пандемия негативно влияет на высокорисковые портфели. Они предлагают метод для оценки ожидаемых и неожиданных потерь в любых экстраординарных событиях.

В статье также упоминается о том, что применение машинного обучения для оценки кре-

дитного рейтинга и скоринга стало предметом недавних исследований. Несмотря на то, что около 10% европейских банков уже используют машинное обучение для расчета регулятивного капитала, авторы отмечают разрыв между академической литературой и банковской практикой в использовании методов машинного обучения для этих целей. В качестве входных данных для прогнозирования ELBE авторы использовали наблюдаемый уровень безработицы в Испании и его прогнозы, выданные Банком Испании, а также некоторые исторические данные, связанные с ипотечным портфелем и спецификой банка. Авторы подчеркивают, что представленный в статье подход является новаторским и может быть адаптирован любым кредитным учреждением для прогнозирования оценок, связанных с внутренними рейтинговыми системами. Следует отметить, что данный подход эффективно справляется с проблемами применения методов машинного обучения, которые мы выделили ранее.

5.4 Кредитный скоринг в условиях социального кредитования

Юань Вана, Янбо Чжан, Мэнкунь Лян, Руйсюэ Юань, Джи Фэнга и Цзюнь Ву [11] рассматривали проблему высокого уровня невозврата студенческих образовательных кредитов, что представляет существенный риск для государства, банков, университетов и самих студентов. Они используют методы машинного обучения для прогнозирования вероятности дефолта по образовательным кредитам. Авторы интегрируют несколько различных моделей машинного обучения для повышения точности прогнозов. Кроме того, они используют интерпретируемый метод SHAP для более глубокого изучения влияния процесса обучения студентов на вероятность дефолта по кредиту. Результаты исследования показывают, что общая сумма стипендий, полученных студентами в течение учебы, значительно влияет на вероятность дефолта. Чем больше стипендий получает студент, тем меньше вероятность того, что он не вернет кредит. Также было обнаружено, что успеваемость студента в университете, отраженная в его среднем балле (GPA), является значимым предиктором дефолта. Кроме того, оказалось, что баллы за вступительные экзамены также могут влиять на риск дефолта. Авторы предлагают использовать полученные результаты для разработки "точных" образовательных программ, которые помогут снизить риск невозврата студенческих кредитов.

Стоит отметить, что в данном исследовании были учтены только характеристики студентов, в то же время есть и другие факторы, влияющие на дефолт по студенческим кредитам, такие как семейная ситуация студента. В будущих исследованиях можно было бы собрать данные о семейных обстоятельствах студентов для более полного анализа.

5.5 Методы предотвращения дефолтов по кредитным картам

Танмай Сринатх и Гурураджа Х.С. [10] исследовали применение машинного обучения для классификации должников по кредитным картам. Они используют различные модели ма-

шинного обучения, которые обычно считаются "черными ящиками" так как не позволяют понять, как они пришли к своим выводам. Основная проблема, которую авторы пытаются решить - это недостаток объяснимости и интерпретируемости в машинном обучении, что приводит к тому, что модели машинного обучения в настоящий момент не могут быть широко использованы в чувствительных областях, таких как диагностика медицинских заболеваний и обнаружение терроризма, поскольку люди не доверяют предсказаниям, не понимая, как они были сделаны. Однако, авторы применяют инструмент DALEX, который помогает интерпретировать результаты этих моделей. Экспериментальные результаты показывают, что модель XGBoost с DALEX дает наилучшие результаты, при этом ее решения основываются на понятной цепочке логики. Они также обсуждают возможные направления для будущих исследований, включая развитие библиотек, подобных DALEX, для большего числа моделей и дальнейшее улучшение способов "заглянуть под капот" моделей машинного обучения.

5.6 Управление кредитными рисками в системе P2P кредитования

Дэвида Мэлоуни, Сунг-Чула Хонга и Барина Н.Нага [7] изучали применение машинного обучения в прогнозировании возврата займов среди заемщиков с низким кредитным рейтингом в системе P2P (peer-to-peer) кредитования. P2P кредитование - это альтернативная форма кредитования, где заемщики и кредиторы взаимодействуют напрямую, минуя традиционные банковские учреждения. Авторы разрабатывают подход на основе машинного обучения для проверки его эффективности и полезности в прогнозировании возврата займов. Они используют бинарный алгоритм классификации для построения модели и применяют его к реальным историческим данным о займах для оценки эффективности. Экспериментальные результаты, визуализации и ключевые показатели производительности обсуждаются в работе, чтобы продемонстрировать надежность предложенного метода. Авторы также обсуждают преимущества использования машинного обучения в P2P кредитовании. В частности, они отмечают, что использование данных, основанных на реальных займах, вместо спекуляций, позволяет улучшить принятие решений. Они также упоминают о возможности использования алгоритмов Байеса для уменьшения смещения и улучшения производительности, ограничивая ошибки классификации. В целом, исследование демонстрирует эффективность применения машинного обучения для решения задачи кредитного скоринга в условиях работы с сильно ассиметричной информацией о кредитоспособности агентов и высокой неопределенностью, что показывает.

6 Результаты

В ходе проведенного исследования были изучены различные методы машинного обучения для оценки кредитных рисков, а также выявлены основные проблемы и ограничения,

связанные с их применением. Однако, несмотря на выявленные сложности, использование данных методов позволяет значительно повысить эффективность и точность оценки кредитных рисков. Результаты исследований показали высокую эффективность применения методов машинного обучения для оценки кредитных рисков. Однако, для более широкого и успешного их использования необходимо продолжать работу по устранению выявленных проблем и ограничений. Наблюдается тенденция использования методов, комбинирующих статистические методы, подкрепленные методами машинного обучения для уточнения результатов и выявления скрытых нелинейных зависимостей. Преимущество данного подхода заключается в том, что он использует в качестве базового метода легко интерпретируемый статистический метод, который широко применяется в индустрии, однако за счет дополнительного обучения модели он способен гарантировать результат, сопоставимый по точности с методами, использующими чисто методы машинного обучения. Вероятно, дальнейшие актуальные исследования в области применения методов машинного обучения для анализа кредитных рисков будут направлены на улучшения работы статистических методов при помощи методов машинного обучения и получении моделей машинного обучения с высокой интерпретируемостью результатов.

Список литературы

- [1] Guina Sotomayor Alzamora, Miguel Romilio Aceituno-Rojo и Henry Ivan Condori-Alejo. “An Assertive Machine Learning Model for Rural Micro Credit Assessment in Peru”. В: *Procedia Computer Science* 202 (2022), с. 301—306.
- [2] Siddharth Bhatore, Lalit Mohan и Y Raghu Reddy. “Machine learning techniques for credit risk evaluation: a systematic literature review”. В: *Journal of Banking and Financial Technology* 4 (2020), с. 111—138.
- [3] Alessandro Bitetto и др. “Machine learning and credit risk: Empirical evidence from small- and mid-sized businesses”. В: *Socio-Economic Planning Sciences* (2023), с. 101746.
- [4] Elena Dumitrescu и др. “Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects”. В: *European Journal of Operational Research* 297.3 (2022), с. 1178—1192.
- [5] Marta Ramos González, Antonio Partal Ureña и Pilar Gómez Fernández-Aguado. “Forecasting for regulatory credit loss derived from the COVID-19 pandemic: A machine learning approach”. В: *Research in International Business and Finance* 64 (2023), с. 101907.
- [6] Tie Li, Gang Kou и Yi Peng. “A new representation learning approach for credit data analysis”. В: *Information Sciences* 627 (2023), с. 115—131.
- [7] David Maloney, Sung-Chul Hong и Barin N Nag. “Two class Bayes point machines in repayment prediction of low credit borrowers”. В: *Heliyon* 8.11 (2022).

- [8] Saba Moradi и Farimah Mokhatab Rafiei. “A dynamic credit risk assessment model with data mining techniques: evidence from Iranian banks”. В: *Financial Innovation* 5.1 (2019), с. 1—27.
- [9] Fabio Sigrist и Nicola Leuenberger. “Machine learning for corporate default risk: Multi-period prediction, frailty correlation, loan portfolios, and tail probabilities”. В: *European Journal of Operational Research* 305.3 (2023), с. 1390—1406.
- [10] Tanmay Srinath и HS Gururaja. “Explainable machine learning in identifying credit card defaulters”. В: *Global Transitions Proceedings* 3.1 (2022), с. 119—126.
- [11] Yuan Wang и др. “National student loans default risk prediction: A heterogeneous ensemble learning approach and the SHAP method”. В: *Computers and Education: Artificial Intelligence* 5 (2023), с. 100166.
- [12] Xu Zhu и др. “Explainable prediction of loan default based on machine learning models”. В: *Data Science and Management* (2023).