

МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра МО ЭВМ

ОТЧЕТ
по научно-исследовательской работе
Тема: Исследование параметров рекуррентной нейронной
сети с управляемыми элементами

Студент гр. 8306

Бобохонов А.Х.

Преподаватель

Жукова Н.А.

Санкт-Петербург

2023

ЗАДАНИЕ НА НАУЧНО-ИССЛЕДОВАТЕЛЬСКУЮ РАБОТУ

Студент: Бобохонов А.Х.

Группа: 8306

Тема НИР: Исследование параметров рекуррентной нейронной сети с управляемыми элементами.

Задание на НИР: Изучить различные параметры рекуррентной нейронной сети (РНН) и определить, как они влияют на производительность модели.

Сроки выполнения НИР: 01.09.2023 – 20.12.2023

Дата сдачи отчета: 15.12.2023

Дата защиты отчета: 22.12.2023

Студент

Бобохонов А.Х.

Руководитель

Жукова Н.А.

АННОТАЦИЯ

Исследование параметров рекуррентных нейронных сетей (РНН) с управляемыми элементами представляет собой актуальную область исследований в области искусственного интеллекта и машинного обучения. В данной работе анализируются различные параметры РНН, такие как архитектура сети, глубина сети, размерность скрытых состояний, тип функций активации и методы оптимизации. Основной целью исследования является определение влияния этих параметров на производительность РНН в различных задачах машинного обучения.

Результаты исследования позволяют лучше понять, какие параметры RNN наиболее важны для достижения высокой точности и эффективности обучения. Также для достижения поставленной цели на практике была разработана модель РНН и скорректированы ее параметры.

Исследование параметров РНН с управляемыми элементами поможет дальше совершенствовать техники глубокого обучения и повышать качество решений в области искусственного интеллекта.

SUMMARY

The study of parameters of recurrent neural networks (RNN) with controlled elements is an urgent area of research in the field of artificial intelligence and machine learning. In this paper, various RNN parameters are analyzed, such as network architecture, network depth, dimension of hidden states, type of activation functions and optimization methods. The main purpose of the study is to determine the effect of these parameters on the RNN performance in various machine learning tasks.

The results of the study allow us to better understand which RNN parameters are most important for achieving high accuracy and efficiency of training. Also, in order to achieve this goal in practice, an RNN model was developed and its parameters were adjusted.

The study of RNN parameters with controlled elements will help to further improve deep learning techniques and improve the quality of solutions in the field of artificial intelligence.

СОДЕРЖАНИЕ

	Введение	6
1.	Постановка задачи	7
2.	Результаты работы в осеннем семестре	8
3.	Описание предполагаемого метода решения	
4.	План работы на весенний семестр	14
	Заключение	15
	Список использованных источников	16

ВВЕДЕНИЕ

В современном мире нейронные сети играют ключевую роль в различных областях, от обработки естественного языка до обучения машинного зрения. Однако, несмотря на значительные успехи, вопросы оптимизации производительности и эффективности остаются актуальными. В этом контексте, данное исследование направлено на анализ и оптимизацию параметров рекуррентных нейронных сетей (РНС) с использованием управляемых элементов.

Целью данного исследования является более глубокое понимание влияния различных параметров РНС, таких как архитектура, скрытые слои, функции активации, на их производительность. Особое внимание уделяется внедрению управляемых элементов в структуру РНС и их роль в оптимизации обучения и адаптации моделей к разнообразным задачам.

В рамках этого исследования была разработана программа на языке Python, использующая библиотеки `pandas`, `numpy` и `tensorflow`, для прогнозирования цен на золото с использованием рекуррентных нейронных сетей “Долгая краткосрочная память” (LSTM). Программа включает в себя загрузку данных о цене золота, их нормализацию, создание последовательностей для прогнозирования временных рядов, построение модели LSTM, обучение и визуализацию результатов.

Этот подход не только предоставляет абстрактное понимание исследования, но и демонстрирует практическую реализацию с использованием передовых методов машинного обучения. Проведенный анализ и прогнозирование цен на золото призваны служить вкладом в развитие области нейронных сетей и их применения в финансовой аналитике.

1. ПОСТАНОВКА ЗАДАЧИ

Актуальность: Исследование параметров рекуррентных нейронных сетей с управляемыми элементами актуально в контексте поиска оптимальных решений для повышения эффективности и точности нейронных сетей. Интеграция управляемых элементов предоставляет новые возможности для оптимизации обучения и адаптации моделей к различным задачам, что имеет значительное значение в развивающейся области искусственного интеллекта.

Объектом исследования являются прогнозирование цен на золото с использованием рекуррентных нейронных сетей.

Предметом исследования являются параметры и характеристика рекуррентных нейронных сетей.

Цель: Исследовать влияние параметров рекуррентных нейронных сетей, таких как архитектура, количество слоев, нейронов в слое, функции активации, а также использование управляемых элементов, на процесс прогнозирования цен на золото. Оптимизировать эти параметры с целью улучшения точности и эффективности модели в предсказании временных рядов финансовых данных.

Задачи: Исследовать влияние различных параметров рекуррентных нейронных сетей с управляемыми элементами на точность прогнозирования временных рядов.

2. РЕЗУЛЬТАТЫ РАБОТЫ В ОСЕННЕМ СЕМЕСТРЕ

Поскольку тема диплома была изменена, ранее составленный план стал неактуальным. Чтобы структурировать работу в этом семестре составлен новый план:

- Исследование сетей RNN;
- Изучить различные параметры рекуррентной нейронной сети (РНН) и определить, как они влияют на производительность модели;
- Разработка модели RNN;
- Изучение их влияния на модель путем изменения параметров РНН.

В результате работы осеннего семестра была создана программная модель рекуррентной нейронной сети с использованием языка программирования Python 3, формат блокнота “ipynb”.

Рекуррентные нейронные сети (RNN) - это тип нейронных сетей, спроектированный для работы с последовательными данными или данными переменной длины. Они обладают способностью запоминать предыдущие состояния, что делает их эффективными для обработки последовательностей, таких как тексты, временные ряды и речь. Рекуррентные нейроны в RNN позволяют передавать информацию от предыдущих шагов обработки к следующим, что позволяет учесть контекст и зависимости в данных.

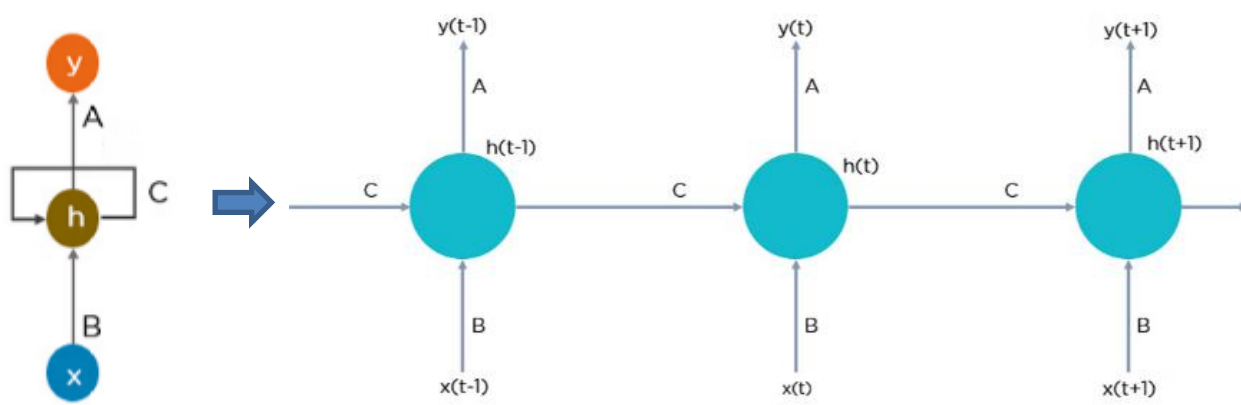
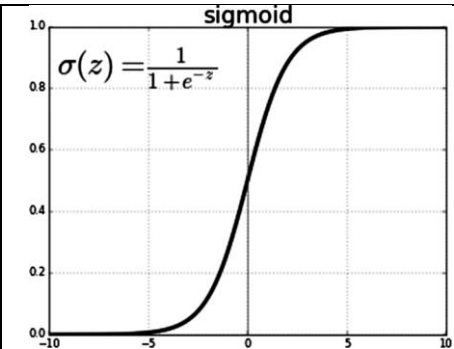
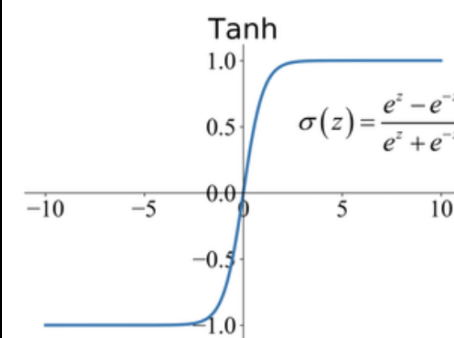
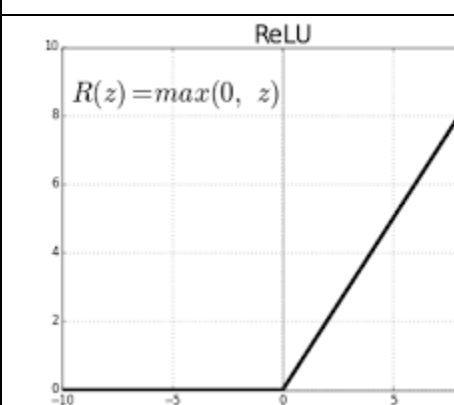


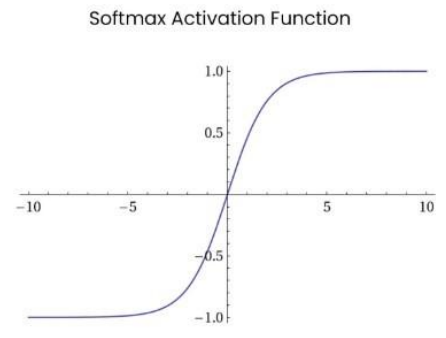
Рисунок 1. RNN и ее развернутое представление

$$h(t) = f_c(h(t - 1), x(t))$$

Здесь “x” - это входной слой, “h” - скрытый слой, а “y” - выходной слой. А, В и С - параметры сети, используемые для улучшения выходных данных модели. В любой момент времени t текущий входной сигнал представляет собой комбинацию входных данных в x(t) и x(t1). Выходные данные в любой момент времени передаются обратно в сеть для улучшения выходных данных [1].

Рекуррентные нейронные сети используют функции активации так же, как и другие нейронные сети, для внесения нелинейности в свои модели. Вот некоторые распространенные функции активации, используемые в RNNs:

Sigmoid Function:	$\sigma(x) = 1 / (1 + e^{(-x)})$	 <p>sigmoid</p> $\sigma(z) = \frac{1}{1 + e^{-z}}$
Hyperbolic Tangent (Tanh) Function:	$\tanh(x) = (e^x - e^{-x}) / (e^x + e^{-x})$	 <p>Tanh</p> $\sigma(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$
Rectified Linear Unit (Relu) Function:	$\text{ReLU}(x) = \max(0, x)$	 <p>ReLU</p> $R(z) = \max(0, z)$

Softmax Function:	$\text{softmax}(x) = \frac{e^x}{\sum(e^x)}$	 <p>Softmax Activation Function</p>
-------------------	---	--

Архитектуры RNN:

- *Простые РНН (Simple RNN)*: Основная форма рекуррентных нейронных сетей, где каждый нейрон передает информацию следующему шагу в последовательности.

- *LSTM (Long Short-Term Memory)*: Расширение РНН, спроектированное для борьбы с проблемой затухающего градиента и способное более эффективно управлять долгосрочными зависимостями в данных.

- *GRU (Gated Recurrent Unit)*: Другая модификация РНН, предложенная для решения проблемы затухающего градиента. GRU обеспечивает некоторые преимущества перед LSTM в смысле вычислительной эффективности.

- *Bidirectional RNN*: Этот тип сетей обрабатывает последовательности в обоих направлениях (вперед и назад), что позволяет учесть как прошлое, так и будущее контексты.

- *Сети с ячейками внимания (Attention-based Networks)*: Используют механизм внимания для фокусировки на определенных частях входных данных, улучшая обработку длинных последовательностей.

Архитектура модели:

- Первый слой LSTM с 100 нейронами, функцией активации ReLU и возвратом последовательностей.

- Слой Dropout с вероятностью отсева 20% для предотвращения переобучения.

- Второй слой LSTM с 100 нейронами и функцией активации ReLU.

- Еще один слой Dropout с вероятностью отсева 20%.

- Полносвязанный слой с 50 нейронами и функцией активации ReLU.
- Выходной слой без функции активации.

Оптимизатор:

- Используется оптимизатор Adam с параметрами `learning_rate=0.001`, `beta_1=0.9` и `beta_2=0.999`.

Функция потерь:

- Модель компилируется с функцией потерь `'mean_squared_error'`, так как решается задача регрессии (прогнозирование числовых значений).

Model: "sequential_21"

Layer (type)	Output Shape	Param #
lstm_40 (LSTM)	(None, 20, 100)	40800
dropout_42 (Dropout)	(None, 20, 100)	0
lstm_41 (LSTM)	(None, 100)	80400
dropout_43 (Dropout)	(None, 100)	0
dense_42 (Dense)	(None, 50)	5050
dense_43 (Dense)	(None, 1)	51

=====
 Total params: 126301 (493.36 KB)
 Trainable params: 126301 (493.36 KB)
 Non-trainable params: 0 (0.00 Byte)

Рисунок 2. Описание модели

Ранняя остановка:

- В коде применяется ранняя остановка (EarlyStopping) с мониторингом потерь на валидационном наборе, терпением 20 эпох и восстановлением лучших весов.

Обучение модели:

- Модель обучается на тренировочных данных в течение 50 эпох с размером пакета 32.

Прогнозы:

- После обучения модель используется для прогнозирования как на существующих данных, так и в будущем.

Визуализация:

- Используются графики для отображения функции потерь в процессе обучения и сравнения фактических цен, прогнозов на существующих данных и прогнозов в будущем.

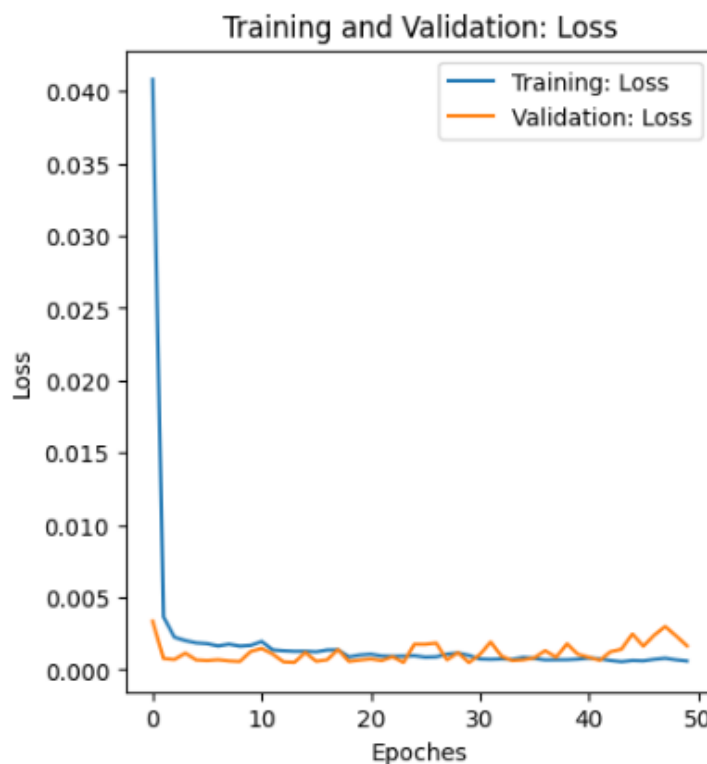


Рисунок 3. Функция потерь

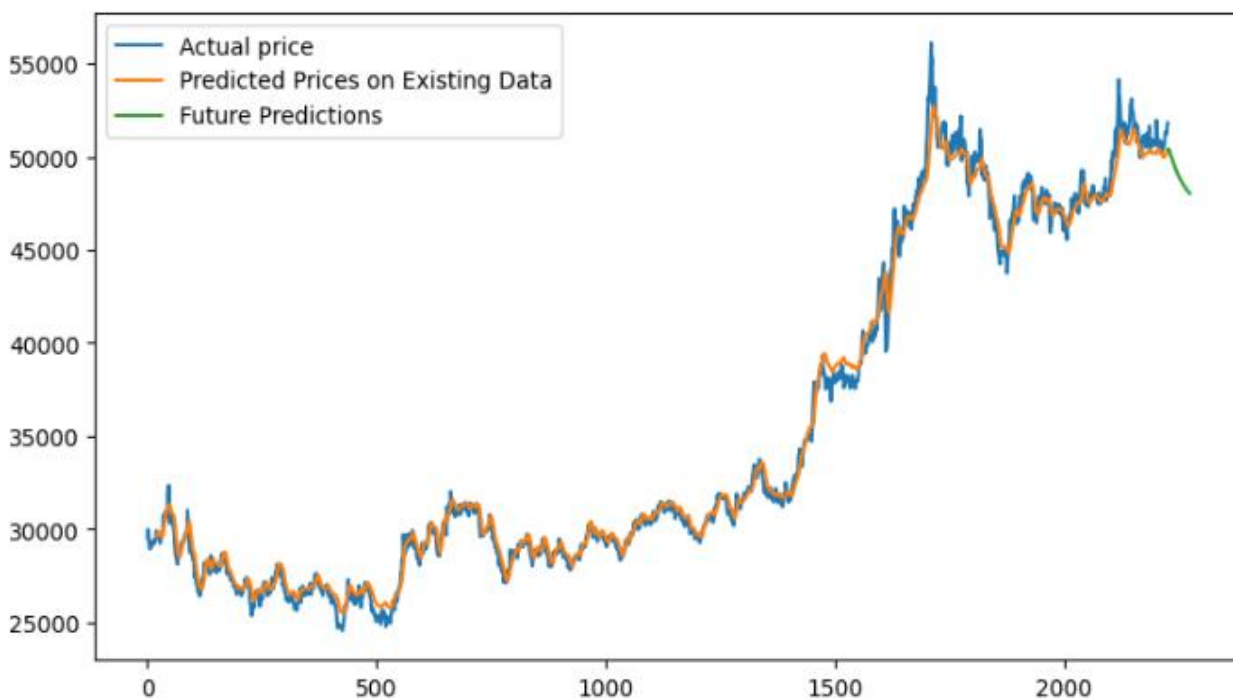


Рисунок 4. Вывод графика фактических цен, прогнозов и прогнозов в будущем

3. ОПИСАНИЕ ПРЕДПОЛАГАЕМОГО МЕТОДА РЕШЕНИЯ

Модель прогнозирования цены на золото была создана с использованием сети Долгая краткосрочная память (LSTM) для исследования параметров RNN и исследование включать в себя следующие аспекты:

1. Анализ данных:

- Изучение структуры и характеристик загруженных данных о цене золота.
- Визуализация временных рядов, выделение трендов и сезонных паттернов.

2. Нормализация данных:

- Разбор процесса нормализации цен и обоснование выбора диапазона (от 0 до 1).

3. Подготовка данных для модели:

- Изучение создания последовательностей для обучения модели.
- Разделение данных на тренировочный и тестовый наборы.

4. Архитектура модели:

- Обоснование выбора архитектуры LSTM с различными слоями и функциями активации.
- Рассмотрение применения слоев Dropout для предотвращения переобучения.

5. Обучение модели:

- Анализ процесса обучения, включая выбор оптимизатора и параметров обучения.
- Рассмотрение использования ранней остановки для предотвращения переобучения.

6. Оценка производительности модели:

- Интерпретация графиков функции потерь на тренировочных и валидационных данных.
- Визуальное сравнение фактических цен, прогнозов на существующих данных и прогнозов в будущем.

7. Прогнозирование в будущем:

- Анализ подхода к прогнозированию цен на будущее на основе текущих данных.
- Визуальное представление фактических цен, прогнозов на существующих

данных и прогнозов в будущем.

8. Оптимизация и дальнейшие шаги:

- Рассмотрение возможных улучшений модели или изменений в параметрах для оптимизации производительности.
- Обсуждение дополнительных аспектов, таких как использование других типов нейронных сетей или улучшенных методов прогнозирования.

Исследование поможет лучше понять процесс построения и обучения временных рядов с использованием рекуррентных нейронных сетей и оптимизировать модель для более точных прогнозов.

3. ПЛАН РАБОТЫ НА ВЕСЕННИЙ СЕМЕСТР

По результатам выполнения заданий осеннего семестра составлен план работы на весенний семестр следующим образом.

- Разработка алгоритма настройки параметров РНН с управляемыми элементами.

- Провести анализ влияния различных параметров модели на качество прогнозирования. Это может включать в себя чувствительность к длине последовательности, числу нейронов и другим параметрам.

- Расширить область применения модели, протестировав её на других временных рядах или изменяя исходные данные для оценки устойчивости и обобщающей способности.

- Создать подробную документацию, описывающую шаги экспериментов, параметры модели, обработку данных и результаты. Это облегчит понимание работы и возможные последующие исследования.

- Подготовить презентацию, в которой кратко представлены результаты исследования, методология работы, выводы и рекомендации для дальнейших шагов.

- Провести обсуждение с научным руководителем, обменяться мнениями, получить обратную связь и рекомендации по улучшению исследования.

- Подготовить и отправить результаты исследования на локальной конференции или в журналы для публикации.

ЗАКЛЮЧЕНИЕ

В результате проведенного исследования была разработана модель для прогнозирования цен на золото с использованием рекуррентной нейронной сети (LSTM). В ходе исследования были изучены параметры модели РНН. Значения параметров РНН были изменены, и их влияние на производительность модели было исследовано. Модель успешно прогнозировала цены на существующих данных, а также демонстрировала способность делать прогнозы на несколько периодов в будущем. Графики функции потерь в процессе обучения, фактических цен, прогнозов на существующих данных и прогнозов в будущем были построены для наглядной оценки результатов. Визуальные результаты позволяют оценить точность и пригодность модели для прогнозирования временных рядов цен на золото.

Исследование подчеркивает эффективность применения рекуррентных нейронных сетей, особенно LSTM, в анализе временных рядов. Помимо этого, важным является постоянное внимание к мониторингу функции потерь и использование методов, таких как ранняя остановка, для управления процессом обучения и предотвращения переобучения модели.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Power of Recurrent Neural Networks (RNN): Revolutionizing AI –
<https://www.simplilearn.com/tutorials/deep-learning-tutorial/rnn>
2. LSTM – сети долгой краткосрочной памяти -
<https://habr.com/ru/companies/wunderfund/articles/331310/>
3. Рекуррентные нейронные сети - Кустикова Валентина, к.т.н., ст.преп. каф.
МОСТ ИИТММ, ННГУ им. Н.И. Лобачевского
4. Применение рекуррентных нейронных сетей в замкнутых следящих
системах с нелинейными объектами - С.В. Феофилов, Д.Л. Хапкин
5. Разработка рекуррентных нейронных сетей в задачах идентификации
систем управления следящими электроприводами крупного радиотелескопа - М. П.
Белов, Ч. Х. Фыонг Санкт-Петербургский государственный электротехнический
университет «ЛЭТИ» им. В. И. Ульянова (Ленина)
6. Прогнозирование стоимости котировок при помощи LSTM и GRU сетей -
Р.С. Ехлаков, В.А. Судаков
7. Репозиторий разрабатываемого приложения:
https://github.com/Akhmadkhan/Research_Work-3/tree/main