

Петрова А. Н., Фролов Д. О.
A. N. Petrova, D. O. Frolov

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ СИСТЕМЫ ПРОГНОЗИРОВАНИЯ, ИСПОЛЬЗУЮЩЕЙ РЕЛЕВАНТНЫЙ ПОИСК В СИСТЕМАХ БОЛЬШИХ ДАННЫХ

USE OF NEURAL NETWORKS FOR A PREDICTION SYSTEM USING RELEVANCE SEARCH IN BIG DATA SYSTEMS

Петрова Анна Николаевна – кандидат технических наук, заведующая кафедрой «Проектирование, управление и развитие информационных систем» Комсомольского-на-Амуре государственного университета (Россия, Комсомольск-на-Амуре). E-mail: PetrovaAN2006@yandex.ru.

Anna N. Petrova – PhD in Engineering, Head of the Department «Design, Management and Development of Information Systems», Komsomolsk-na-Amure State University (Russia, Komsomolsk-on-Amur). E-mail: PetrovaAN2006@yandex.ru.

Фролов Дмитрий Олегович – аспирант Комсомольского-на-Амуре государственного университета (Россия, Комсомольск-на-Амуре). E-mail: optcompanys@mail.ru.

Dmitriy O. Frolov – Graduate Student, Komsomolsk-na-Amure State University (Russia, Komsomolsk-on-Amur). E-mail: optcompanys@mail.ru.

Аннотация. В настоящее время искусственные нейронные сети (ИНС) успешно решают множество задач, используя различные виды информации в качестве входных данных. Однако возникают сложности в обучении ИНС из-за возможного использования неподходящих наборов данных с большим количеством параметров. Одна из ключевых проблем подготовки данных – оценка их эффективности для обучения конкретной ИНС. В данном исследовании проведён анализ методов обучения искусственных нейронных сетей, включая разнообразные подходы к оптимизации параметров при подготовке данных для обучения ИНС. Также изучено влияние выбора метода оптимизации и количества параметров нейронной сети на эффективность обучения. Главная цель обучения состоит в достижении баланса между способностью сети правильно реагировать на входные данные, использованные в обучении (запоминание), и способностью выдавать правильные результаты на входные данные, аналогичные, но не идентичные тем, что использовались в обучении (принцип обобщения).

Summary. At present, artificial neural networks (ANNs) effectively tackle numerous tasks utilizing diverse types of information as input vectors. Nevertheless, there exists a challenge in training ANNs due to the potential utilization of improperly prepared datasets containing a vast array of parameters. Evaluating the efficacy of data preparation in training a particular ANN stands as one of the primary issues. The analysis in the paper delves into training methods for artificial neural networks, examining various approaches to optimizing parameters during data preparation for training an ANN. The study also scrutinizes the impact of optimization method selection and the quantity of neural network parameters on learning efficiency. The objective of training is to strike a balance between the network's capacity to provide accurate responses to input data employed during training (memorization) and its ability to yield correct outcomes when faced with input data that is akin but not identical to that used during training (the principle of generalization).

Ключевые слова: искусственные нейронные сети, алгоритмы обучения, анализ данных.

Key words: artificial neural networks, learning algorithms, data analysis.

УДК 517.95

Введение. Перед началом подготовки данных важно изучить различные методы обучения нейронных сетей. С увеличением сложности задач, а также нелинейным ростом объёмов и размерности данных, становится критически важным повышение эффективности алгоритмов обучения нейронных сетей. Существует несколько алгоритмов, таких как метод обратного распростра-

нения, метод упругого распространения и генетический алгоритм обучения. Однако все они основаны на двух основных принципах: обучении с учителем и без него.

В процессе добычи руд и песков драгоценных металлов компания ООО «Золото» нуждается в информации от своих сотрудников, основанной на их опыте. Возникает необходимость заменить экспертное мнение опытных сотрудников на прогноз, полученный от обученной нейронной сети на основе данных из системы больших данных предприятия. Для улучшения качества обработки и снижения затрат на добычу драгоценных металлов необходимо внедрить систему прогнозирования процесса добычи, использующую соответствующий поиск во внутренних системах больших данных предприятия.

Подготовка данных для обучения нейронных сетей занимает значительное количество времени из-за сложности процессов ручной обработки информации. Объём поступающих данных избыточен и разнообразен. Полученные данные хранятся в корпоративной информационной системе в виде таблиц с различными типами данных и большим количеством параметров процесса добычи. Методы формирования обучающих и тестовых выборок зависят от поставленной задачи: для классификации данные разделяются таким образом, чтобы численное соотношение объектов разных классов в полученных наборах было сопоставимо с исходной генеральной совокупностью; для регрессионного анализа необходимо обеспечить равномерное распределение целевой переменной в полученных наборах, используемых для обучения и проверки качества.

Алгоритм обратного распространения ошибки представляет собой метод обучения многослойных перцептронов, т. е. нейронных сетей с множеством слоёв, включая скрытые. При этом процессе обучения необходимо иметь заданные целевые значения для примеров. Основным принципом этого алгоритма заключается в использовании ошибки на выходе нейронной сети для вычисления коррекции весов нейронов в её скрытых слоях:

$$E = \frac{1}{2} \sum_{i=1}^k (y - y')^2,$$

где k – число выходных нейронов сети; y – целевое значение; y' – фактическое выходное значение.

Обобщённый процесс обучения включает в себя множество примеров и сложную сеть нейронов. На каждом этапе тренировки происходит два прохода через нейронную сеть: сначала данные передаются от входного к выходному слою, затем выходные значения обратно распространяются для коррекции весов. Подготовка данных для обучения также важна. Она включает в себя сбор и обработку информации для создания набора данных, который соответствует требованиям алгоритмов машинного обучения. Хотя процесс обучения ориентирован на определённые цели, такие как классификация, уменьшение количества параметров с помощью корреляции данных также является важным шагом перед реализацией метода:

$$r_{xy} = \frac{\sum(x_i - \bar{x}) * (y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 * \sum(y_i - \bar{y})^2}}$$

где x_i – значение переменной X ; y_i – значение переменной Y ; \bar{x} – среднее арифметическое для переменной X ; \bar{y} – среднее арифметическое для переменной Y .

В результате количество параметров для обучения снизится с 16 до 9. Проверяем данные на разных эпохах обучения. Результаты приведены в табл. 1.

Значения потери – функции и точность нейронной модели – приведены в табл. 2 для уменьшенного количества параметров.

По результатам проведённой валидации можно сделать вывод об успешном обучении нейронной модели. Дальнейшая проверка качества обучения тесно связана с прогнозированием работы на обученной модели Sequential и последующим сравнением с ручными расчётами.

Прогноз формируется на основе предыдущих данных из внутренней системы, осуществляющей релевантный поиск в больших данных. Затем с помощью обученной нейронной сети оценивается успешность выполнения поиска данных. Рассчитанные значения добавляются к исходной

выборке данных, и далее производится прогноз с учётом предыдущих значений. Этот процесс повторяется до достижения заданного периода прогнозирования.

Таблица 1

Исследование до и после корреляции данных

Этап	Количество строк на каждом шаге до корреляции	Количество строк на каждом шаге после корреляции
Базовая	6024	6024
1 шаг	4138	3877
2 шаг	2239	1793
3 шаг	1421	712
4 шаг	648	97

Таблица 2

Качество обучения модели на уменьшенном количестве параметров

Потеря		Точность	
Обучение	Валидация	Обучение	Валидация
0,0482	0,0969	0,9893	0,9897

Важно отметить, что на каждом шаге прогнозирования значения управляющих параметров считаются известными и передаются на вход нейронной сети. Это позволяет производить дальнейшее изменение их значений в поисках оптимального режима работы.

После окончательного обучения модели мы вычисляем ошибку прогнозирования для каждой строки в тестовых данных в виде абсолютной процентной ошибки.

Для оценки результатов рассчитывается процентная ошибка (APE):

$$APE = (\text{оценочное значение} - \text{фактическое значение}) / \text{фактическое значение} \times 100.$$

Пример работы приложения для прогнозирования представлен на рис. 1. Из расчёта процентной ошибки видно, что на данном наборе данных процентная ошибка не превышает установленного порога в 5 %.

relationship	PredictedRelationship	APE
0.915874	0.902503	1.459902
0.951825	0.906832	4.726996
0.968373	0.940506	2.877753
0.916418	0.903856	1.370751
0.960972	0.942439	1.928556

Рис. 1. Расчёт процентной ошибки

Для оценки точности прогнозирования сети было проведено сравнение с данными, полученными в результате ранее осуществлённого релевантного поиска в системах больших данных. В результате этого сравнения была построена гистограмма, иллюстрирующая сопоставление оценок, полученных от сотрудника, и прогнозов, сделанных системой прогнозирования (см. рис. 2).

ПРОГНОЗИРОВАНИЕ

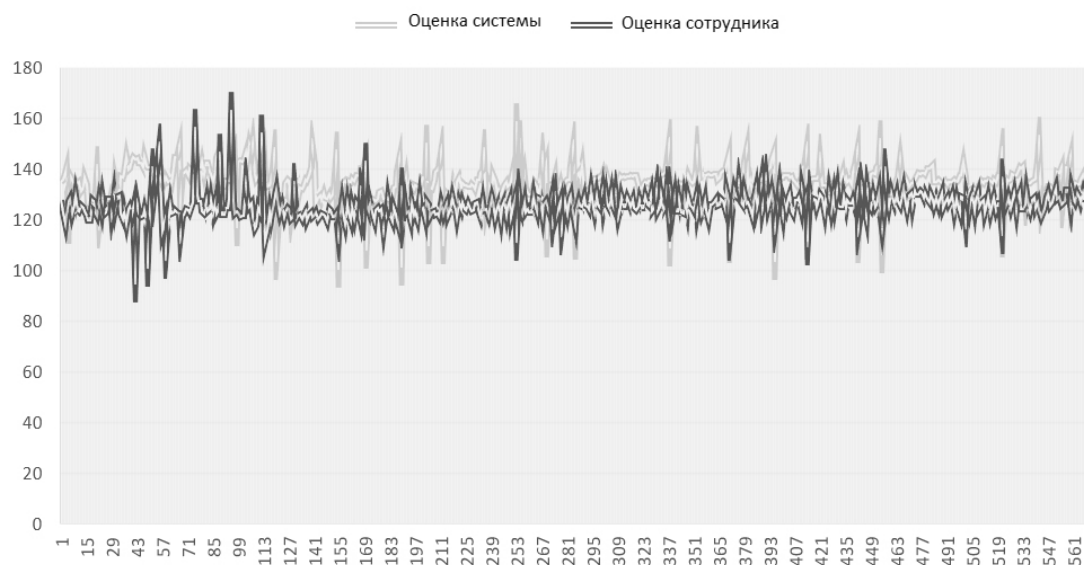


Рис. 2. График прогноза

Качество прогнозирования оценивалось по величине средней квадратичной ошибки.

Заключение. В рамках проекта были созданы как математическая, так и компьютерная модели, обученные для релевантного поиска в структурах больших данных для обнаружения аномалий в процессе добычи драгоценных металлов. Данные, полученные в результате поиска, существенно влияют на прогнозирование эффективности обнаружения аномалий. Также были предложены методы уменьшения объёма данных с использованием корреляции. После анализа процентной ошибки стало ясно, что прогнозируемые значения в среднем отличаются от фактических на 3 %.

ЛИТЕРАТУРА

1. Гудфеллоу, Я. Deep Learning / Я. Гудфеллоу, И. Бенджио, А. Курвилль; пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2018. – 652 с.
2. Хант, Э. Python и машинное обучение / Э. Хант. – М.: ДМК Пресс, 2019. – 414 с.
3. Шолле, Ф. Глубокое обучение на Python / Ф. Шолле. – СПб.: Питер, 2018. – 400 с.
4. Шолле, Ф. Программирование нейронных сетей на Python / Ф. Шолле. – М.: ДМК Пресс, 2019. – 403 с.
5. Нильсен, М. Нейронные сети и глубокое обучение / М. Нильсен. – М.: ДМК Пресс, 2019. – 308 с.
6. Аршакян, А. И. Применение нейронных сетей для анализа больших данных в бизнесе / А. И. Аршакян, А. С. Пугачев, В. Ю. Варгулёв // Мир науки, культуры, образования. – 2018. – № 4. – С. 36-38.
7. Брэдбери, С. Глубокое обучение с PyTorch / С. Брэдбери. – М.: ДМК Пресс, 2019. – 254 с.
8. Рашка, С. Python и машинное обучение / С. Рашка. – М.: ДМК Пресс, 2016. – 418 с.
9. Чарки, Ш. Глубокое обучение и TensorFlow / Ш. Чарки. – М.: ДМК Пресс, 2018. – 480 с.
10. Применение методов машинного обучения в задачах оценки технологических процессов / О. В. Попова, Я. Ю. Григорьев, Е. П. Жарикова, А. Л. Григорьева // Учёные записки Комсомольского-на-Амуре государственного технического университета. Науки о природе и технике. – 2021. – № VII (55). – С. 68-72.
11. Попова, О. В. Моделирование оценки параметров в технологических задачах / О. В. Попова, А. В. Кириллов // Наука, инновации и технологии: от идей к внедрению: материалы Международной научно-практической конференции, Комсомольск-на-Амуре, 07-11 февраля 2022 года. – Комсомольск-на-Амуре: ФГБОУ ВО «КнаГУ», 2022. – С. 244-246.
12. Кизянов, А. О. Корреляция Пирсона на языке программирования Python / А. О. Кизянов // Постулат. – 2018. – № 8 (34). – С. 14.
13. Рекуррентные нейронные сети и методы оптимизации проектных технологических процессов в АСТП машиностроительного производства / С. Н. Поезжалова [и др.] // Вестник Уфимского государственного авиационного технического университета. – 2011. – Т. 15. – № 5 (45). – С. 36-46.