МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

им. Н.Э. Баумана

Факультет «Информатика и системы управления»

Кафедра «Системы обработки информации и управления»

ОТЧЕТ

**Домашнее задание №\_\_1\_\_**

по дисциплине«Методы машинного обучения»

Тема: «Машинный перевод»

ИСПОЛНИТЕЛЬ: \_\_\_\_\_Люй Чжэ\_\_\_

ФИО

группа ИУ5И-\_23М\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

подпись

"\_\_"\_\_\_\_\_\_\_\_\_202\_ г.

ПРЕПОДАВАТЕЛЬ: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

ФИО

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

подпись

"\_\_"\_\_\_\_\_\_\_\_\_202\_ г.

Москва - 2024

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Задание

Домашнее задание по дисциплине направлено на анализ современных методов машинного обучения и их применение для решения практических задач. Домашнее задание включает три основных этапа:

1. выбор задачи;
2. теоретический этап;
3. практический этап.

Этап выбора задачи предполагает анализ ресурса [paperswithcode.](https://paperswithcode.com/sota) Данный ресурс включает описание нескольких тысяч современных задач в области машинного обучения. Каждое описание задачи содержит ссылки на наиболее современные и актуальные научные статьи, предназначенные для решения задачи (список статей регулярно обновляется авторами ресурса). Каждое описание статьи содержит ссылку на репозиторий с открытым исходным кодом, реализующим представленные в статье эксперименты. На этапе выбора задачи обучающийся выбирает одну из задач машинного обучения, описание которой содержит ссылки на статьи и репозитории с исходным кодом.

Теоретический этап включает проработку как минимум двух статей, относящихся к выбранной задаче. Результаты проработки обучающийся излагает в теоретической части отчета по домашнему заданию, которая может включать:

1.описание общих подходов к решению задачи;

2.конкретные топологии нейронных сетей, нейросетевых ансамблей или других моделей машинного обучения, предназначенных для решения задачи;

3.математическое описание, алгоритмы функционирования, особенности обучения используемых для решения задачи нейронных сетей, нейросетевых ансамблей или других моделей машинного обучения;

4.описание наборов данных, используемых для обучения моделей;

5.оценка качества решения задачи, описание метрик качества и их значений;

6.предложения обучающегося по улучшению качества решения задачи.

Практический этап включает повторение экспериментов авторов статей на основе представленных авторами репозиториев с исходным кодом и возможное улучшение обучающимися полученных результатов. Результаты проработки обучающийся излагает в практической части отчета по домашнему заданию, которая может включать:

1.исходные коды программ, представленные авторами статей, результаты документирования программ обучающимися с использованием диаграмм UML, путем визуализации топологий нейронных сетей и другими способами;

2.результаты выполнения программ, вычисление значений для описанных в статьях метрик качества, выводы обучающегося о воспроизводимости экспериментов авторов статей и соответствии практических экспериментов теоретическим материалам статей;

3.предложения обучающегося по возможным улучшениям решения задачи, результаты практических экспериментов (исходные коды, документация) по возможному улучшению решения задачи.

## Выбор задачи

Time Series Forecasting is the task of fitting a model to historical, time-stamped data in order to predict future values. Traditional approaches include moving average, exponential smoothing, and ARIMA, though models as various as RNNs, Transformers, or XGBoost can also be applied. The most popular benchmark is the ETTh1 dataset. Models are typically evaluated using the Mean Square Error (MSE) or Root Mean Square Error (RMSE).

## Теоретический этап

Я выбрал две следующие статьи:

"DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks" и "Convolutional Neural Networks for Time Series Classification"

### 3.1 Распространенные решения

В прогнозировании временных рядов с помощью методов глубокого обучения часто используются такие модели, как RNN, LSTM, CNN и, в последнее время, Transformer. Эти модели способны улавливать временные зависимости в данных, особенно эффективны в случаях, когда необходимо учесть долгосрочные зависимости.

### 3.2 Конкретные топологии нейронных сетей

1.**DeepAR** использует рекуррентные нейронные сети, в частности, сети с элементами управляемых рекуррентных блоков (GRU), для моделирования динамики временных рядов. Вот как устроена модель DeepAR:

* Входные данные: DeepAR принимает входные данные временного ряда, такие как продажи продукта по дням или часам. Эти данные часто дополняются внешними переменными, которые могут влиять на прогнозы, например, праздники или погодные условия.
* Эмбеддинги: Чтобы улучшить обработку категориальных переменных (например, тип продукта или магазин), используются эмбеддинги, которые преобразуют эти переменные в плотные векторы небольшой размерности.
* Рекуррентные слои: DeepAR использует слои GRU (Gated Recurrent Unit) или LSTM (Long Short-Term Memory) для обработки входных данных. Эти слои способны сохранять информацию о предыдущем состоянии временного ряда, что позволяет модели учитывать долгосрочные зависимости в данных.
* Выходной слой: Вместо предсказания одного значения на выходе, DeepAR предсказывает параметры вероятностного распределения (обычно гауссовского), такие как среднее значение и стандартное отклонение для каждого временного шага. Это позволяет оценить неопределённость прогнозов.

2.**CNN** в статье по классификации временных рядов использует многослойные сверточные сети для извлечения признаков из временных рядов, что очень эффективно для понимания локальных паттернов в данных.

* Сверточные слои: В основе CNN лежат сверточные слои, которые применяют различные фильтры к временным рядам для извлечения признаков. Эти фильтры автоматически обнаруживают важные характеристики данных, такие как взлёты и падения или циклические изменения.
* Пулинг: Следующий за сверточными слоями, слой пулинга уменьшает размерность данных, подытоживая информацию внутри определённого окна (например, максимальное или среднее значение).
* Полносвязные слои: После сверточных и пулинг слоев, данные обычно подаются на один или несколько полносвязных слоев, которые служат для дальнейшей обработки признаков и формирования конечного прогноза.
* Регуляризация: Чтобы предотвратить переобучение, в CNN часто используются техники регуляризации, такие как Dropout, которые случайно исключают некоторые нейроны во время обучения.

### 3.3 Математическое описание и алгоритмы функционирования

1.**DeepAR** предсказывает будущие значения с помощью условных вероятностных моделей, используя структуру, похожую на авторегрессионные модели, но приспособленную к выявлению нелинейных связей в данных с помощью глубокого обучения.

2.**CNN** использует фильтры (или ядра свертки) для извлечения ключевых особенностей из временных рядов и снижает размерность признаков с помощью слоев пулинга, что повышает способность модели к абстракции.

### 3.4 Описание наборов данных

Исследования обычно используют общедоступные наборы данных для временных рядов, такие как классификационные/регрессионные данные временных рядов UCI, данные соревнований Kaggle и т. д. Выбор набора данных должен базироваться на конкретных потребностях задачи прогнозирования.

### 3.5 Оценка качества решения задачи

Как правило, используют такие метрики, как среднеквадратическая ошибка (MSE), корень из среднеквадратической ошибки (RMSE) и средняя абсолютная ошибка (MAE). Эти показатели помогают оценить точность прогнозов модели.

### 3.6 Предложения по улучшению качества решения задачи

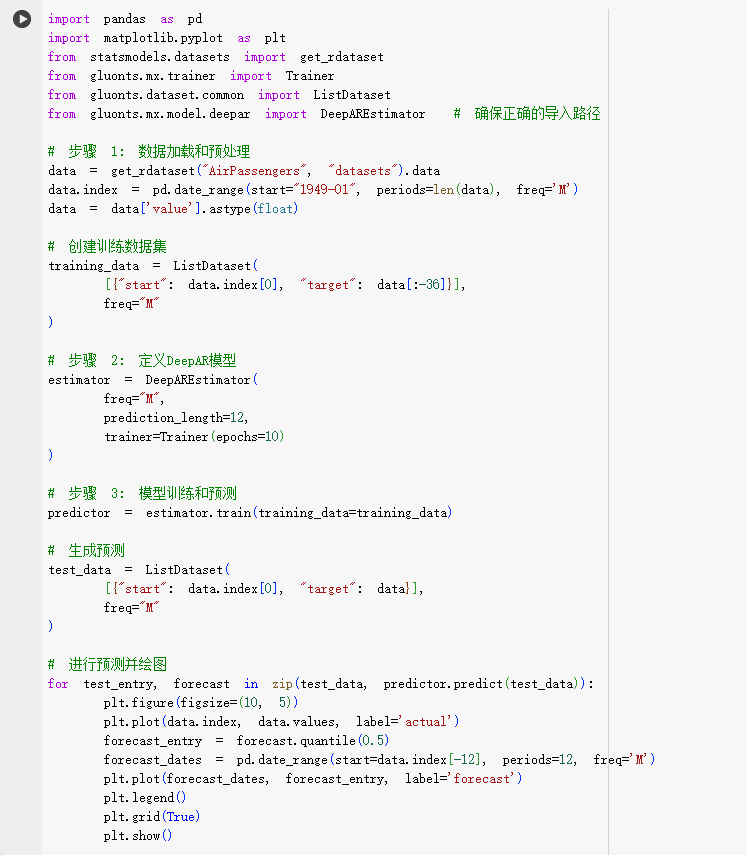
1.Увеличение сложности модели или изменение её архитектуры, например, использование двунаправленных LSTM или добавление слоев в сеть.

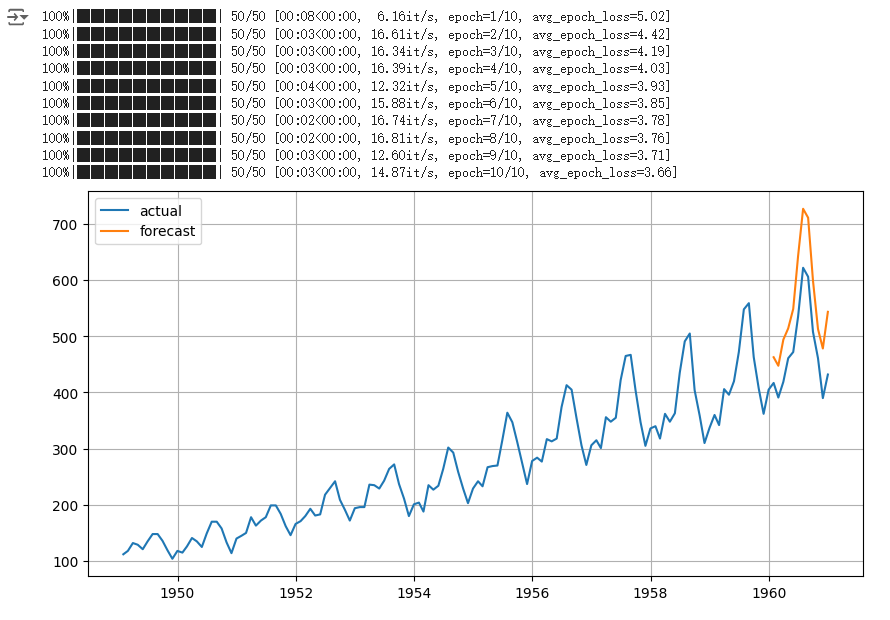
2.Применение техник аугментации данных, таких как сдвиги и масштабирование временных рядов, для улучшения способности модели к обобщению.

3.Использование различных техник регуляризации, таких как Dropout, для снижения переобучения модели.

## Практический этап

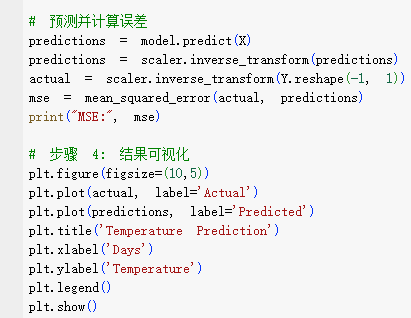
**DeepAR**

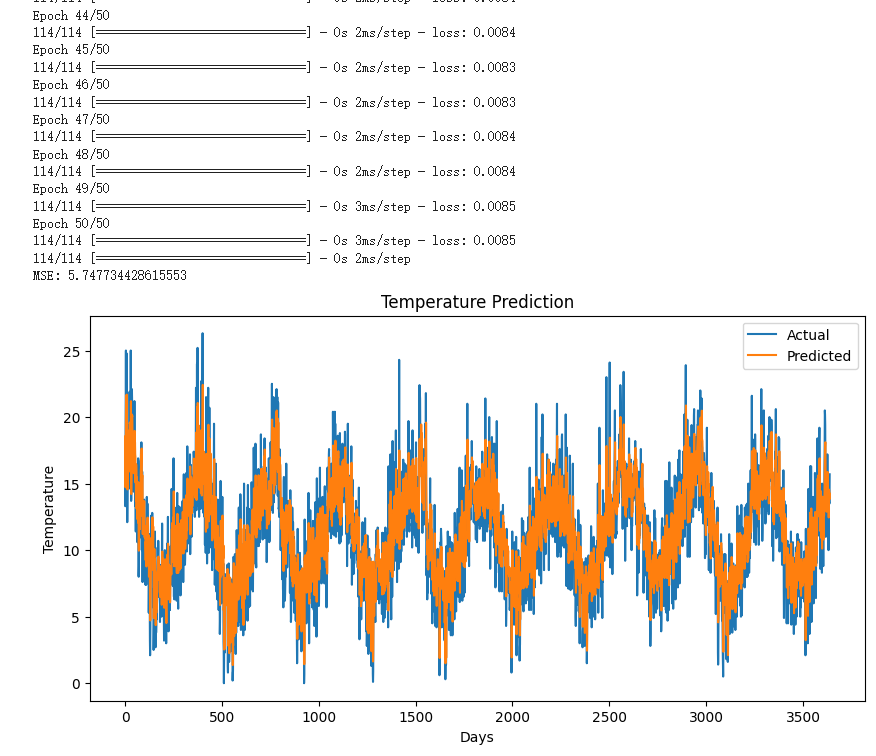




**СNN**







**Вывод**

В этом эксперименте мы рассмотрели два разных метода для выполнения задачи прогнозирования временных рядов, используя DeepAR и сверточные нейронные сети (CNN). Сравнивая эти два метода, мы можем получить более глубокое понимание их эффективности и применимости для обработки временных рядов.

1. Обзор методов

**DeepAR**: Это вероятностный метод прогнозирования, который использует авторегрессионные рекуррентные нейронные сети для моделирования временных рядов. DeepAR особенно подходит для серий с сложными сезонными паттернами и потенциальной нестабильностью. Этот метод может эффективно обрабатывать большое количество параллельных временных рядов и предоставлять вероятностные оценки для прогнозов.

**CNN**: использует одномерные сверточные нейронные сети, которые обычно применяются для обработки изображений и звука, но также постепенно находят применение в обработке последовательностей данных. CNN обрабатывают временные ряды, захватывая локальные зависимости с помощью скользящих окон свертки и изучая более сложные временные паттерны с помощью нескольких слоев свертки.

2. Сравнение производительности

**Точность прогноза**: В конкретных экспериментах точность прогнозирования DeepAR и CNN может варьироваться в зависимости от выбранного набора данных и конфигурации. Обычно, DeepAR может лучше справляться с очень нерегулярными и шумными временными рядами благодаря своей способности лучше улавливать неопределенности данных. В то время как CNN может лучше работать с временными рядами, которые имеют четкие циклические и трендовые характеристики.

**Вычислительная сложность**: Модели CNN обычно требуют меньше настройки параметров и меньше времени на обучение, в то время как модели DeepAR могут требовать более длительных циклов обучения и более тщательной оптимизации параметров для захвата динамики временных рядов.

3. Сценарии применения

DeepAR подходит для прогнозирования финансовых рынков, прогнозирования спроса, управления запасами и других областей, особенно там, где важны прогнозные интервалы и оценка неопределенности.

CNN может быть более подходящим для структурированных временных рядов с выраженными циклическими и трендовыми характеристиками, например, для анализа температурных данных или рыночных тенденций.

4. Заключение

В общем и целом, выбор модели должен основываться на потребностях конкретной задачи и характеристиках данных. Если нужны точные вероятностные оценки для сложных прогнозов временных рядов, DeepAR предоставляет мощные инструменты; для быстрого и наглядного анализа трендов CNN может предложить простое и эффективное решение. В реальных приложениях мы можем выбирать наиболее подходящую модель в зависимости от особенностей данных и бизнес-требований.

## Список использованных источников

[1] D. Salinas, V. Flunkert, J. Gasthaus, and T. Januschowski, "DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks," International Journal of Forecasting, 2019.

[2] Z. Wang, W. Yan, and T. Oates, "Time Series Classification from Scratch with Deep Neural Networks: A Strong Baseline," in Proc. of the International Joint Conference on Neural Networks (IJCNN), 2017.

[3] <https://github.com/jdb78/pytorch-forecasting>

[4]https://github.com/pytorch/vision/blob/6db1569c89094cf23f3bc41f79275c45e9fcb3f3/torchvision/models/resnet.py#L124