

Сравнение методов обучения и архитектур рекуррентных нейронных сетей в задаче прогнозирования

Выполнил:

Руководитель:

Консультант:

Зуев Даниил Владимирович, гр. 7383

Лисс Анна Александровна, к.т.н., доцент

Григорьевна Наталия Юрьевна, к.ф.-м.н

Цель и задачи

Актуальность: применение в предупреждении экологических проблем, использование полученных методологий для построения моделей рекуррентных нейронных сетей, решающих задачу прогнозирования, в других областях.

Цель: исследование рекуррентных нейронных сетей, решающих задачу прогнозирования значения содержания хлорофилла *a* в водоеме.

Задачи:

1. Определить параметры, от которых зависит результат обучения моделей нейронных сетей
2. Обучить наборы моделей нейронных сетей с заданными параметрами
3. Определить показатели качества моделей нейронных сетей
4. Исследовать и сравнить полученные модели

Исходные данные

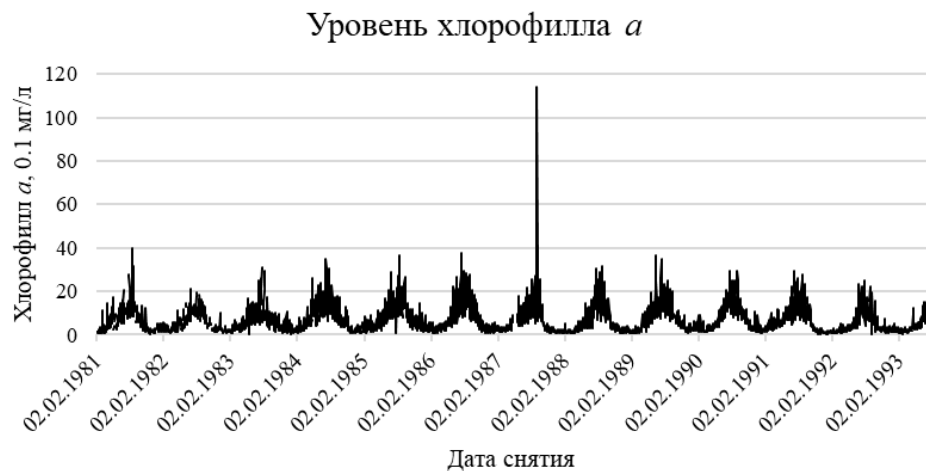


Рисунок 1 – Значения содержания хлорофилла *a* из первого набора



Рисунок 2 – Значения содержания хлорофилла *a* из второго набора

Рекуррентные нейронные сети

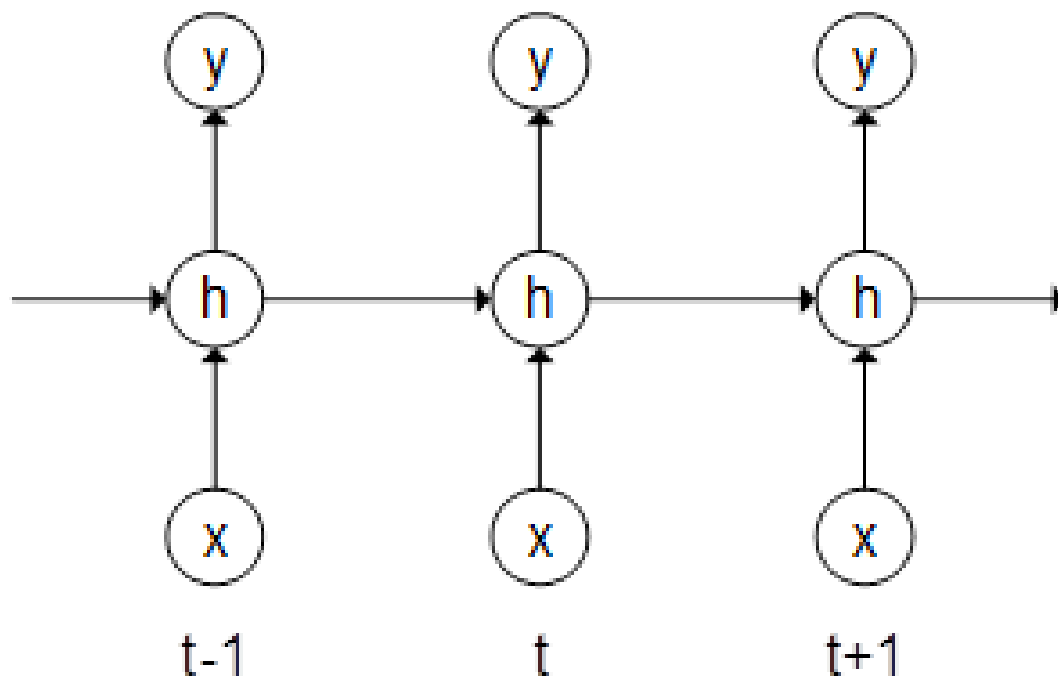
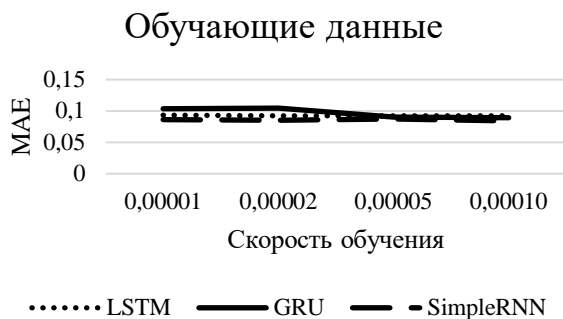
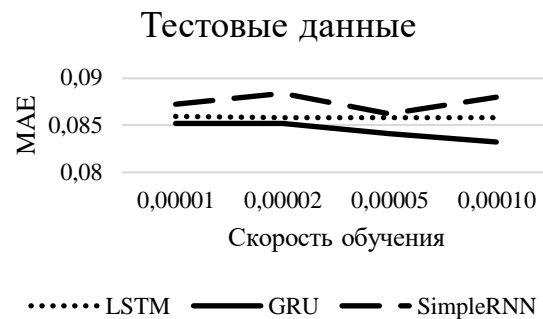


Рисунок 3 – Принцип работы нейрона рекуррентной нейронной сети

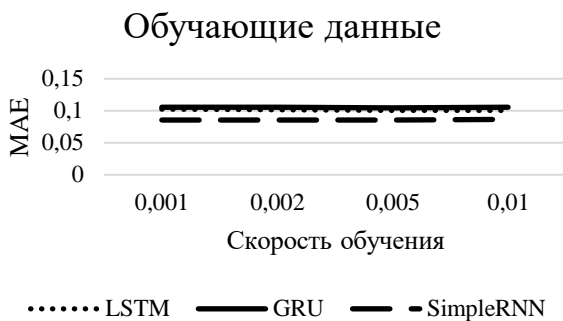
Анализ скорости обучения



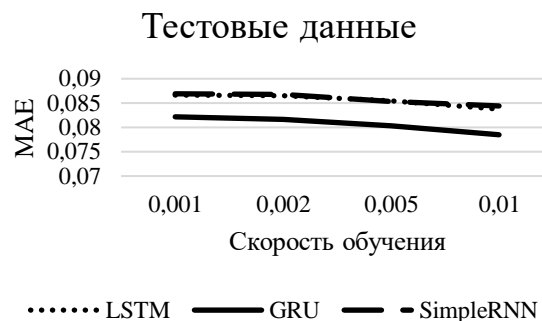
а)



б)



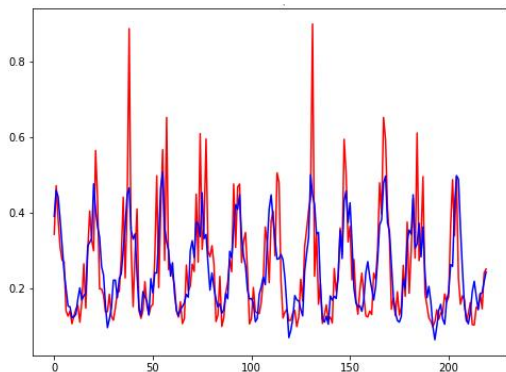
в)



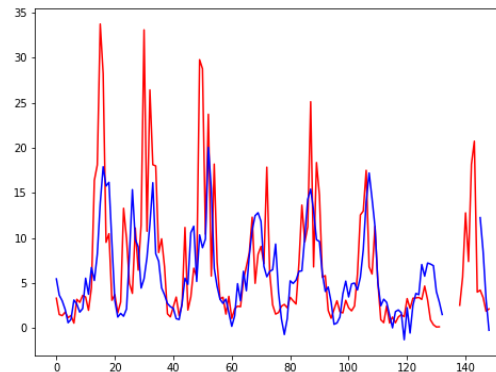
г)

Рисунок 4 - Зависимость MAE от скорости обучения: а) на обучающих данных при использовании оптимизатора Adam, б) на тестовых данных при использовании оптимизатора Adam, в) на обучающих данных при использовании оптимизатора SGD, г) на тестовых данных при использовании оптимизатора SGD

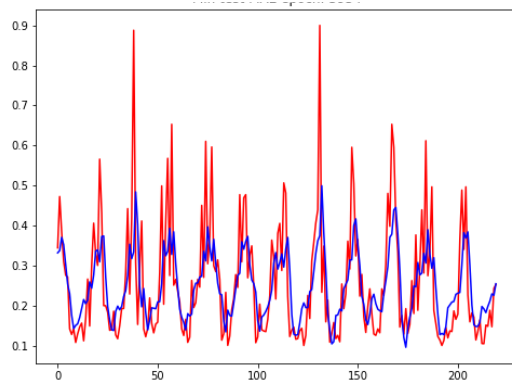
Анализ графиков предсказания



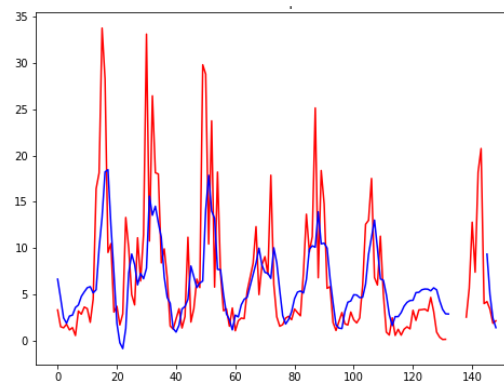
а)



б)



в)



г)

Рисунок 5 - Графики предсказания моделями с архитектурой 1-10-5-1: а) с типом блока SimpleRNN на обучающих и тестовых данных, б) с типом блока SimpleRNN на контрольных данных, в) с типом блока GRU на обучающих и тестовых данных, г) с типом блока GRU на контрольных данных.

Анализ графиков предсказания

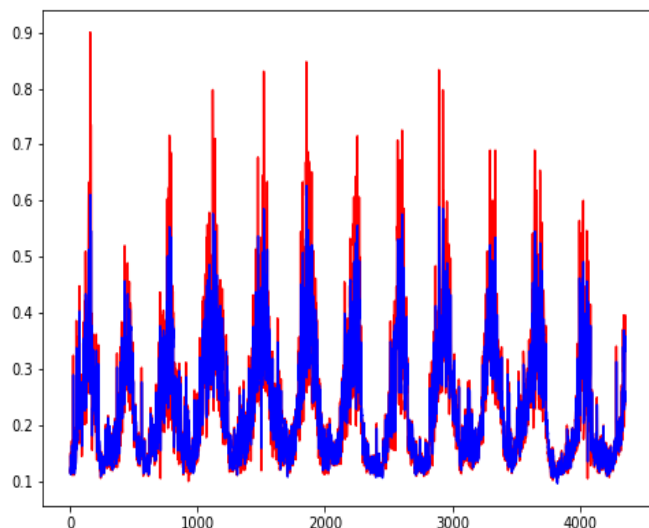
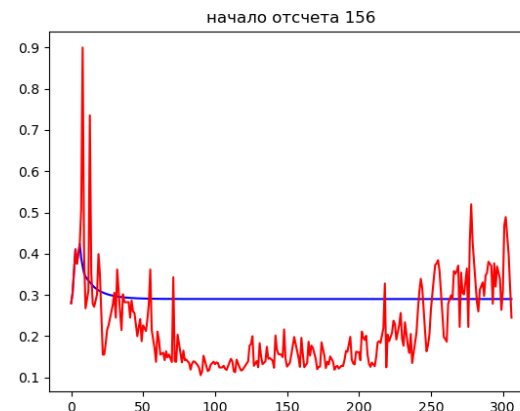
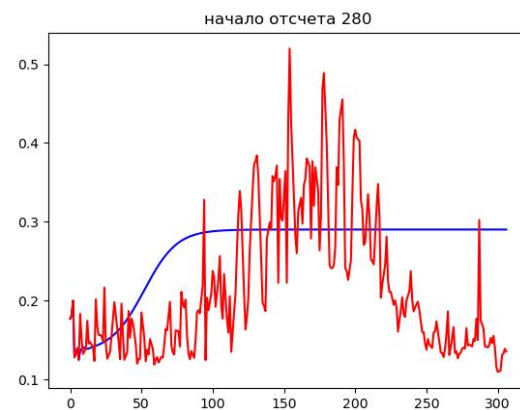


Рисунок 6 - График предсказания на ежедневных данных



а)



б)

Рисунок 7 - Графики рекуррентного предсказания на данных без выбросов: а) моделью с типом блока SimpleRNN, б) моделью с типом блока GRU

Анализ графиков предсказания

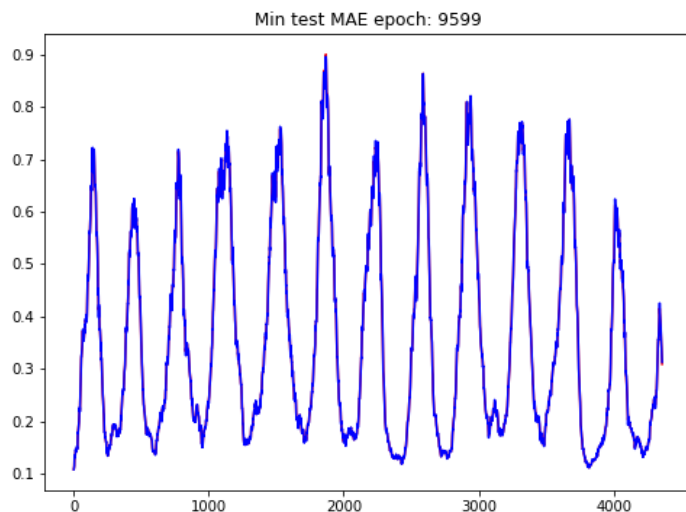
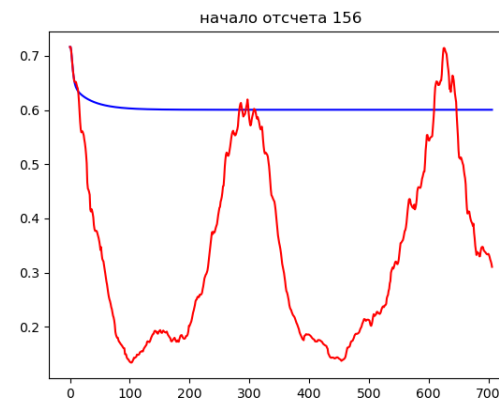
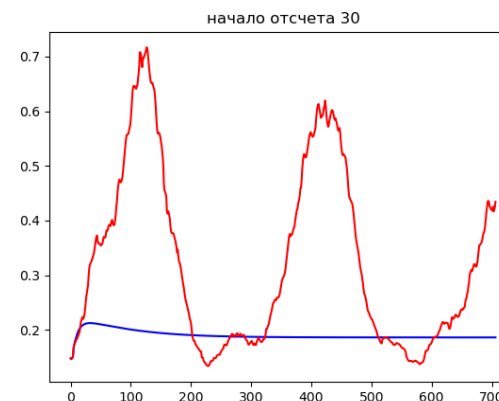


Рисунок 8 - Предсказание по сглаженным данным



а)



б)

Рисунок 9 - Графики рекуррентного предсказания на сглаженных данных: а) моделью с рекуррентным блоком типа SimpleRNN, б) моделью с рекуррентным блоком типа GRU

Тип данных «гребенка»

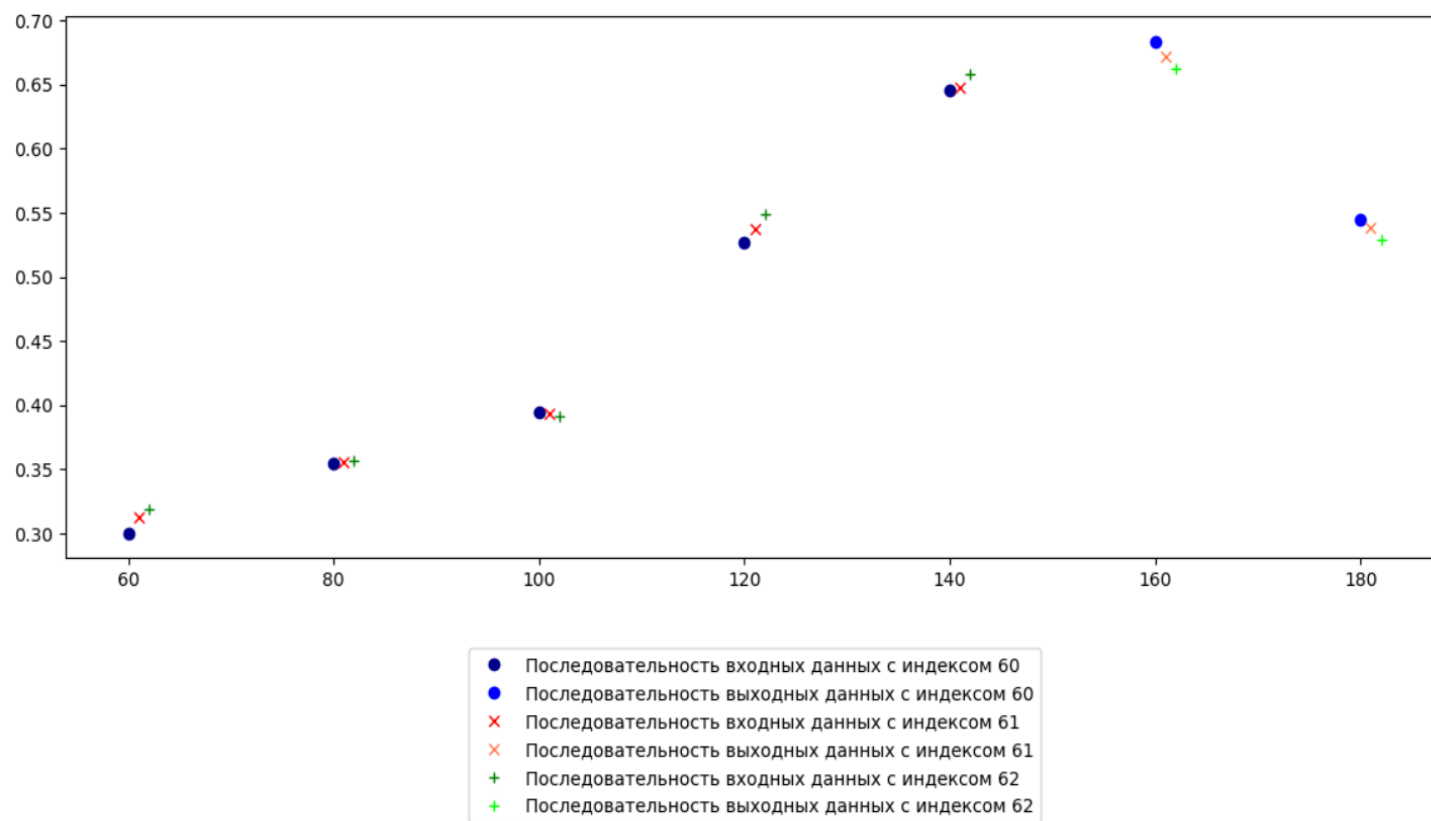


Рисунок 10 - Графическое представление
разреженных последовательностей

Анализ графиков предсказания по «гребенке»

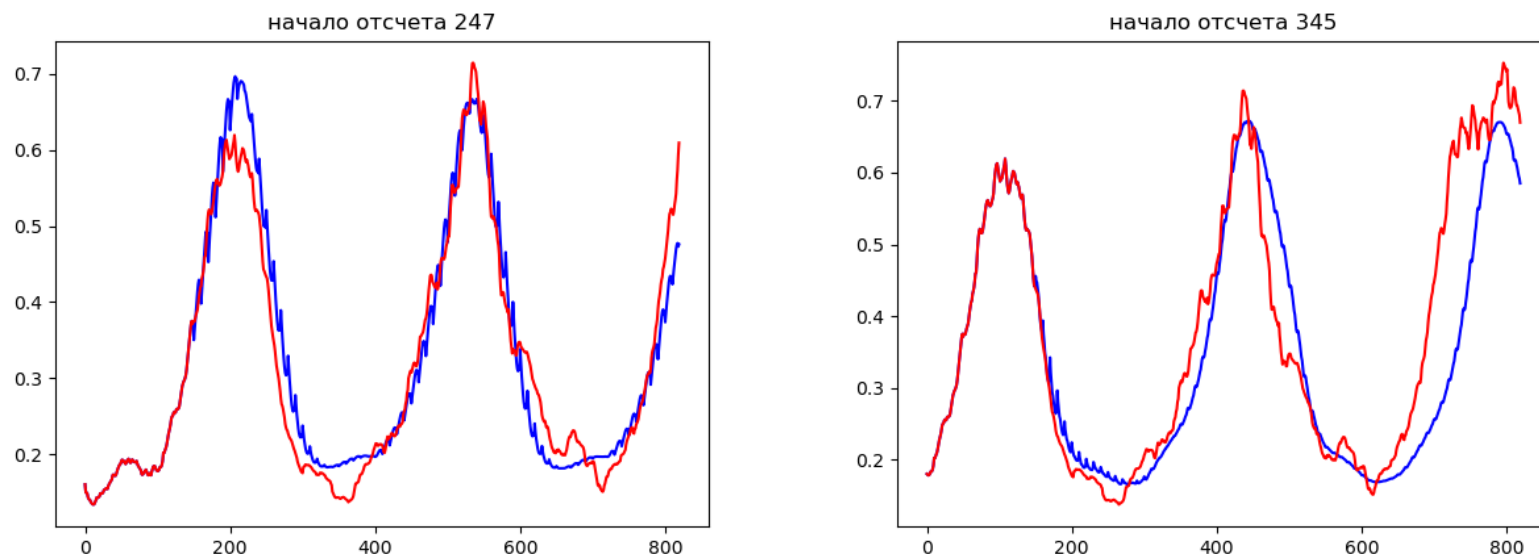


Рисунок 4.15 - Графики рекуррентного предсказания: а) моделями с типом блока GRU, б) моделями с типом блока LSTM.

Выводы

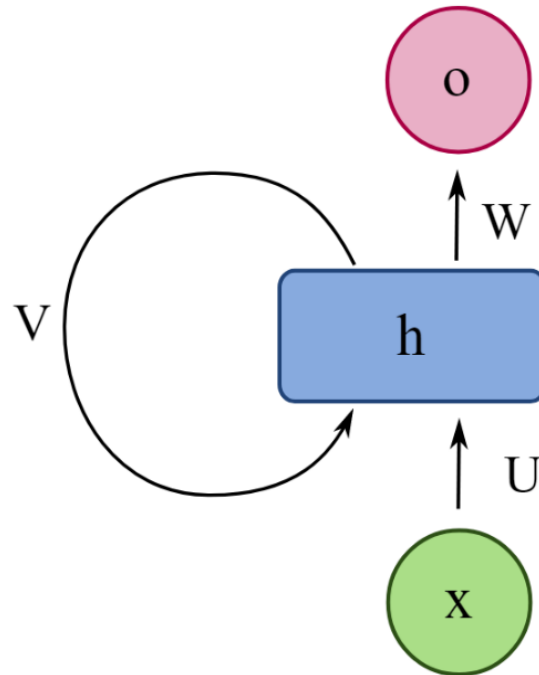
- Были определены параметры, влияющие на обучение и результат обучения моделей: метод обучения, скорость обучения, архитектура модели, архитектура рекуррентного блока и тип данных.
- Зависимость от метода обучения показала, что SGD лучше работает на малом количестве данных, однако он нестабилен, а Adam работает стабильно и на малых, и на больших количествах данных.
- При недостаточной скорости обучения модель может не достигнуть глобального минимума, а при слишком большой скорости обучения обучение модели происходит нестабильно.
- При слишком сложной архитектуре модели или рекуррентного блока количество весов модели может быть слишком большим и модели сложнее достичь глобального минимума, однако если она его достигнет, то она будет предсказывать лучше, чем менее сложная.
- При входных данных, охватывающих большой промежуток времени, модель способна рекуррентно предсказывать значения, чего она не может при небольших промежутках времени.

Апробация

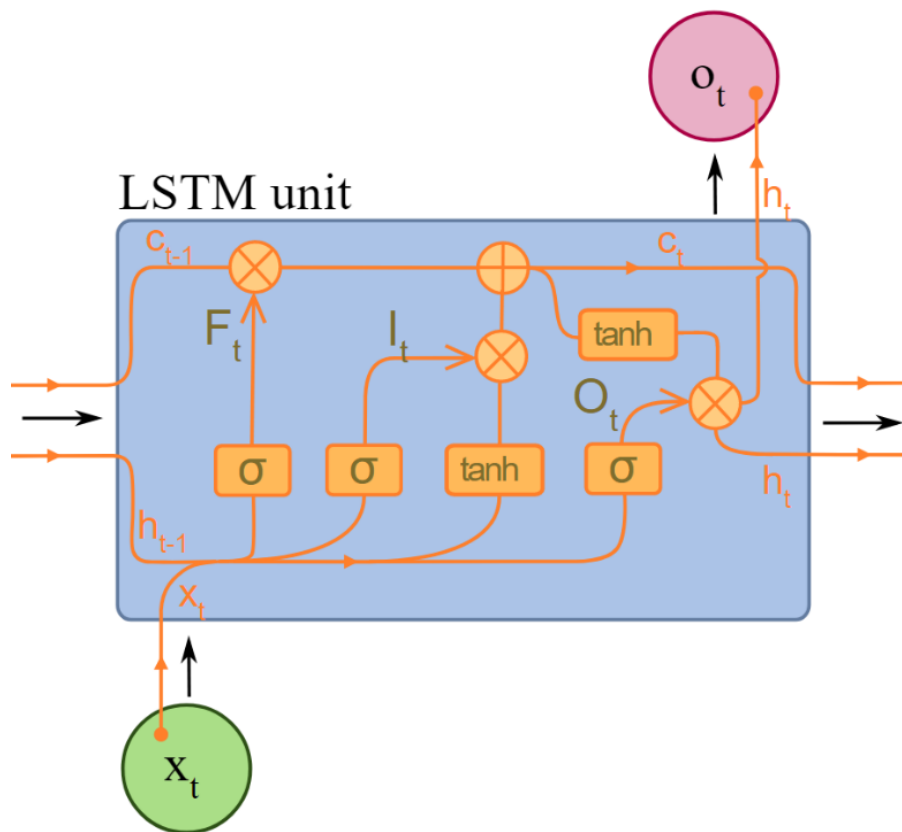
- Зуев Д.В., Жангиров Т.Р. Сравнение методов обучения и архитектур рекуррентных нейронных сетей в задаче прогнозирования / Сборник докладов студентов и аспирантов на научно-техническом семинаре кафедры МОЭВМ. СПб, 2021 г. / СПбГЭТУ «ЛЭТИ», СПб, 2021, С. 49-55.

Запасные слайды

Архитектура блока полностью рекуррентной нейронной сети



Архитектура блока LSTM



Архитектура блока GRU

