

# СРАВНЕНИЕ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ<sup>1</sup>

**Иванюк В.А., Комков Н.И., Володина Н.Н.**

*Институт народнохозяйственного прогнозирования РАН, Москва, Россия  
ivaver6@gmail.com, komkov\_ni@mail.ru, nn\_volod@mail.ru*

**Шувалов К.И., Мартынова А.Н.**

*Институт проблем управления имени В.А. Трапезникова РАН, Москва, Россия  
shuval@ipu.ru, nastyaa232510@mail.ru*

*Аннотация. Нейронные сети являются одним из наиболее эффективных методов прогнозирования финансовых временных рядов. Они могут быть использованы для создания более точных прогнозов и помочь трейдерам и инвесторам принимать более осознанные решения на рынке. В данной работе было проведено практическое исследование эффективности нейросетевого моделирования на временных рядах энергетического сектора.*

*Ключевые слова: нейронные сети, прогнозирование временных рядов, моделирование.*

## Введение

Искусственная нейронная сеть — это программная реализация математической модели, которая имитирует поведение биологических нейронных сетей. Одной из ключевых особенностей искусственных нейронных сетей является их способность к обучению и адаптации путем корректировки весовых коэффициентов синаптических связей. Это позволяет им решать сложные проблемы, требующие аналитического мышления [1-3].

## 1. Построение нейронной сети для прогнозирования временных рядов

Рассмотрим шаги, необходимые для построения нейронной сети для прогнозирования временных рядов:

- Подготовка данных.
- Выбор подходящей архитектуры нейронной сети.
- Проектирование: определение количество слоев, нейронов, функций активации и других гиперпараметров.
- Обучение: использование обучающего набора, вычисление ошибки или потери, а также корректировку весов и смещений сети для минимизации потерь.
- Тестирование. Обученная сеть используется для прогнозирования набора тестовых данных, а также оценивается точность прогнозов.

При правильном подходе нейронные сети могут обеспечить точные и надежные прогнозы для задач прогнозирования временных рядов [4-6].

Архитектура нейронных сетей может быть классифицирована на различные типы в зависимости от их структуры. Наиболее распространенные типы архитектур нейронных сетей следующие:

- Нейронные сети с прямой связью. Они состоят из одного или нескольких слоев нейронов, которые обрабатывают входные данные в прямом направлении без каких-либо циклов или обратной связи.
- Свёрточные нейронные сети (CNN). Они предназначены для обработки изображений и приложений компьютерного зрения.
- Рекуррентные нейронные сети (RNN). Чаще всего используются в задачах последовательной обработки данных, таких как обработка естественного языка и распознавание речи. У них есть циклы, которые позволяют информации сохраняться с течением времени и ячейки памяти, в которых хранится контекстная информация.
- Сети с длительной кратковременной памятью (LSTM): являются вариантом RNNs, которые предназначены для решения проблемы исчезающих градиентов. Они используют ячейки памяти и механизмы для выборочного хранения и обновления контекстной информации с течением времени.
- Автокодеры. Они используются для неконтролируемых задач обучения, таких как уменьшение размерности и сжатие данных.
- Генеративные состязательные сети (GAN). Они предназначены для генерации новых данных, аналогичных данным обучения. Они состоят из сети генератора, которая создает новые данные и

<sup>1</sup> Работа поддержанна РНФ 23-41-10001 “Математические модели и компьютерные технологии календарного планирования производства и энергетики в условиях экономической неопределенности”

сети дискриминатора, которая проводит различие между сгенерированными данными и реальными данными.

Главной силой CNN является их способность обнаруживать сложные закономерности в многомерных данных. При применении к данным временных рядов они могут идентифицировать закономерности, такие как периодичность и тенденции. Данная сеть может быть применена как к одномерным, так и к многомерным данным временных рядов.

Основная идея использования CNN для прогнозирования временных рядов состоит в том, чтобы рассматривать временной ряд как одномерное изображение и использовать свёрточные фильтры для извлечения из него признаков. Входные данные для CNN – это последовательность временных шагов, а выходные данные – прогноз для будущего временного шага.

Свёрточная нейронная сеть обычно состоит из трех типов слоев:

- Convolution. В случае временных рядов входные данные для CNN обычно представляют собой одномерную последовательность точек данных, представляющих временной ряд. Свёрточные фильтры скользят по последовательности, чтобы идентифицировать шаблоны и извлекать функции.
- Pooling – данный слой уменьшает размерность данных за счет суммирования выходных данных свёрточных слоев.
- Fully connected – этот слой берёт выходные данные свёрточных слоев и слоев объединения и сглаживает их в одномерный вектор. Затем этот одномерный вектор проходит через один или несколько полно связанных слоев, которые можно рассматривать как традиционные слои нейронной сети. Эти слои используются для создания окончательных прогнозов на основе извлеченных функций.

Сети LSTM, в свою очередь, это тип рекуррентной нейронной сети (RNN), которые специально разработаны для преодоления проблемы исчезающего градиента. Проблема исчезающего градиента возникает в RNNs, когда градиенты, используемые для обновления весов сети во время обратного распространения, становятся слишком малыми, чтобы быть эффективными. LSTM решают эту проблему, вводя ячейку памяти, которая хранит информацию с течением времени и позволяет сети выборочно забывать или запоминать информацию с предыдущих временных шагов. Одним из преимуществ использования LSTM для прогнозирования временных рядов является их способность фиксировать долгосрочные зависимости в данных. Также LSTM способна обрабатывать входные последовательности переменной длины.

После выбора архитектуры нейронной сети происходит обучение модели. Существует несколько методов расчета весовых коэффициентов в нейронной сети [7-9]. Вот некоторые из наиболее распространенных:

Градиентный спуск: это наиболее распространенный метод оптимизации, используемый в нейронных сетях. Цель градиентного спуска – найти набор весов, который минимизирует функцию потерь. Алгоритм работает путем вычисления градиента функции потерь по отношению к весам и обновления весов в направлении отрицательного градиента.

Стochastic gradient descent (SGD): это вариант градиентного спуска, при котором веса обновляются после каждого обучающего примера, а не после каждой эпохи. Это может быть быстрее и эффективнее, чем обычный градиентный спуск, особенно для больших наборов данных.

Мини-пакетный градиентный спуск: это еще один вариант градиентного спуска, в котором веса обновляются после обработки небольшой партии обучающих примеров, а не после каждого отдельного примера. Это может обеспечить хороший баланс между скоростью SGD и стабильностью обычного градиентного спуска.

Adagrad: этот алгоритм адаптирует скорость обучения каждого веса на основе его прошлых градиентов. Идея состоит в том, чтобы давать меньшие обновления весов с большими градиентами и большие обновления весов с малыми градиентами.

Momentum: это метод, который помогает ускорить градиентный спуск в соответствующем направлении и гасить колебания. Он делает это, поддерживая экспоненциально взвешенное скользящее среднее градиентов и используя его для обновления весов.

RMSprop: это еще один алгоритм оптимизации, который адаптирует скорость обучения каждого веса на основе его прошлых градиентов. Отличие от Adagrad в том, что он использует скользящее среднее квадратов градиентов для управления скоростью обучения.

Adam: сочетание адаптивной оптимизации скорости обучения и импульса. Он адаптирует скорость обучения для каждого веса на основе исторических градиентов, как в Adagrad и поддерживает

экспоненциально взвешенное скользящее среднее градиентов, как в Momentum. Кроме того, он использует коррекцию смещения, чтобы скорректировать смещение скользящих средних.

Эти методы различаются по сложности, скорости сходимости и возможности обхода локальных минимумов.

## 2. Реализация нейронных сетей

Для построения моделей прогнозирования на основе нейронных сетей были выбраны следующие активы: цена на нефть BRENT и природный газ NYMEX.NG. Перед составлением и обучения моделей был произведен предварительный анализ для каждого временного ряда.

На рисунках 1-5 представлен прогноз временных рядов на основе различных моделей нейронных сетей.

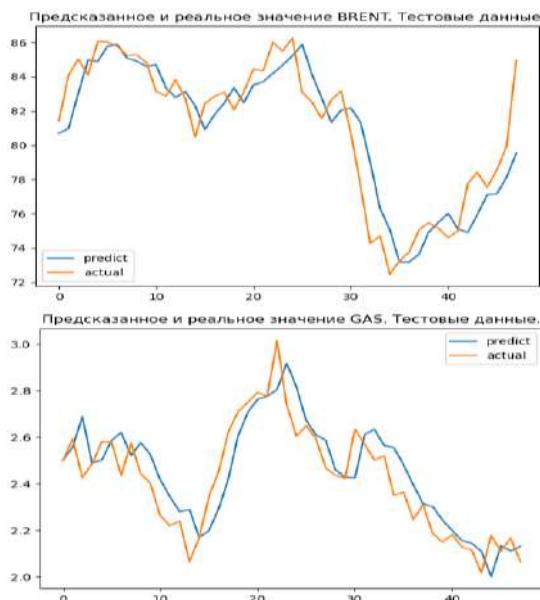


Рис. 1. Модель нейронной сети с полносвязными слоями на тестовых данных

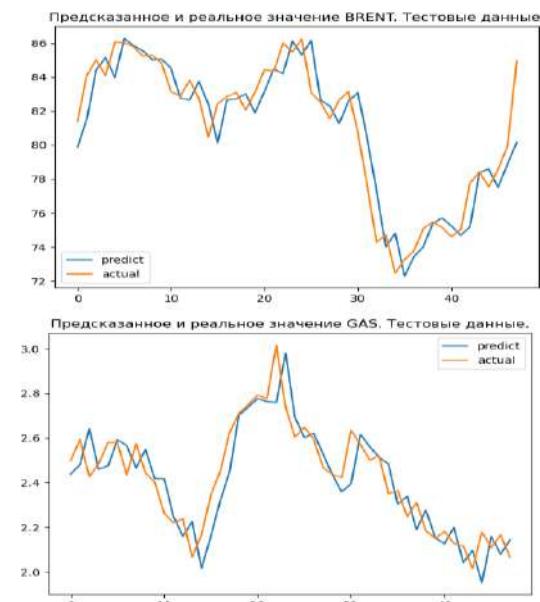
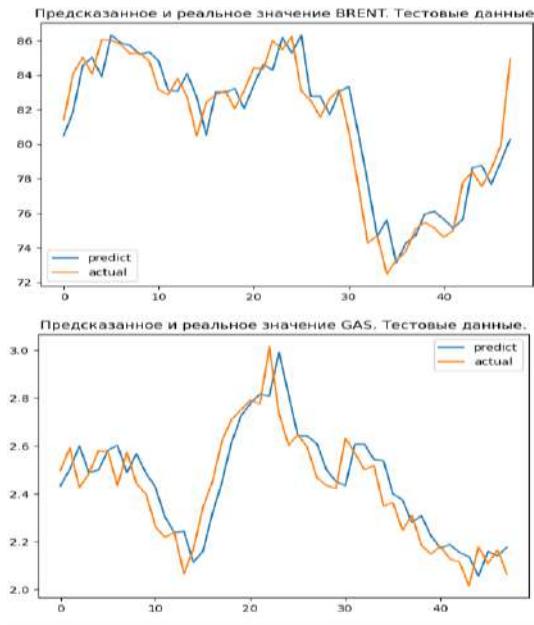
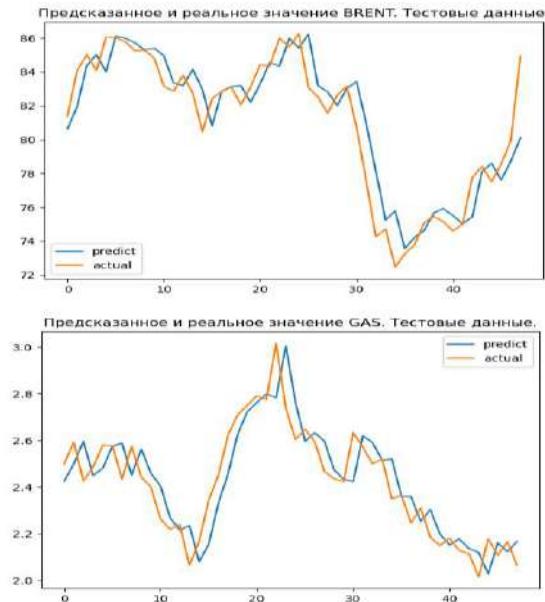


Рис. 2. Модель нейронной сети с ячейками RNN на тестовых данных



*Рис. 3. Модель нейронной сети с ячейками LSTM на тестовых данных*



*Рис. 4. Модель нейронной сети с ячейками GRU на тестовых данных*

### 3. Заключение

В настоящее время нейронные сети являются мощным инструментом для прогнозирования финансовых временных рядов. Они могут использоваться для прогнозирования цен на акции, курсов валют, доходности облигаций и других финансовых инструментов. Однако стоит помнить, что эффективность применения данного инструмента зависит от нескольких факторов, таких как качество данных, выбор архитектуры нейронной сети и методов обучения [10-11]. Одним из основных преимуществ нейронных сетей является их способность к обучению на больших объемах данных и выявлению скрытых зависимостей в этих данных. В результате модели нейронных сетей могут выделять сложные тренды и зависимости в финансовых временных рядах, которые могут быть незаметны при использовании других методов. Однако, использование нейронных сетей также имеет некоторые недостатки. Например, нейронные сети могут быть склонны к переобучению, что может привести к ошибкам в прогнозировании. При этом проблема с затухающим градиентом и построением верной архитектуры также могут приводить к затруднениям в обучении и использовании нейронной сети на практике.

## **Литература**

1. Абдикеев Н.М., Иванюк В.А., Пащенко Ф.Ф. Долгосрочное прогнозирование макроэкономических показателей (на примере ВВП) // Фундаментальные исследования. – 2017. – №. 8–1. – С. 110–114.
2. Архипова А.А. Применение нейронных сетей в задаче прогнозирования финансовых временных рядов // Экономика и бизнес: теория и практика. – 2023. – №6–1 (100). – С. 134–137.
3. Воййко О.А. Анализ временных рядов и прогнозирование: практикум / О. А. Воййко. – Москва; Берлин: Директ-Медиа, 2019. – 175 с.
4. Иванюк В.А., Андропов К.Н., Цвиркун А.Д. Методология совокупного прогнозирования активов и их рисков //Фундаментальные исследования. – 2014. – №. 12–5. – С. 1028–1031.
5. Иванюк В.А., Андропов К.Н., Цвиркун А.Д. Разработка методологии долгосрочного прогнозирования на основе мультитрендового прогноза // Фундаментальные исследования. – 2014. – №. 12–5. – С. 1032–1035.
6. Иванюк В.А., Арутюнов А.Л., Цвиркун А.Д. Разработка инструментальных средств прогнозирования в социально-экономических системах // Препринт доклада. –ИПУ РАН. – 2012. – С. 43.
7. Иванюк В.А., Цвиркун А.Д. Разработка многофакторной системы прогнозирования на основе имитационно-оптимизационного подхода // Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2012. – Т. 10. – №. 1. – С. 104–108.
8. Иванюк В.А. Разработка инновационных систем прогнозирования для решения сложных экономических задач // Известия Волгоградского государственного технического университета. – 2010. – №. 13. – С. 153–157.
9. Fiesler E., Choudry A., John Caulfield H. Weight discretization paradigm for optical neural networks // The International Congress on Optical Science and Engineering. – April 1990. – № 1281. – Р. 164–173.
10. Ivanyuk V., Tsvirkun A. Intelligent system for financial time series prediction and identification of periods of speculative growth on the financial market // IFAC Proceedings Volumes. – 2013. – Т. 46. – №. 9. – С. 1128–1133.
11. Karlsson N. Comparison of linear regression and neural networks for stock price prediction: Dissertation. – 2021. – 31 p. — UPTEC F, ISSN 1401–5757.