

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ
Н.Г.ЧЕРНЫШЕВСКОГО»**

Кафедра теории функций и стохастического анализа

Нейронные сети в портфельном инвестировании

АВТОРЕФЕРАТ

студента (ки) 2 курса 248 группы

направления 09.04.03 «Прикладная информатика»

механико-математического факультета

Кашкиной Марии Владимировны

Научный руководитель
доцент, к.ф.-м. наук

Шаталина А.В.

Зав. кафедрой
доцент, д.ф.-м. наук

Сидоров С.П.

Саратов 2019

Введение. В магистерской работе изучаются вопросы, связанные с анализом временных рядов, а именно использование искусственных нейронных сетей для прогнозирования экономических показателей с учетом специфики данных.

Актуальность работы связана с широким применением искусственных нейронных сетей в различных сферах деятельности общества, таких, как автоматизация процессов распознавания образов, адаптивное управление, аппроксимация функционалов, прогнозирование, создание экспертных систем, организация ассоциативной памяти и др.

Целью магистерской работы является изучение теоретического материала и применение его для оптимизации бизнес-процесса купли-продажи топлива в организации.

В соответствии с данной целью были поставлены следующие **задачи**:

- 1) изучить различные методы прогнозирования;
- 2) изучить архитектуру нейронных сетей;
- 3) проанализировать инструментальные средства для реализации задачи прогнозирования;
- 4) разработать модель искусственных нейронных сетей для прогнозирования экономических показателей топливной компании.

Объектом данной работы является бизнес-процесс закупки в компании.

Предметом работы является построение модели искусственной нейронной сети для прогнозирования экономических показателей топливной компании.

Структуру магистерской работы можно разделить на два больших раздела: теоретический и практический.

Теоретическая часть состоит из двух параграфов. В первом параграфе изложены результаты теоретического исследования современных методов прогнозирования, выполнен анализ методов прогнозирования и выбор наилучшей модели для решения задачи прогнозирования. Второй параграф посвящен

инструментальным средствам для моделирования нейросетевых моделей, рассмотрены особенности нейронных сетей и алгоритмов обучения, определен программный комплекс для реализации задачи прогнозирования.

В практической части приведен анализ предприятия и его отдельных бизнес-процессов, выявлены слабые места и предложен процесс оптимизации бизнес-процесса. Спроектирована модель для прогнозирования закупок топлива, разработана математическая модель для прогнозирования показателей ООО "Мегаполис - Топливная группа".

Результаты работы опубликованы в статье Шаталиной А.В., Кашкиной М.В. "Нейронные сети в портфельном инвестировании"[1].

Основное содержание работы. Первый раздел содержит информацию об основных методах прогнозирования, их классификации и применении к различным видам данных.

Основными адаптивными методами прогнозирования являются:

- экспоненциальное сглаживание;
- АРСС (ARIMA, autoregressive integrated moving average);
- регрессионные модели;
- искусственные нейронные сети.

Методы прогнозирования во времени

Одним из главных критериев выбора прогнозирования является временной промежуток, на который данные прогноз ориентирован.

По времени прогнозирования различают:

- краткосрочный;
- среднесрочный;
- долгосрочный прогноз.

Для наглядного представления выбора моделей в зависимости от интересующей длительности прогноза данные были сведены в таблице на рисунке 1.

Метод:	Тип прогноза:		
	Краткосрочный	Среднесрочный	Долгосрочный
АРПСС	+	+/-	-
Экспоненциальное сглаживание	+/-	+	-
Регрессия	+/-	+/-	+
Нейронные сети	+	+	+

Рисунок 1 – Длительность прогноза различных моделей

Во втором разделе изложена история возникновения биологических и искусственных нейронных сетей, проведен сравнительный анализ основных методов обучения, подробно доказана теорема Колмогорова и ее частные случаи.

Искусственный нейрон в первом приближении имитирует свойства биологического нейрона. В соответствии с рисунком 2, от сомы – центральной части нейрона – отходят древовидные отростки - дендриты. Они играют роль рецепторов, то есть проводят сигналы от других нейронов. Задача аксона, самого крупного отростка, заключается в том, чтобы передавать сигнал активности от сомы другим нейронам. Место соединения аксона с дендритами других нейронов разделено малым расстоянием, порядка 200 нм. Этот промежуток называется синапсом.

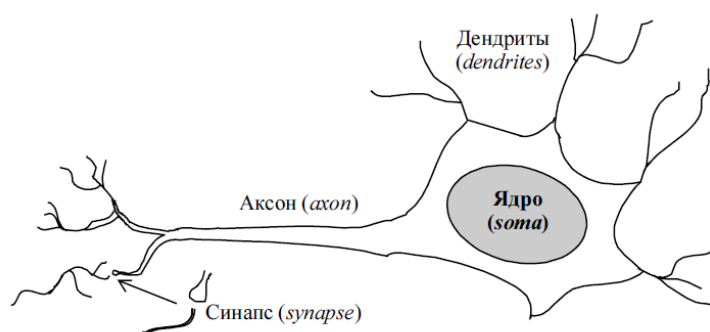


Рисунок 2 – Структура нейрона

Искусственные нейронные сети (ИНС) – это математические модели, которые решают задачу нелинейной глобальной оптимизации. В соответствии

с рисунком 3 множество входных сигналов, обозначенных x_1, x_2, \dots, x_n , поступает на искусственный нейрон. Эти входные сигналы, в совокупности обозначаемые вектором X , соответствуют сигналам приходящим в синапсы биологического нейрона. Каждый синапс характеризуется величиной синаптической связи или ее весом w_i .

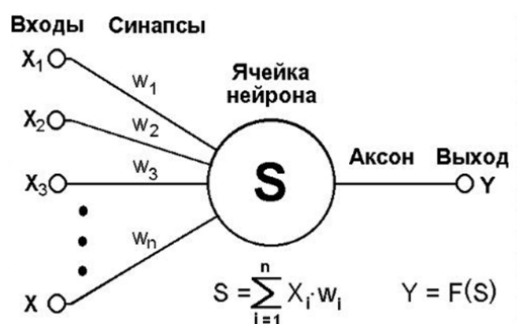


Рисунок 3 – Искусственный нейрон

Каждый сигнал умножается на соответствующий вес w_1, w_2, \dots, w_n , и поступает на суммирующий блок. Каждый вес соответствует «силе» одной биологической синаптической связи. Суммирующий блок, соответствующий телу биологического элемента, складывает взвешенные входы алгебраически, создавая величину S .

Таким образом, текущее состояние нейрона определяется как взвешенная сумма его входов

$$S = \sum_{i=1}^n x_i w_i. \quad (2.1.1)$$

Выход нейрона - это функция его состояния, она имеет вид

$$y = f(s), \quad (2.1.2)$$

где f - активационная функция, моделирующая нелинейную передаточную характеристику биологического нейрона и представляющая нейронной сети большие возможности.

Обучение нейронной сети - это процесс, в котором свободные параметры нейронной сети настраиваются посредством моделирования среды, в

многих переменных достаточно операций суммирования и композиции функций одной переменной [11].

Теорема Колмогорова (основная формулировка). Любая непрерывная функция от n переменных $f(x_1, x_2, \dots, x_n)$ на замкнутом ограниченном множестве может быть представлена в виде

$$f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n+1} \chi_q \left(\sum_{p=1}^n \psi^{pq}(x_p) \right), \quad (2.3.3)$$

где χ_q и ψ^{pq} — непрерывные функции, причем ψ^{pq} не зависят от функции f .

Далее проводилась оценка нейросетевой модели в программном комплексе MATLAB, в котором существует визуальный инструмент для тестирования и оценки искусственных нейронных сетей - Neural Network Start.

Neural Network Start позволяет создать сеть и обучить ее на выбранном массиве данных. После обучения выводится среднеквадратичное отклонение и множественный коэффициент корреляции. Neural Network Start предоставляет возможность выбрать алгоритм обучения из представленных [13].

Для прогнозирования временных рядов представлены следующие **алгоритмы обучения**:

- Levenberg-Marquard;
- Bayesian Regularization;
- Scaled Conjugate Gradient.

В практической части работы была решена задача оптимизации планирования закупок топлива для компании ООО "Мегаполис - Топливная группа" с использованием аналитических инструментов, таких как нейронные сети. Благодаря этому, можно будет спрогнозировать сезонные скачки спроса и заранее договориться с поставщиками на предоставление выделенного количества топлива.

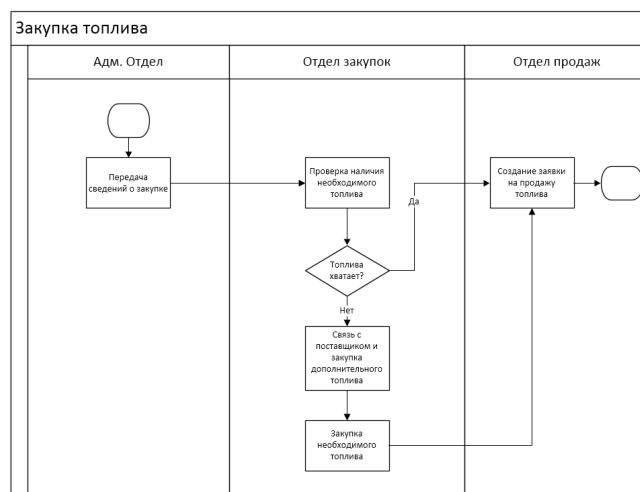


Рисунок 5 – Бизнес-процесс закупки топлива

Была построена и проанализирована модель искусственной нейронной сети для решения поставленной задачи прогнозирования для компании ООО «Мегаполис - Топливная группа».

Построение модели искусственной нейронной сети можно разделить на несколько этапов:

- подготовка репрезентативной выборки;
- тестирование различных моделей;
- анализ полученных результатов;
- написание функционального файла для построения сети.

Для задачи прогнозирования была создана выборка, содержащая количество закупленного топлива по дням за 2018 год. Объем выборки значений 365, что соответствует количеству ежедневных отчетов о поступлениях денежных средств.

В ходе анализа различных моделей искусственных нейронных сетей рассматривались:

- различные типы моделей нейронных сетей (NARX, NAR, NONL);
- различное количество нейронов в модели;
- различная задержка прогнозирования, связанная с особенностью рассматриваемых моделей.

Тестовая выборка была поделена на три части в соотношении 70/15/15 (тренировочную, тестовую и проверочную).

Главной задачей при выборе модели является выбор количества нейронов в слое. В случае временных рядов учитывается задержка, которая показывает период на сколько шагов назад искать связи между входами и выходами,

Тестирование сетей проводилось методом перебора параметров. Значение задержки для тестирования было выбрано от 1 до 10 с шагом единица от 1 до 6 и с шагом два от 8 до 10. Конечное значение 10 было выбрано в соответствии с желанием компании видеть прогноз на 10 дней вперед.

Количество нейронов подбиралось по принципу усложнения сети, т.к. данных для анализа не очень много. Производился перебор от 2 до 8 нейронов. Оценка качества модели заключалась в количественном сравнении значения среднеквадратичной ошибки и значения регрессии.

Наилучшей была выбрана модель NARX при задержке 10 и количестве персептронов на слое 6. Остальные сети показывают относительно большую ошибку, в особенности на тестовом множестве, что означает худшую по сравнению с NARX способность к обобщению [19]

Оценка функции производилась по следующим графикам:

- график регрессии;
- график производительности;
- гистограмма ошибок;
- time series response;
- график автокорреляции.

График регрессии показывает, что отображение сетевых выходов по отношению к исходящим данным для всех трех множеств распределены вдоль линии в 45 градусов. Множественный коэффициент корреляции зафиксирован не ниже 0.92 для всех подвыборок.

График производительности показывает, что ошибки на всех трех множествах стабилизировались (высокая работоспособность модели).

График Time Series Response показывает, что со временем ошибка стабилизируется, и исключаются выбросы на линии.

Гистограмма ошибок выбранной модели показала нормальное распределение вокруг нуля.

График автокорреляции показывает, что корреляция ошибок прогнозирования преимущественно равна нулю. Ошибки больше нуля распределены по нормальному закону, что является хорошим показателем.

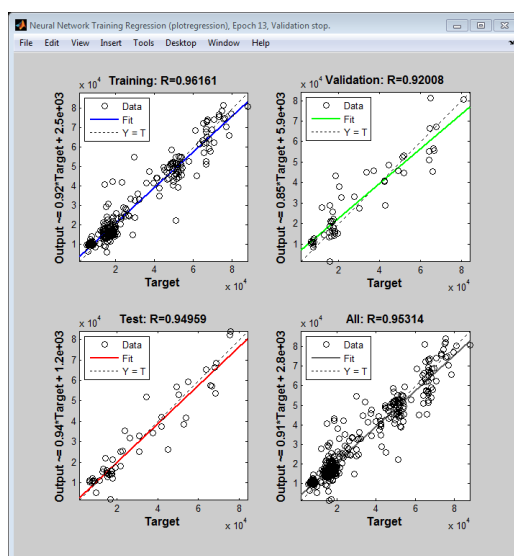


Рисунок 6 – График регрессии для выбранной модели

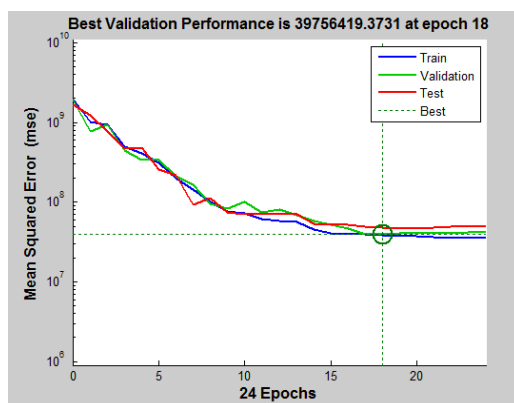


Рисунок 7 – График производительности для выбранной модели

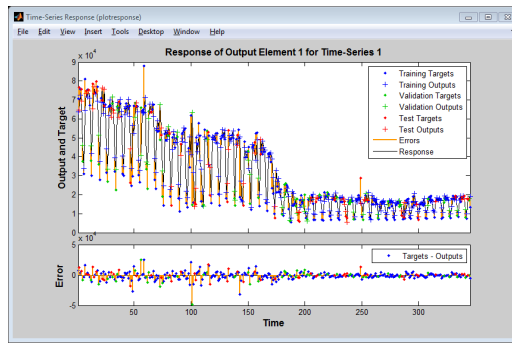


Рисунок 8 – График Time Series Response для выбранной модели

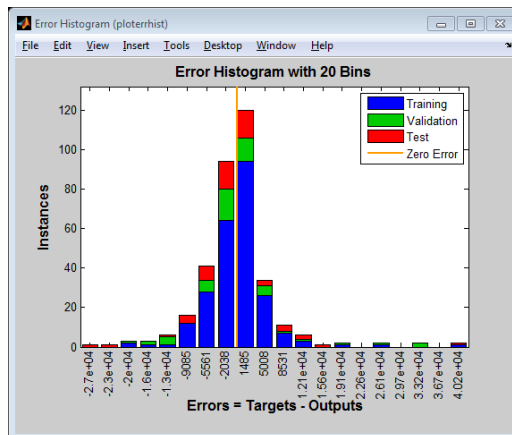


Рисунок 9 – Гистограмма ошибок выбранной модели

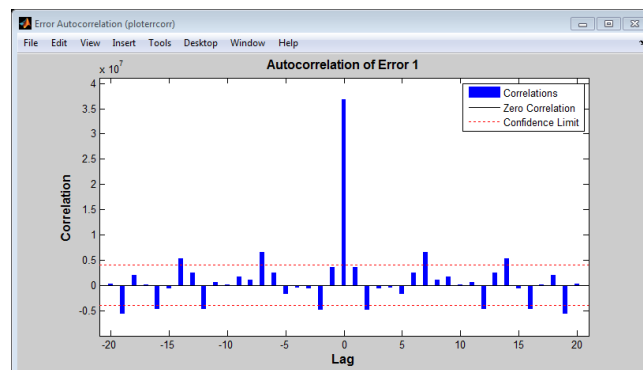


Рисунок 10 – График автокорреляции для выбранной модели

Заключение. В результате выполнения выпускной квалификационной работы проведен анализ методов прогнозирования. Предлагается использование искусственных нейронных сетей в связи с их универсальностью, возможностями к расширению, работой с самыми различными типами данных и возможностью учета различных факторов внешней среды.

Для сбора и подготовки данных была разработана программа ETL-процесса. В качестве инструмента хранения данных использовалась система управления

базами данных MySQL. Полученная база данных позволяет быстро и удобно обрабатывать данные, иметь наглядное представление связи между данными и иметь доступ к работе с запросами для ускорения формирования репрезентативной выборки.

Проанализированы инструментальные средства для реализации задачи прогнозирования. На основе анализа был выбран программный комплекс MATLAB версии R2019b 9.6, для 64-bit-ных систем. Преимуществами данного продукта является большое количество различных моделей и методов, интегрированных в приложение, наличие возможности изменения моделей для конкретных задач, уникальные модели, которые применяются только в MATLAB, в частности сети NARX.

Была разработана модель искусственной нейронной сети для прогнозирования экономических показателей. Данная модель отвечает требованиям компании. Модель построена на основе архитектуры сети NARX и содержит 6 нейронов на скрытом слое и имеет числовое значение задержки 10.

Дальнейшим развитием данной модели может стать новое исследование, направленное не только на анализ закупок, но и размещения заказов, каждого заказчика в отдельности.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Шаталина, А. В., Кашкина, М.В. Нейронные сети в портфельном инвестировании / А.В. Шаталина, М.В. Кашкина // Математическое и компьютерное моделирование в экономике, страховании и управлении рисками: материалы VII Междунар. молодежной науч.-практ. конф. — Саратов: ООО Изд-во “Научная книга”, 2018. — С. 269-272.
- 2 Галушкин, А.И. Нейронные сети: основы теории /А.И.Галушкин. М.: Горячая линия-Телеком, 2010. 496 с.
- 3 Каллан, Р. Основные концепции нейронных сетей. /Р.Каллан. М:Вильямс, 2002. 287 с.
- 4 Аксенов С.В., Новосельцев В.Б. Организация и использование нейронных сетей (методы и технологии)/В.Б.Новосельцев. М.:НТЛ, 2006. - 128 с.
- 5 Барский, А.Б. Нейронные сети: распознавание, управление, принятие решений./ А.Б. Барский. М.: Финансы и статистика, 2004. — 176 с.
- 6 Хайкин, С. Нейронные сети. Полный курс / С. Хайкин. М.: Вильямс, 2006. 1104 с.
- 7 Круг, П.Г. Нейронные сети и нейрокомпьютеры/ П.Г.Круг. М.: МЭИ, 2002. 176 с.
- 8 Каллан, Р. Основные концепции нейронных сетей. /Р.Каллан. М:Вильямс, 2002. 287 с.
- 9 Самарский, А. А. Введение в численные методы. / А.А. Самарский. СПб.: Лань, 2005. 288 с.
- 10 Кирьянов, И. В. Методика формирования непараметрических портфелей./И.В.Кирьянов.// Сибирская финансовая школа, 2011. № 2. с. 78-83.

- 11 Колмогоров, А.Н. О представлении непрерывных функций нескольких переменных в виде суперпозиции непрерывных функций одного переменного. Доклад АН СССР. Т.114. с.953-956.
- 12 Matlab, Third Edition: A Practical Introduction to Programming and Problem Solving / Stormy Attaway Ph.D. Boston University – Butterworth-Heinemann, 2013.
- 13 MATLAB for Neuroscientists, Second Edition: An Introduction to Scientific Computing in MATLAB / Pascal Wallisch and Michael E. Lusignan –Academic Press, 2013.
- 14 1С:Предприятие:Практическое пособие разработчика. Примеры и типовые приемы. / Радченко М.Г., Хрусталева Е.Ю. – М.: 1С-Паблишинг, 2016.
- 15 Теория и практика построение баз данных. / Кренке Д.М.: Питер, 2014.
- 16 Язык запросов "1С:Предприятия 8"/ Хрусталева Е. Ю. – М.: 1С-Паблишинг, 2013.
- 17 Программирование на языке Go. Разработка приложений XXI века/ Марк Саммерфильд – М.: ДМК-Пресс, 2017.
- 18 Трофимова, Е. А., Мазуров, В.Д., Гилёв, Д.В. Нейронные сети в прикладной экономике/ Е.А. Трофимова, В.Д. Мазуров, В.Д. Гилёв. М.:Урал. федер. ун-т, 2017. 96 с.
- 19 Агеев А.Д., Балухто А.Н., Бычков А.В. и др. «Нейроматематика» – Кн. 6 серии Нейрокомпьютеры и их применение, М: ИПРЖР, 2002. – 448 с.
- 20 Artificial Neural Networks Applied For Digital Images with Matlab Code: The Applications of Artificial Intelligence In Image Processing Field Using Matlab. / Othman A. Khfagy, Muhammad Atta – LAP LAMBERT Academic Publishing, 2015.