

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.383

Сенькович
Дмитрий Сергеевич

Нейросетевая система поддержки принятия решения о выдаче кредита

АВТОРЕФЕРАТ

на соискание академической степени
магистра информатики и вычислительной техники

по специальности 1-40 80 04 – Обработка больших объемов информации

подпись магистранта

Научный руководитель

Жвакина А.В.

к.т.н., доцент

подпись научного руководителя

Минск 2020

КРАТКОЕ ВВЕДЕНИЕ

Одним Залогом выгодного кредитования является оценка потенциального заемщика с точки зрения его платежеспособности и социальной благонадежности. Необходимо проанализировать возможность возврата кредита, опираясь на данные о предыдущей кредитной истории клиента, его платежеспособности и социальных особенностях.

Для решения данной задачи используются различные подходы. Один из них базируется на применении кредитного скоринга. При этом рассматривается информация об имевшихся кредитах, и на основании статистической и математической обработки данных из базы кредитного регистра определяются:

- класс рейтинга;
- скорбалл;
- вероятность просрочки более 90 дней в год на определенную сумму.

Класс рейтинга присваивается на основании оценки в баллах (скорбаллов), характеризующей вероятность того, что кредит не будет возвращен, и вероятности просрочки выплат по кредиту. Чем ниже скорбалл, тем выше вероятность нарушения своевременного погашения кредита. Важными факторами являются время, прошедшее после заключения первого договора о кредитной сделке, количество таких сделок, а также овердрафтовых и потребительских договоров, суммах и продолжительности просрочек.

Недостатком данного подхода является то, что его нельзя использовать по отношению к лицам, которые вообще не брали кредиты или брали их более 5 лет назад, заключали сделки о кредитах с лизинговыми или микрофинансовыми организациями.

Таким образом, чтобы принять решение о выдаче кредита, недостаточно опираться лишь на сведения о кредитном рейтинге, необходимо оценить также множество факторов, влияющих на возврат денег заемщиком: платежеспособность, которую характеризует уровень дохода и надежность компании-работодателя, демографические и социальные данные.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является разработка программного обеспечения с использованием нейронных сетей для поддержки принятия решения о выдаче кредита.

В ходе работы для достижения поставленной цели выполнены следующие задачи:

- выбраны и исследованы алгоритмы машинного обучения и возможность их использования для решения поставленной задачи.
- получены и подготовлены для дальнейшей обработки данные с решениями о выдаче кредитов.
- исследовано качество предсказания и производительность систем принятия решения на основе различных алгоритмов машинного обучения.
- агрегированы данные исследования и сделан вывод о лучшем алгоритме для построения системы.
- разработано программное обеспечение с использованием нейронных сетей для поддержки принятия решения о выдаче кредита.

Объектом исследования являются алгоритмы машинного обучения и их применимость к задаче построения системы принятия решения о выдаче кредита.

Предметом исследования является математическое и программное обеспечение компьютерных систем для поддержки принятия решения о выдаче кредита.

Основной *гипотезой*, положенной в основу диссертационной работы, является возможность использования нейронных сетей в качестве основы для построения систем принятия решений о выдаче кредита. Машинное обучение развивается очень быстро, порядка пяти лет назад глубокие нейронные сети изменили способы построения интеллектуальных систем, в том числе и в финтех секторе. Это значит, что нейронные сети могут стать наилучшим решением для построения систем поддержки принятия решений. Изучая огромные объемы данных, такие системы могут распознавать сложные взаимосвязи, которые человеку увидеть просто невозможно.

Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики

Разработка программного обеспечения поддержки принятия решения о выдаче кредита является одним из самых перспективных направлений в финтех секторе как минимум для стран СНГ, где такие решения:

Внедрены с использованием некоторых простейших статистических методов наподобие логистической регрессии.

Используют выявленные людьми зависимости для принятия решения.\

Отсутствуют вовсе, в частности, отсутствия систем сбора и обработки больших объемов данных.

Такая система могла бы ускорить процесс принятия или отказа решения о выдаче кредита, автоматизировать до некоторой степени этот процесс, оптимизировать затраты предприятия на существующие процессы принятия решения, качественно улучшить результаты принятия решения с учетом сложнейших связей в исторических данных.

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично

Опубликованность результатов диссертации

По теме диссертации опубликованы 2 печатные работа, из них 2 работы в сборниках трудов и материалов международных конференций.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, трех глав, заключения, списка использованных источников и приложений. В первой главе представлен обзор прикладной области, а именно процесса принятия решения о выдаче кредита. Во второй главе представлен обзор различных алгоритмов машинного обучения и математического аппарата, на котором они основываются. В третьей главе представлен обзор решения, полученного в ходе работы, и результатов практических экспериментов, подведение итогов, идеи для дальнейшего развития полученного результата, реализация системы поддержки решения.

Общий объем работы составляет 80 страниц, из которых основного текста – 65 страниц, 33 рисунков на 24 страницах, 4 таблиц на 5 страницах, список использованных источников из 39 наименований на 4 страницах и 2 приложения на 8 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ

Во **введении** определена область и указаны основные направления исследования, показана актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначена практическая ценность работы.

В **первой главе** проведен анализ прикладной области: рассмотрен процесс принятия решения о выдаче кредита, данные, наиболее часто используемые в процессе, правила, влияющие на решение.

Вторая глава посвящена обзору алгоритмов машинного обучения, применимых для построения системы поддержки принятия решения о выдаче кредита. Глава описывает такие алгоритмы машинного обучения как логистическая регрессия, метод опорных векторов, случайный лес и нейронные сети. Вкратце рассмотрим математический аппарат, на котором основывается каждый из методов. Приведены проблемы использования методов и некоторые решения этих проблем.

В **третьей главе** описаны обработка данных для реализации системы поддержки принятия решения о выдаче кредита, исследованы различные модели машинного обучения с различными конфигурациями, в результате чего на основе F1 оценки и точности модели выбран наилучший алгоритм с наилучшей конфигурацией. Реализовано программное средство для поддержки принятия решения о выдаче кредита. Программное средство представляет из себя глубокую нейронную сеть, использующую результаты исследования из третьей главы и основывается на наилучшей конфигурации нейронной сети, дающей наибольшую точность и F1 оценку. Нейронная сеть улучшена для получения еще большей точности в сравнении с наилучшим результатом предыдущей главы.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

Рассмотрена применимость различных алгоритмов машинного обучения к решению задачи создания системы поддержки принятия решения при выдаче кредита. Исследованы различные конфигурации и архитектуры моделей, что позволяет сделать некоторые выводы о наиболее перспективных результатах для использования.

Показано, что нейронные сети дает наилучший результат при нахождении сложных взаимосвязей между данными о клиенте, заявке на кредит и его кредитной истории. Результат оценен несколькими подходами, в результате чего можно сделать вывод о применимости нейронных сетей для построения системы.

Разработана нейронная сеть, которая может быть использована в качестве системы поддержки принятия решения о выдаче кредита. Нейронная сеть позволяет принять решение исходя из данных о клиенте и его заявке в формате, рассмотренной в работе.

Рекомендации по практическому использованию результатов

Полученная в работе модель очень хорошо нашла взаимосвязи в исходных данных, что может означать, что данных недостаточно. Действительно, выборка из 1000 векторов недостаточно и может вызывать значительное переобучение, поэтому для получения наиболее полезных результатов следует значительно увеличить объем обрабатываемых данных.

Собранных характеристик о клиенте и заявке на кредит может не хватать: в работе рассмотрены лишь 18 различных характеристик, в то время как система, используемая банком, может иметь тысячи таких характеристик.

Результаты обучения в значительной мере зависят от региона, в котором собрана информация. Поэтому системы поддержки принятия решения о выдаче кредита должны строиться на данных, как можно лучше описывающих социально-демографическую обстановку в регионе.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1-А. Сенькович, Д.С. Нейросетевая система поддержки принятия банковских решений о выдаче кредитов / Д.С. Сенькович // Big Data and Advanced Analytics: Международная научно-практическая конференция. – 2020. – Часть 1 (20). – с. 358–366.

2-А. Сенькович, Д.С. Нейросетевая система поддержки принятия банковских решений о выдаче кредитов / Д.С. Сенькович // 56-ая научная конференция аспирантов, магистрантов и студентов БГУИР. – 2020.

Библиотека БГУИР