

УДК 004.852

5.2.2. Математические, статистические и инструментальные методы экономики (физико-математические науки, экономические науки)

ИССЛЕДОВАНИЕ РЕКУРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧЕ ПРОГНОЗИРОВАНИЯ ФИНАНСОВОГО ВРЕМЕННОГО РЯДА ПРИ ДЕФИЦИТЕ ТРЕНИРОВОЧНЫХ ДАННЫХ

Яхонтова Ирина Михайловна
Кандидат экономических наук, доцент

Нехов Сергей Алексеевич
Студент 2 курса магистратуры
Кубанский Государственный Аграрный университет имени И.Т.Трубилина, Краснодар, Россия

Рекуррентные нейронные сети являются мощным инструментом в области прогнозирования временных рядов благодаря их способности учитывать последовательность данных и сохранять информацию о прошлых состояниях. В данной статье рассматривается применение RNN для прогнозирования финансовых временных рядов в условиях дефицита тренировочных данных. Финансовые рынки характеризуются высокой степенью неопределенности и сложной динамикой, что делает прогнозирование их поведения сложной задачей. Одной из основных проблем при работе с финансовыми данными является недостаток данных для обучения моделей прогнозирования. В данном исследовании мы исследуем эффективность применения рекуррентных нейросетей в условиях ограниченного объема тренировочных данных и исследуем возможные способы улучшения качества прогнозов в таких условиях. Целью данного исследования является оценка точности и надежности прогнозов, полученных с помощью RNN, LSTM, GRU, при ограниченном объеме и качестве тренировочных данных. Для этого мы проводим серию экспериментов на реальных финансовых данных, анализируем полученные результаты и предлагаем рекомендации по улучшению прогнозов в условиях дефицита данных

Ключевые слова: ПРОГНОЗИРОВАНИЕ, НЕЙРОННЫЕ СЕТИ, ИИ, ФИНАНСЫ, ДЕФИЦИТ ДАННЫХ, ПРОГНОЗИРОВАНИЕ ФИНАНСОВОГО ВРЕМЕННОГО РЯДА, ВРЕМЕННЫЕ РЯДЫ, ПРОГНОЗИРОВАНИЕ ВРЕМЕННОГО РЯДА, СКОЛЬЗЯЩЕЕ ОКНО, LSTM, RNN, GRU

<http://dx.doi.org/10.21515/1990-4665-199-022>

UDC 004.852

5.2.2. Mathematical, statistical and instrumental methods of economics (physical and mathematical sciences, economic sciences)

THE STUDY OF RECURRENT NEURAL NETWORKS IN THE TASK OF FORECASTING A FINANCIAL TIME SERIES WITH A SHORTAGE OF TRAINING DATA

Yakhontova Irina Mikhailovna
Candidate of Economic Sciences, Associate Professor

Nehov Sergey Alekseevich
2nd year student of the Master's degree
Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, Russia

Recurrent neural networks are a powerful tool in the field of time series forecasting due to their ability to take into account the sequence of data and store information about past states. This article discusses the use of RNN for forecasting financial time series in conditions of a shortage of training data. Financial markets are characterized by a high degree of uncertainty and complex dynamics, which makes forecasting their behavior a difficult task. One of the main problems when working with financial data is the lack of data for training forecasting models. In this study, we investigate the effectiveness of using recurrent neural networks in conditions of a limited amount of training data and explore possible ways to improve the quality of forecasts in such conditions. The purpose of this study is to evaluate the accuracy and reliability of forecasts obtained using RNN, LSTM, GRU, with limited volume and quality of training data. For this purpose

Keywords: PREDICTION, NEURAL NETWORKS, AI, FINANCE, DATA DEFICIT, FINANCIAL TIME SERIES FORECASTING, TIME SERIES, TIME SERIES PREDICTION, SLIDING WINDOW, LSTM, RNN, GRU

<http://ej.kubagro.ru/2024/05/pdf/22.pdf>

Подбор данных для обучения нейронной сети является самым важным действием, поскольку нейросеть учится на зависимостях данных, которые мы ей предоставляем. Исходя из этого, необходимо сделать первоначальную обработку данных, то есть – препроцессинг (preprocessing).

Поскольку, для корректной оценки нейросети необходимо огромное количество данных, а обучение нейросети на данных из одного временного ряда является некорректным с точки зрения разбиения на тренировочную и тестовые выборки, то, в качестве тестовой выборки будет использоваться часть схожего временного ряда. То есть тестовый временной ряд будет примерно схожей динамики роста и падения, а, так же, в одном секторе финансовых услуг.

Данные с ценами курса акций будут браться из API Yahoo finance, благодаря библиотеке yfinance (для языка программирования Python). Для тренировки нейросети были выбраны цены акций следующих компаний:

- Coca-Cola (тикет: KO);
- PepsiCo (тикет: PEP).

Данные акций были выбраны исключительно из-за их долгой истории цен акций на бирже (с 1980 года) и из-за возможности протестировать устойчивость к «поворотам тренда» и устойчивость к выбросам.

Благодаря используемому API в данных отсутствуют None-ячейки (так же проверено, используя библиотеку pandas), следовательно, целостность не нарушена. Далее необходимо проводить нормализацию, поскольку данные имеют значительные размерности, относительно друг друга, следовательно, модель может оказаться слишком чувствительна к их изменениям в размерности.

$$X = \frac{x_i - \min}{\max - \min},$$

где X – нормализованное значение, $0 \dots 1$;

x_i – численное значение из набора данных;

\min – минимальное значение выборки;

\max – максимальное значение выборки.

Также, стоит отметить, что в задачи прогнозирования временных рядов присутствуют различия в количестве входных данных: одномерные и многомерные. При наличии данных: Data, Open, High, Low, Close, Volume, Adj Close, стоит выделить, что на финальное прогнозирование в большей степени влияет именно данные Open, следовательно, остальные данные будут приводить к нежелательным помехам в обучении нейросети (аномалии/выбросы).

Далее необходимо произвести разбиение данных на подвыборки:

— обучающую (тренировочную);

— валидационную;

— тестовую.

Спецификация разбиения данных временного ряда на данные подвыборки – достаточно сложная задача, поскольку, относительно других данных (картинки, численно исчислимые, текстовые данные), данные слишком малы. И, исходя из данной проблемы, появился способ искусственно «разделить» данные, поделив их методом «скользящее окно» (sliding window).

Суть метода заключается в том, что набор данных разбивается поэтапно на данные «предшествующие» и «результатирующие». То есть, мы берём 3 первоначальных данных, которые будем подавать на вход нейросети, а значение, которое следует за первоначальными данными будет являться «результатирующим» или «предсказываемым». И, впоследствии, данное окно будет двигаться от более старых данных, к более новым с определёнными «оконными рамками». В качестве визуального примера приведу Рисунок 1.

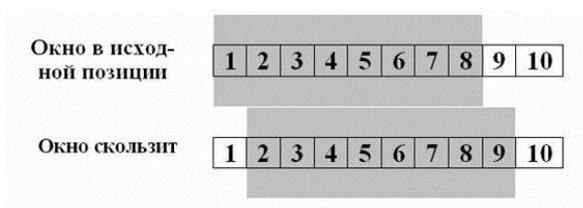


Рисунок 1 – Визуализация работы алгоритма скользящего окна.

Выбор функции потерь

Функция потерь – это функция, которая измеряет расхождение между предсказанными значениями модели и фактическими значениями в процессе обучения модели. Основная цель функции потерь заключается в том, чтобы минимизировать ошибку и улучшить качество модели.

Для выбора функции потерь и оптимизатора необходимо сравнить их между собой и проанализировать их виды, преимущества и недостатки. Видов функций потерь есть множество, но все они подразделяются для задач классификации и регрессии.

Поскольку данная задача относится к задачам регрессии, то на рассмотрение попадают следующие функции потерь:

- 1) Среднеквадратическая ошибка (MSE);
- 2) средняя абсолютная ошибка (MAE).

И, уравнения которых, соответственно равны:

$$MSE = \frac{1}{n} * \sum_i^n (y - y_{pred})^2,$$

где, MSE – результат работы функции, среднеквадратическая ошибка;

n – количество значений в выборке;

y – фактическое (верное) значение;

y_{pred} – предсказанное значение.

$$MAE = \frac{1}{n} * \sum_i^n |y - y_{pred}|,$$

где, MAE – результат работы функции, средняя абсолютная ошибка;

n – количество значений в выборке;

y – фактическое (верное) значение;

y_{pred} – предсказанное значение.

Положительные и отрицательные качества данных функций приведены в Таблице 1.

Проведя сравнение, можно прийти к выводу, что, и MSE, и MAE имеют как значимые преимущества, так и недостатки. Но, поскольку основные оптимизаторы нейросетей, которые предназначены для задачи регрессии (прогнозирование временного ряда), имеют в себе дифференцирование, то функция потерь MSE, в данном случае, будет лучшим выбором, по сравнению с MAE.

Таблица 1 – Качества рассматриваемых функций потерь

Наименование функции потерь	Преимущества	Недостатки
MSE	1) Является распространённой метрикой для задач прогнозирования; 2) учёт сильных отклонений (выбросов).	1) Чувствительность к выбросам; 2) возможное завышение показателей точности; 3) сложная интерпретация получившихся значений.
MAE	1) Устойчивость к выбросам; 2) лёгкая интерпретируемость (прямолинейная).	1) Не учитывает относительную важность ошибок; 2) не является дифференцируемым.

Учитывая выбранную функцию потерь, необходимо выбрать функцию-оптимизатор, которая будет удовлетворять поставленной задаче. Но, поскольку видов оптимизаторов существует огромное количество, и, при большом желании, в совокупности с пониманием, можно дополнить существующий алгоритм или составить собственный, то их называют «эвристическим» параметром. И, алгоритмы выбираются по следующим критериям:

- 1) Точность за итерацию;
- 2) скорость оптимизации за итерацию (скорость).

Устоявшиеся функции-оптимизаторы для задач регрессии (в данной ситуации – прогнозирования временного ряда) являются следующие:

- 1) SGD (возможны использования импульсов);
- 2) RMSProp;
- 3) Adam (возможны использования импульсов).

Преимущества и недостатки выбранных оптимизаторов предоставлены в Таблице 2.

Таблица 2 – Преимущества и недостатки функций-оптимизаторов

Наименование функции-оптимизатора	Преимущества	Недостатки
SGD	<ol style="list-style-type: none"> 1) Простота реализации; 2) Эффективность на больших наборах данных; 3) Устойчив к шумам. 	<ol style="list-style-type: none"> 1) Низкая скорость; 2) Вероятность «увязнуть» в локальных минимумах.
RMSProp	<ol style="list-style-type: none"> 1) Адаптация к изменяющейся структуре 	<ol style="list-style-type: none"> 1) Высокая вероятность «застрять» в

Наименование функции-оптимизатора	Преимущества	Недостатки
	функции потерь; 2) Высокая начальная скорость сходимости.	локальных минимумах; 2) Высокое требование к подбору гиперпараметров.
Adam	1) Быстрая сходимость; 2) Автоматическая адаптация скорости обучения.	1) Задействование сравнительного большого объёма памяти; 2) Высокая чувствительность к гиперпараметрам.

Изучив принцип работы каждого оптимизатора по-отдельности, становится очевидно, что, для решения данной задачи, необходимо использовать Adam, поскольку он имеет преимущества алгоритма RMSProp, а, учитывая небольшое количество данных, то затраты по памяти будут незначительные.

Архитектуры нейросетей, рассматриваемые в исследовании

Поскольку обучение нейросети – это способ формирования аппроксимирующей функции на основе данных (в нашем случае заранее известных или размеченных), то данную нейросеть можно воспринимать как частный случай обучения с учителем.

В данной научно-исследовательской работе будет рассматриваться исключительно архитектура на рекуррентных блоках, следовательно, она будет являться многослойными RNN, GRU, LSTM нейросетями.

С учётом выбранных нейросетевых слоёв и выходной функции активации, следует выбрать количество скрытых слоёв (в многослойных сетях) и количество ядер (нейронов) в каждом слое. Учитывая поставленные выше критерии, построю следующие нейронные сети:

- 1) OnLrRNN(1, 128);
- 2) MnLrRNN(4, 128|64|32|16);
- 3) BiRNN(4, 128|64|32|16);
- 4) OnLrGRU(1, 128);
- 5) MnLrGRU(4, 128|64|32|16);
- 6) BiGRU(4, 128|64|32|16);
- 7) OnLrLSTM(1, 128);
- 8) MnLrLSTM(4, 128|64|32|16);
- 9) BiLSTM(4, 128|64|32|16);

Где значения в скобках соответствуют числу слоёв (без учёта выходного), количество ядер в слоях, соответственно, через символ «|».

Сокращения, используемые выше означают:

- 1) Сети с одним слоем (OneLayer – OnLr);
- 2) сети с многими слоями (ManyLayer – MnLr);
- 3) двунаправленные (Bidirectional – Bi).

Результаты экспериментов

Проведя обучение на выбранных, нормированных, данных для различных рекуррентных архитектур, были выделены критерии, по которым проводился эксперимент:

- 1) Loss-параметр;
- 2) Ассурасу-параметр;
- 3) время обучения.

Вышеперечисленные критерии были визуализированы в виде графиков, которые были разбиты по принципу видов нейросетевого слоя, которые предоставлены на Рисунке 2 – Рисунок 4.

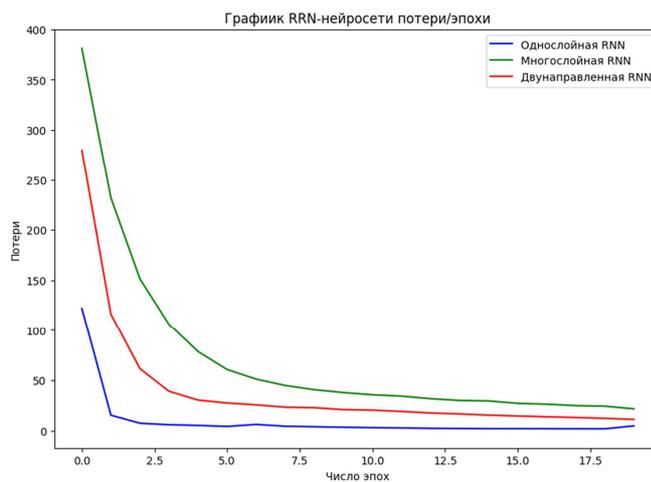


Рисунок 2 – Визуализация изменения потерь при обучении RNN-нейросетей

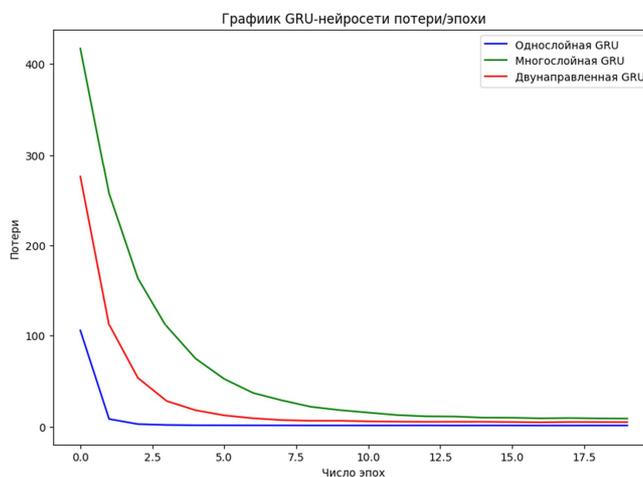


Рисунок 3 – Визуализация изменения потерь при обучении GRU-нейросетей

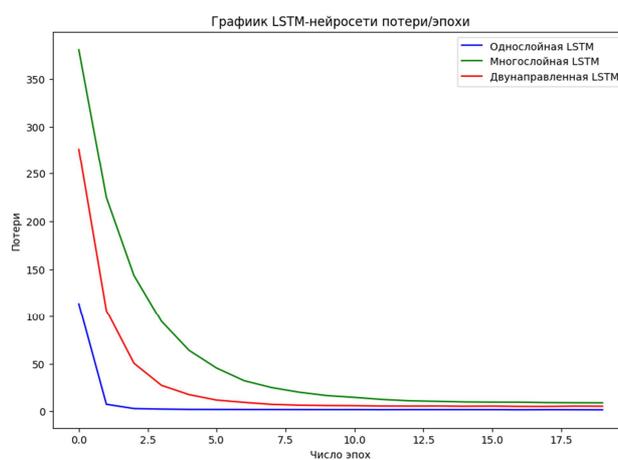


Рисунок 4 – Визуализация изменения потерь при обучении LSTM-нейросетей

Исходя из вышепредставленных графиков, можно сделать вывод, касательно видов нейросетевого слоя, что, в условиях малого количества данных для обучения и прогнозирования, во всех случаях обучение многослойной, двунаправленной и однослойной нейронной сети, соответственно является коррелирующими друг относительно друга. То есть многослойная модель начинает с большого значения потерь из-за большого количества слоёв (относительно однослойной) и однонаправленными весами зависимостей, двунаправленная имеет среднее значение потерь благодаря способности в начале обучения сообщаться двум соседним слоям в обоих направлениях, и, однослойная модель является наиболее привлекательной, в плане потерь, но это максимум обработки данной нейронной сети из-за малого количества.

Для наглядности представления графиков всех нейросетей, можно обратиться к Рисунку 5.

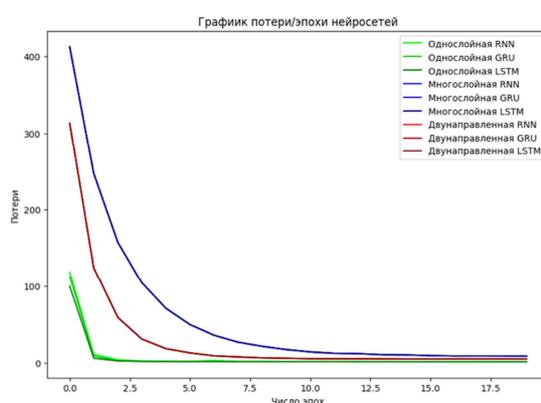


Рисунок 5 – Визуализация изменения потерь при обучении всех нейросетей

На Рисунке 5 можно заметить, что значительные (относительно многослойных и двунаправленных моделей) расхождение на начальных стадиях обучения присутствуют у однослойных моделей. Это связано с внутренней структурой (архитектурой) самих моделей, и, высокая чувствительность к выбросам временной последовательности данных, в то время как другие оперируют долгосрочными зависимостями.

Для простоты восприятия данных, форматирую их в таблицу 3.

Таблица 3 – Результаты критериев для нейросетей

Наименование нейросети	Loss	Ассурасу, %	Время обучения, мин.
OnLrRNN	1174.143799	0.030786	3
MnLrRNN	1330.898315	0.000000	16
BiRNN	1213.132690	0.000000	28
OnLrGRU	1339.341919	0.030786	1
MnLrGRU	1095.978027	0.030786	1
BiGRU	1024.598389	0.030786	3
OnLrLSTM	1212.563110	0.030786	1
MnLrLSTM	1080.634399	0.030786	2
BiLSTM	1021.408203	0.030786	3

Исходя из данных, предоставленных в таблицу 3, можно сделать выводы, что:

1) Время обучения в моделях увеличивается с увеличением количества скрытых слоёв и наличием более, чем одного, направления изменения весов;

2) скорость обучения у GRU-моделей выше, по сравнению с LSTM, при схожих условиях;

3) наличие двунаправленного скрытого слоя уменьшает потери, по сравнению с остальными моделями;

4) точность предсказаний является плохой у RNN-моделей.

Для более полной оценки точности необходимо предоставить визуализацию графика цен акций двух компаний на Рисунке 6.

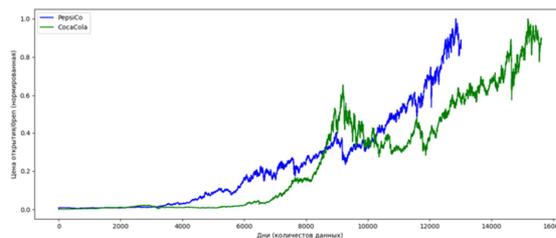


Рисунок 6 – Визуализация данных открытия (Open) цен на биржевых торгах

Из графика, на рисунке 6 можно заметить, что в данных присутствуют шумы, из-за которых точность предсказаний может быть сильно неверной, даже при наличии предпроцессинга – нормализации. Так же, на графиках можно увидеть часть данных, которые находятся в «антитренде», то есть, если данные одного временного ряда «движутся» в одну сторону, то данные предсказываемого – в обратную. Это довольно важное замечание, поскольку в условиях малого количества данных, эти «окна антитренда» при прогнозировании могут расцениваться моделью как чрезмерные выбросы данных, и будет стараться их не учитывать, а продолжать прогнозировать данные по тренировочным, но не являются переобучением, поскольку не повторяют в полной мере данные тренировочного паттерна. Это можно увидеть на Рисунках 7 – 9.

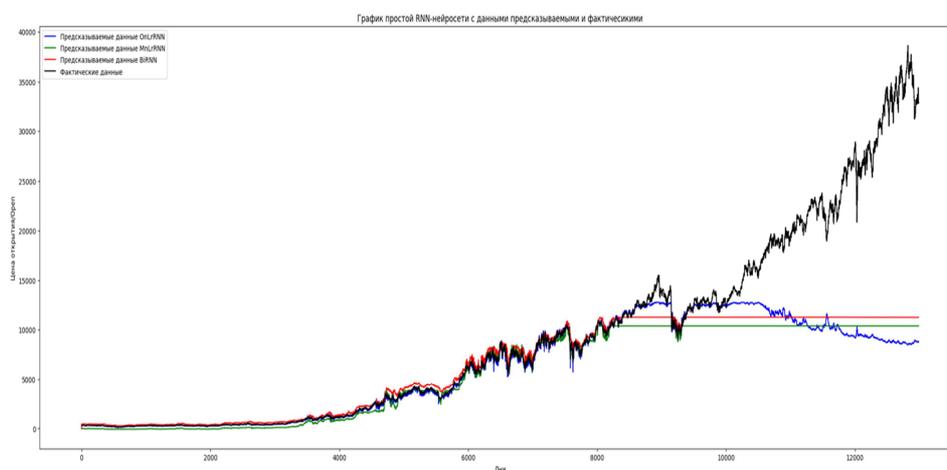


Рисунок 7 – Визуализация работы по предсказанию RNN-нейросетей

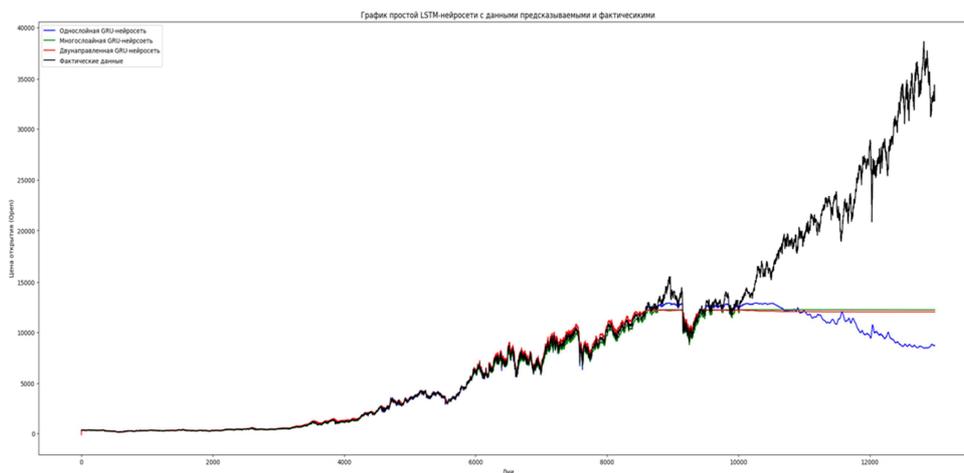


Рисунок 8 – Визуализация работы по предсказанию GRU-нейросетей

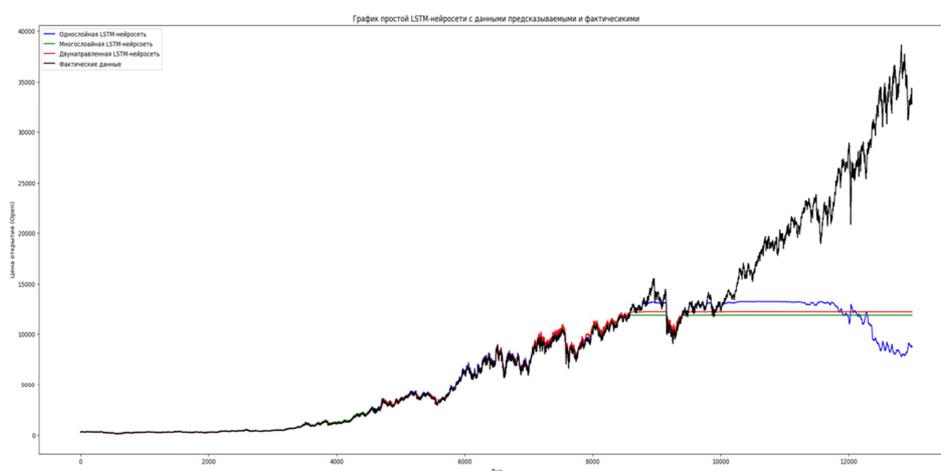


Рисунок 9 – Визуализация работы по предсказанию LSTM-нейросетей

Вывод

Принимая во внимание всё вышеперечисленные результаты, с уверенностью можно сказать, что обучение нейросети для задачи прогнозирования временного ряда в условиях малого количества тренировочных данных, наилучшим выбором будет использовать:

- нормированные временные ряды;
- временные ряды без значительных выбросов, либо антитрендовых последовательностей (использовать схожие временные ряды);
- использовать LSTM-слои для возможности поиска длительных зависимостей временного ряда;

— использовать разновидность LSTM моделей как двунаправленную.

По итогам проделанной работы были получены численные значения потерь, точности и времени тренировки нейросети в зависимости от условия изначального недостатка тренировочных данных (одного временного ряда). Была выделена лучшая нейросетевая модель, были объяснены причины результатов тестового прогнозирования временного ряда.

Литература

1) Афанасьев В. Н. Анализ временных рядов и прогнозирование : учебник / В. Н. Афанасьев; Ай Пи Ар Медиа – Саратов, Оренбургский гос. ун-т. – Оренбург : 2020. – 286 с.

2) Голубев Ю. В. Рекуррентные нейронные сети для анализа временных рядов // Научно-технический вестник информационных технологий, механики и оптики. 2016. Т. 16. № 4. С. 691-696.

3) Занина Д. Н., Занин В. Т. Методы рекуррентных нейронных сетей в задачах прогнозирования временных рядов // Вестник пермского национального исследовательского политехнического университета. 2019. Т. 15. № 4. С. 678-693.

4) Исследование проблемы оценки эффективности инвестиционных проектов в условиях цифровой трансформации бизнеса / Л. О. Великанова, И. М. Яхонтова, А. В. Коваленко, А. С. Маликов // Современная экономика: проблемы и решения. – 2022. – № 1(145). – С. 38-48.

5) Ковалева Я. В. Использование VBA при решении финансовых задач / Я. В. Ковалева, И. М. Яхонтова // Информационное общество: современное состояние и перспективы развития : сборник материалов VIII международного форума, Краснодар, 26–30 декабря 2016 года. – Краснодар: Кубанский государственный аграрный университет, 2017. – С. 328-329.

6) Кормен, Томас Х., Лейзерсон, Чарльз И., Ривест, Рональд Л., Штайн, Клиффорд. В24 Алгоритмы: построение и анализ, 2-е издание. : Пер. с англ. — М. : Издательский дом “Вильямс”, 2011. – 1296 с.

7) Куковинцева В. Б., Куковинцев А. В. Применение рекуррентных нейронных сетей для анализа финансовых рынков // Журнал Европейской экономической ассоциации. 2016. № 2(31). С. 128-133.

8) Минаев В. А., Сенько Е. С., Богданов А. В. Прогноз временных рядов с использованием рекуррентных нейронных сетей // Компьютерные инструменты в образовании. 2019. Т. 2. № 2(30). С. 82-88.

References

1) Afanas'ev V. N. Analiz vremennyh rjadov i prognozirovanie : uchebnik / V. N. Afanas'ev; Aĭ Pi Ar Media – Saratov, Orenburgskii gos. un-t. – Orenburg : 2020. – 286 s.

2) Golubev Ju. V. Rekurrentnye nejronnye seti dlja analiza vremennyh rjadov // Nauchno-tehnicheskij vestnik informacionnyh tehnologij, mehaniki i optiki. 2016. T. 16. № 4. S. 691-696.

3) Zanina D. N., Zanin V. T. Metody rekurrentnyh nejronnyh setej v zadachah prognozirovanija vremennyh rjadov // Vestnik permskogo nacional'nogo issledovatel'skogo politehnicheskogo universiteta. 2019. T. 15. № 4. S. 678-693.

4) Issledovanie problemy ocenki jeffektivnosti investicionnyh proektov v uslovijah cifrovoj transformacii biznesa / L. O. Velikanova, I. M. Jahontova, A. V. Kovalenko, A. S. Malikov // Sovremennaja jekonomika: problemy i reshenija. – 2022. – № 1(145). – S. 38-48.

5) Kovaleva Ja. V. Ispol'zovanie VBA pri reshenii finansovyh zadach / Ja. V. Kovaleva, I. M. Jahontova // Informacionnoe obshhestvo: sovremennoe sostojanie i perspektivy razvitija : sbornik materialov VIII mezhdunarodnogo foruma, Krasnodar, 26–30 dekabnja 2016 goda. – Krasnodar: Kubanskij gosudarstvennyj agrarnyj universitet, 2017. – S. 328-329.

6) Kormen, Tomas H., Lejzerson, Charl'z I., Rivest, Ronal'd L., Shtajn, Klifford. V24 Algoritmy: postroenie i analiz, 2-e izdanie. : Per. s angl. — M. : Izdatel'skij dom “Vil'jams”, 2011. – 1296 s.

7) Kukovinceva V. B., Kukovincev A. V. Primenenie rekurrentnyh nejronnyh setej dlja analiza finansovyh rynkov // Zhurnal Evropejskoj jekonomicheskoj asociacii. 2016. № 2(31). S. 128-133.

8) Minaev V. A., Sen'ko E. S., Bogdanov A. V. Prognoz vremennyh rjadov s ispol'zovaniem rekurrentnyh nejronnyh setej // Komp'juternye instrumenty v obrazovanii. 2019. T. 2. № 2(30). S. 82-88.