

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

УДК 004.4

Белов
Кирилл Павлович

Оценка кредитоспособности клиента на основе нахождения модели и весов
нейронной сети с применением генетического алгоритма

АВТОРЕФЕРАТ

на соискание степени магистра информатики и вычислительной техники
по специальности 1-40 81 04 Обработка больших объемов информации

Научный руководитель
Нестеренков Сергей Николаевич
кандидат технических наук, доцент

Минск 2018

Введение

В условиях быстро растущих потребительских способностей людей увеличивается оборот предоставляемых товаров и услуг, перед компаниями занимающимися продажами продукции встают вопросы об автоматизации своей деятельности для извлечения наибольшей выгоды из самого процесса продажи продукции. В современном обществе можно иметь высокие доходы, если вести свою деятельность быстро используя хорошо отлаженные процессы работы с клиентами, тем самым оставляя у них хорошее впечатление о вашей компании, что позволит им возвращаться к вам снова и снова [1].

Потребность в современных методах обработки информации при помощи вычислительной техники во всех сферах производства и услуг является бесспорной.

Необходимость прогнозирования, анализа и получения выводов, исходя из накопленных данных приводит к поискам схожих методов анализа с целью их массового последующего применения. Таким образом, автоматизация деятельности обусловлено общим подходом к организации статистических наблюдений, определению статистических показателей и получению на основе этих данных прогнозов для дальнейшей работы.

Внедрение современного программного обеспечения позволяет решить такие вопросы, как:

- Автоматизации.
- Снижение трудоемкости работы.
- Повышение оперативности и достоверности данных.
- Создание предпосылок для дальнейшего совершенствования существующих программных средств.
- Широкое использование экономико-математических методов для получения более состоятельного экономического анализа.

Таким образом, на текущий момент лидирующей сферой деятельности является сфера обработки информации. Постоянное увеличение объемов информации, требующей обработки, способствует повышению требований ко времени обработки, а также ее качеству. В соответствии с этим появляются все новые и развиваются уже созданные программные средства автоматизированной обработки информации.

Актуальность темы подтверждается отсутствием на данный момент внедренных или находящихся в процессе внедрения схожих систем кредитования, которые были бы широкодоступны и помогали обычным гражданам в подборе подходящего для них по условиям и размеру кредита без необходимости обращаться к специалистам.

Краткое содержание работы

В начале производится анализ предметной области с целью определения дальнейшего плана разработки и тестирования программного обеспечения, а также составления наиболее полных требований к будущему программному продукту.

Производится исследование проблемы на обобщенном примере. Далее идет обзор аналогов среди уже имеющегося на данный момент программного обеспечения различных организаций.

Для обеспечения достоверности результатов работы программного средства производится анализ возможных параметров, которые могут оказывать влияние на итоговую сумму кредитования, а также его продолжительность, происходит формирование требований для будущего продукта, определяется состав и назначение разработки, определяются требования к входным и выходным данным, временные требования работы программного средства, требования к надежности, а также составу и параметрам.

Во второй главе описываются возможные подходы машинного обучения к решению поставленной ранее задачи. Описываются основные принципы машинного обучения (постановка задачи, способы машинного обучения и линейная регрессия), после чего производится описание инструментов, используемых в машинном обучении (фреймворков, баз данных, аналитических платформ и прочего программного обеспечения). Представляется описание алгоритмов машинного обучения, которые делятся на контролируемые и неконтролируемые. Описывается общий алгоритм работы программного средства, алгоритмов нейронной сети для нахождения возможности выдачи кредита, входные и выходные данные необходимые для обучения. Приводятся параметры генерации данных для обучения, описывается общая суть работы генетического алгоритма, после чего описывается сам алгоритм.

В третьей главе происходит непосредственно генерирование выборки для обучения, описание граничных условий параметров сгенерированной выборки, приводится общее описание работы алгоритма, многомерной оптимизации, структуры нейронной сети.

В четвертой главе описывается исследование полученной концепции, описываются все возможные функции оптимизации, которые были приняты во внимание при обучении сети, возможное количество нейронов, и приводятся выводы.

Заключение выводы

В результате разработки диссертации по теме «Оценка кредитоспособности клиента на основе нахождения модели и весов нейронной сети с применением генетического алгоритма» было реализовано приложение, позволяющее упростить процедуру принятия решения на допустимую сумму кредита основываясь на истории уже прошедших кредитование людей.

Предварительно был произведен анализ литературы по нужной тематике, изучены аналогичные разработки в данной области. В результате были выделены достоинства и недостатки, а также установлено направление для дальнейшей проработки в рамках настоящей дипломной работы.

На основе проведенного исследования можно сделать следующие выводы:

1. Нынешнему обществу необходимо развитие сферы автоматизации предоставления услуг любого рода.
2. Требуется автоматизация процессов выдачи кредитов с целью увеличения оборота капитала и развития экономики.
3. Пользовательский интерфейс аналогов приложения требует развития, чтобы привлечь дополнительную массу клиентов.
4. Необходимо существенное расширение функциональности системы.

Кроме того, были изучены возможности современных библиотек, фреймворков, а также языков программирования и платформ с целью подбора наиболее подходящего варианта. Итогом данного этапа стал выбор основного языка программирования – Python, фреймворка Tensorflow, операционной системы Ubuntu 18.04 LTS в качестве среды разработки и дальнейшего использования приложения и веб-сервера Nginx.

Были изучены принципы создания генетических алгоритмов, их работа и применение. Успешно реализован генетический алгоритм расчета показателя кредитования.

По окончанию этапа проектирования системы была получена архитектура приложения, позволяющая реализовать гибкую и достаточно легковесную систему с возможностью расширения по необходимости.

По завершению анализа и проектирования было разработано программное средство, включающее в себя:

1. Модуль работы с базой данных.
2. Модуль бизнес-логики, через который осуществляется взаимодействие с системой, и который организует основные процессы серверной части при передаче управления от клиента (от модуля графического интерфейса).

3. Модуль генерирования псевдослучайной выборки данных для обучения системы.

4. Модуль генетического алгоритма для начального обучения системы.

Для обеспечения качества был разработан ряд тест-кейсов. По завершению этапа разработки было успешно выполнено 80% тест кейсов. Все найденные ошибки были успешно устранены. По результату финального тестирования все тесты были пройдены успешно, что говорит о высоком качестве разработанного ПС.

В итоге реализации удалось получить гибкое решение как с точки зрения разработчиков, так и с точки зрения пользователей системы. Одними из главных достоинств системы является ее легковесность, доступность, гибкость и простота.

Разработанное программное средство позволяет значительно ускорить время принятия решения конечным пользователем и снять нагрузку с персонала обслуживающей компании.

Список опубликованных работ

1) Нестеренков, С. Н. Генетический алгоритм как способ решения задач многомерной оптимизации / С. Н. Нестеренков, К. П. Белов // Информационные технологии и системы 2017 (ИТС 2017) = Information Technologies and Systems 2017 (ITS 2017) : материалы междунар. науч. конф. (Республика Беларусь, Минск, 25 октября 2017 года) / редкол. : Л. Ю. Шилин [и др.]. - Минск : БГУИР, 2017. - С. 196 - 197.

2) Нестеренков, С. Н. Использование генетического алгоритма для нахождения весовых коэффициентов нейронной сети в финансовом секторе / С. Н. Нестеренков, К. П. Белов // Информационные технологии и системы 2017 (ИТС 2017) = Information Technologies and Systems 2017 (ITS 2017) : материалы междунар. науч. конф. (Республика Беларусь, Минск, 25 октября 2017 года) / редкол. : Л. Ю. Шилин [и др.]. - Минск : БГУИР, 2017. - С. 198 - 199.

3) Нестеренков, С. Н. Модифицированный генетический алгоритм для обучения нейронной сети / С. Н. Нестеренков, К. П. Белов // Информационные технологии и системы 2017 (ИТС 2017) = Information Technologies and Systems 2017 (ITS 2017) : материалы междунар. науч. конф. (Республика Беларусь, Минск, 25 октября 2017 года) / редкол. : Л. Ю. Шилин [и др.]. - Минск : БГУИР, 2017. - С. 204 - 205.

4) Нестеренков, С. Н. Использование генетического алгоритма для нахождения коэффициентов и структуры нейронной сети / С. Н. Нестеренков, К. П. Белов // Информационные технологии и системы 2018 (ИТС 2018) = Information Technologies and Systems 2018 (ITS 2018) : материалы международной научной конференции, Минск, 25 октября 2018 г. / Белорусский государственный университет информатики и

радиоэлектроники ; редкол. : Л. Ю. Шилин [и др.]. – Минск, 2018. – С. 124
- 125.