

УДК 004.8

***ИСПОЛЬЗОВАНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРЕДСКАЗАНИЯ
КРЕДИТНОГО РЕЙТИНГА ОРГАНИЗАЦИЙ***

Киселев В.В.

к.т.н., доцент,

*Московский Государственный Технический Университет им. Н.Э. Баумана,
Москва, Россия*

Федоров И.А.

магистрант,

*Московский Государственный Технический Университет им. Н.Э. Баумана,
Москва, Россия*

Аннотация

В статье рассматриваются задача присвоения кредитного рейтинга организациям с помощью методов машинного обучения. Используются методы случайного леса и искусственной нейронной сети. В статье предлагается решить задачу этими методами и сравнить результаты с помощью таких метрик как аккуратность, MAE, MSE.

Ключевые слова: машинное обучение, нейронные сети, кредитный рейтинг.

***USING MACHINE LEARNING TO PREDICT THE CREDIT RATING OF
ORGANIZATIONS***

Kiselev V. V.

Ph. D., associate Professor, Bauman

Moscow State Technical University,

Moscow, Russia

Fedorov I.A.

master's student, Bauman

*Moscow State Technical University,
Moscow, Russia*

Abstract

The paper deals with the task of assigning credit rating to organizations using machine learning methods. Random forest and artificial neural network methods are used. The article proposes to solve the problem using these methods and compare the results using such metrics as accuracy, MAE, MSE.

Key words: machine learning, neural networks, credit rating.

Кредитные рейтинги служат важным инструментом для оценки финансового состояния организаций, привлечения капитала, управления рисками и повышения доверия на финансовых рынках. Они основываются на анализе предшествующей и текущей финансовой деятельности субъектов рынка, а также на оценке величины их активов и принятых долговых обязательств. Основная цель таких оценок заключается в предоставлении потенциальным кредиторам и инвесторам информации о вероятностной оценке исполнения субъектами своих финансовых обязательств в установленные сроки. В последние годы область применения кредитных рейтингов значительно расширилась: они начали использоваться для расчета страховых премий, установления условий аренды, а также для оценки надежности потенциальных работников при приеме на работу [2].

Организации получают кредитные рейтинги по ряду причин:

1. Оценка кредитоспособности. Кредитный рейтинг представляет собой независимую оценку способности организации своевременно выполнять свои финансовые обязательства перед кредиторами и инвесторами. Это важный индикатор финансового здоровья компании.

2. Доступ к рынку капитала. Организации с высокими кредитными рейтингами имеют больше шансов привлечь капитал через выпуск облигаций

Дневник науки | www.dnevniknauki.ru | СМИ Эл № ФС 77-68405 ISSN 2541-8327

или получение кредитов на выгодных условиях. Высокий рейтинг сигнализирует о низком уровне риска для инвесторов и кредиторов.

3. Управление рисками. Инвесторы и кредиторы используют кредитные рейтинги для управления своими портфелями рисков. Рейтинги помогают определить, какие активы следует включить в портфель, а какие избегать.

4. Регулирование и требования законодательства. Во многих странах регуляторы требуют, чтобы финансовые учреждения имели минимально допустимые кредитные рейтинги для участия в определённых видах операций. Например, банки должны соблюдать нормативы достаточности капитала, основанные на кредитных рейтингах.

5. Международная торговля и инвестиции. Для международных организаций кредитные рейтинги играют важную роль в установлении деловых отношений с иностранными партнёрами. Высокий рейтинг повышает доверие иностранных контрагентов и облегчает доступ к международным рынкам.

6. Повышение прозрачности. Присвоение кредитных рейтингов способствует повышению прозрачности финансового состояния организации. Это особенно важно для публичных компаний, акции которых торгуются на фондовом рынке.

7. Конкуренентоспособность. Компании с хорошими кредитными рейтингами часто имеют конкурентное преимущество на рынке, поскольку инвесторы и кредиторы предпочитают работать с организациями, имеющими стабильное финансовое положение [4].

Задача присвоения кредитного рейтинга является частным случаем задачи небинарной классификации. Исторически, кредитные рейтинги определяются специализированными агентствами и являются консенсусным мнением финансовых экспертов.

В современном мире, благодаря развитию технологий машинного обучения и увеличению вычислительных мощностей, к решению данной задачи

Дневник науки | www.dnevniknauki.ru | СМИ Эл № ФС 77-68405 ISSN 2541-8327

стали применять различные методы науки о данных. При должном внимании к первоначальной обработке и разметке данных, а также при правильном обучении на достаточном массиве данных об уже присвоенных кредитных рейтингах, мнение модели можно считать, в каком-то смысле, более объективным и непредвзятым.

Помимо решения проблемы конфликта интересов, машинное обучение:

- 1) позволяет автоматизировать сбор и обработку больших объемов данных, что значительно сокращает время на анализ;
- 2) способно работать с огромными массивами данных включая неструктурированные источники информации, такие как новости, социальные сети и отраслевые отчеты, что помогает получить более полное представление о компании и ее деятельности;
- 3) может использовать методы feature selection, чтобы автоматически определить наиболее значимые факторы, влияющие на кредитоспособность компании;
- 4) требуют меньше человеческих ресурсов для обработки данных, что снижает затраты на содержание аналитического отдела.

В данной статье для решения задачи присвоения кредитного рейтинга будут использованы метод случайного леса и метод искусственной нейронной сети.

Основная идея метода случайного заключается в использовании большого ансамбля решающих деревьев, каждое из которых само по себе даёт очень невысокое качество классификации, но за счёт их большого количества результат получается приемлемым. Классификация объектов проводится путём голосования: каждое дерево ансамбля (леса) относит классифицируемый объект к одному из классов, а побеждает класс, за который проголосовало наибольшее число деревьев [3]. Оптимальное число деревьев подбирается таким образом, чтобы минимизировать ошибку классификатора на тестовой выборке. В случае

её отсутствия, минимизируется оценка ошибки на не вошедших в обучающий набор образцах.

Метод искусственной нейронной сети (ИНС) заключается в правильном построении и обучении этой нейросети — математической модели, а также её программного воплощения. Нейросети строятся по принципу организации биологических нейронных сетей — сетей нервных клеток живого организма. В данной статье использовалась полносвязная нейронная сеть, концептуальную модель которой можно увидеть на рисунке 1 [1].

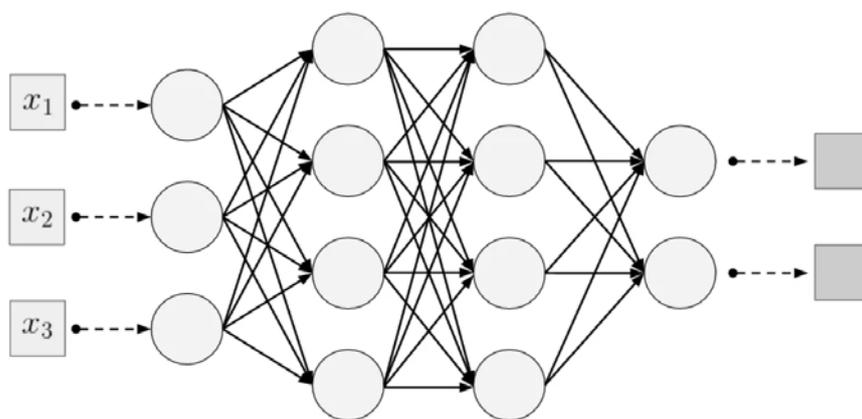


Рис. 1. Полносвязная нейронная сеть

В качестве обучающей и тестовой выборки был выбран заранее размеченный набор данных [5]. В качестве основного инструмента был выбран язык программирования Python, а также его библиотеки NumPy, SciPy, Tensorflow.

После обучения моделей на обучающей выборке, были проведены эксперименты на тестовой выборке в 10 000 организаций. Каждой из них был присвоен определенный кредитный рейтинг, а также были выделены главные признаки, влияющие на повышение или понижение рейтинга. В данном наборе данных ими оказались прибыльность, размер существующих долгов и соотношение долга к собственным средствам.

Если говорить о временных показателях, то обучение модели заняло около 1,5 часов в случае нейронной сети и около 40 минут для метода случайного леса. Вычисление рейтингов на тестовой выборке заняло 6 минут 35 секунд для метода случайного леса и 15 минут 12 секунд для нейронной сети. Оба результата существенно превосходят во времени возможности классических методов, поскольку в настоящее время для присвоения кредитного рейтинга только для одной организации затрачивается около полугода.

Кроме того, были замерены метрики точности обучения: MSE (англ. mean square error, средняя квадратичная ошибка), MAE (англ. mean absolute error, средняя абсолютная ошибка) и аккуратность (accuracy). Значения метрик можно посмотреть в таблице 1:

Таблица 1. Метрики качества обучения

	Accuracy	MAE	MSE
Случайный лес	0.87	0.17	0.26
ИНС	0.81	0.23	0.34

Библиографический список:

1. Галушкин А. И. Нейронные сети. — М.: Большая российская энциклопедия, т. 35, 2017.
2. Cantor R., Packer F. The Credit Rating Industry. FRBNY Quarterly Review/Summer-Fall, 1994.
3. Deng H., Runger G., Tuv E. Bias of importance measures for multi-valued attributes and solutions. Proceedings of the 21st International Conference on Artificial Neural Networks (ICANN), 2011.
4. Fitch T.P. Rating industrial securities. Special Bulletin No. 41. New York: The Fitch Publishing Company, 1920.
5. Starter: Corporate Credit Rating 724a4d11-b [Электронный ресурс] — Режим доступа — URL: <https://www.kaggle.com/code/kerneler/starter-corporate-credit-rating-724a4d11-b/input> (Дата обращения 20.12.2024)

Оригинальность 75