

Министерство образования и науки Российской Федерации
Московский физико-технический институт (государственный
университет)

Физтех-школа прикладной математики и информатики
Кафедра дискретной математики

Выпускная квалификационная работа магистра

Применение моделей глубокого обучения для прогнозирования финансовых временных рядов

Автор:

Студент М05-217р группы
Зелезецкий Даниил Владимирович

Научный руководитель:

к.ф.-м.н.
Куликов Александр Владимирович

Москва 2024

Оглавление

1	Обзорная часть	5
1.1	Понятие, свойства и характеристики временных рядов	5
1.2	Предобработка временного ряда	6
1.3	Эволюция моделей	6
1.4	Оптимизация портфеля	6
2	Теоретическая часть	7
2.1	Подготовка данных	7
2.2	Выбор моделей и подготовка к испытаниям	7
2.3	Постановка задачи оптимизации портфеля	7
3	Результаты работы	9
3.1	Описание результатов	9
3.2	Анализ и обсуждение	9

Глава 1

Обзорная часть

1.1 Понятие, свойства и характеристики временных рядов

Временной ряд представляет собой упорядоченный набор данных, измеренных в последовательные моменты времени. Эти данные могут быть собраны с определенной периодичностью, например, ежедневно, еженедельно, ежемесячно или в ином временном разрезе. Основное свойство временных рядов заключается в их зависимости от времени, что отличает их от обычных независимых и однородных наборов данных.

Временные ряды широко используются в различных областях для анализа и прогнозирования. Например, в экономике они могут отражать динамику финансовых рынков, показатели инфляции, объемы производства и другие важные экономические показатели. В метеорологии временные ряды могут содержать информацию о погоде, климатических изменениях и т. д.

Для анализа временных рядов применяются различные методы, включая статистические методы, машинное обучение, временные ряды и эконометрику. Целью анализа является выявление закономерностей, трендов, сезонных колебаний, цикличности и других характеристик данных, а также построение моделей для прогнозирования будущих значений временного ряда.

Изучение временных рядов имеет важное практическое значение, поскольку позволяет делать прогнозы и принимать обоснованные решения на основе анализа и интерпретации данных, содержащих информацию о динамике явлений во времени.

Временные ряды обладают рядом особых свойств, которые делают их уникальными и требующими специального подхода к анализу. Вот основные свойства временных рядов:

1. Зависимость от времени: Основное свойство временных рядов заключается в том, что значения в ряду упорядочены по времени. Это означает, что каждое значение зависит от предыдущих и (или) последующих значений в ряду.

2. Тренд: Тренд представляет собой общее направление изменения временного ряда в течение длительного периода времени. Он может быть восходящим, нисходящим или отсутствовать вообще.

3. Сезонность: Сезонность отражает циклические изменения во временном ряде, которые повторяются в течение определенного периода времени. Например, сезонные колебания могут наблюдаться ежегодно, ежемесячно и т.д.

4. Цикличность: Цикличность представляет собой изменения временного ряда, которые происходят не с фиксированной периодичностью, но имеют длительные циклы изменений.

5. Шум (случайная составляющая): Шум или случайная составляющая временного ряда представляет собой непредсказуемые и случайные колебания, которые могут быть вызваны различными факторами и вносят некоторую степень неопределенности в данные.

6. Стационарность: Стационарность временного ряда означает, что его статистические свойства (среднее, дисперсия, автокорреляция и т.д.) не меняются со временем. Это важное свойство, которое облегчает анализ и прогнозирование временных рядов.

7. Автокорреляция: Автокорреляция представляет собой степень линейной зависимости между значениями временного ряда в различные моменты времени. Высокий уровень автокорреляции может указывать на наличие закономерностей в данных и использоваться для прогнозирования.

8. Среднее: Среднее значение временного ряда отражает общую тенденцию данных и может быть использовано для определения тренда или базового уровня ряда.

9. Дисперсия: Дисперсия временного ряда измеряет степень изменчивости данных вокруг их среднего значения. Высокая дисперсия может указывать на большую изменчивость данных, в то время как низкая дисперсия указывает на более стабильные значения ряда.

Временной ряд цен финансовых активов далеко не всегда обязан иметь некоторые перечисленные свойства. Например, типичный ряд цен не имеет постоянного тренда, так как ему свойственно регулярно меняться, в нём не наблюдается явной сезонной составляющей и циклической компоненты, так как доподлинно неизвестны истинные факторы, влияющие на настроения игроков. Финансовые временные ряды персистентны. Это означает, что последующие показатели сильно зависят от предыдущих. К этому близка характерная для хаоса чувствительность к начальным условиям. Определить, персистентен ли временной ряд, позволяет показатель Херста [!!!]. Если данный показатель колеблется в пределах от 0,5 до 1, то временной ряд является персистентным. Херст называл такого рода процессы «процессами с долговременной памятью». Расчет показателя Херста для много численных временных рядов подтверждает их персистентный характер [!!!, !!!].

1.2 Предобработка временного ряда

1.3 Эволюция моделей

1.4 Оптимизация портфеля

Глава 2

Теоретическая часть

2.1 Подготовка данных

2.2 Выбор моделей и подготовка к испытаниям

2.3 Постановка задачи оптимизации портфеля

Глава 3

Результаты работы

3.1 Описание результатов

3.2 Анализ и обсуждение

Литература

- [1] *Zinenko, Anna*. Forecasting financial time series using singular spectrum analysis / Anna Zinenko // *Business Informatics*. — 2023.
- [2] *Б.Б., Иламанов*. МАРТИНГАЛЫ В ФИНАНСОВОЙ МАТЕМАТИКЕ: ВВЕДЕНИЕ И ОСНОВЫ / Иламанов Б.Б. // *ВЕСТНИК НАУКИ*. — 2023. — Vol. 12, no. 3.
- [3] *Staudemeyer, Ralf C*. Understanding LSTM – a tutorial into Long Short-Term Memory Recurrent Neural Networks. — 2019.
- [4] *Chung, Junyoung*. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. — 2014.
- [5] *Vaswani, Ashish*. Attention Is All You Need. — 2023.
- [6] *Chen, Si-An*. TSMixer: An All-MLP Architecture for Time Series Forecasting. — 2023.
- [7] A survey on long short-term memory networks for time series prediction / Benjamin Lindemann, Timo Müller, Hannes Vietz et al. // *Procedia CIRP*. — 2021. — Vol. 99. — Pp. 650–655. — 14th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 15-17 July 2020. <https://www.sciencedirect.com/science/article/pii/S2212827121003796>.
- [8] *moexalgo*. MoexAlgo. — <https://github.com/moexalgo/moexalgo>. — 2023.
- [9] *Bubeck, Sébastien*. Sparks of Artificial General Intelligence: Early experiments with GPT-4. — 2023.
- [10] *SuperGLUE*. GLUE Benchmark leaderboard. — <https://gluebenchmark.com/leaderboard>. — 2024. — Accessed: March 13, 2024.
- [11] *SQuAD*. SQuAD Benchmark leaderboard. — <https://rajpurkar.github.io/SQuAD-explorer/>. — 2024. — Accessed: March 13, 2024.
- [12] *ImageNet*. ImageNet Benchmark leaderboard. — <https://paperswithcode.com/sota/image-classification-on-imagenet>. — 2024. — Accessed: March 13, 2024.
- [13] *Garza, Azul*. TimeGPT-1. — 2023.