

Анализ влияния количества переобучений нейронной сети на качество прогнозирования временных рядов

Курдошев З.М.

Томский государственный университет

2023

Аннотация

Прогнозирование временных рядов является одной из важнейших проблем на сегодняшний день, и вопрос прогнозирования не утратил своей актуальности. Точность прогнозов сложна даже для рекуррентных нейронных сетей из простых нейронных сетей.

Введение

Основываясь на результатах этих исследований, он широко используется в статистике, обработке сигналов, обнаружении закономерностей, эконометрике, математических финансах, прогнозировании погоды, предсказании землетрясений, электроэнцефалографии, технике управления, астрономии, технике связи. Приложения для прогнозирования временных рядов с помощью искусственных нейронных сетей можно найти в [1,2,3].

Введение

Возникает вопрос, насколько важен выбор количества переобучений в процессе прогнозирования такими рекуррентными нейронными сетями. Поскольку возник этот вопрос, для определения точности прогнозов была изучена взаимосвязь между количеством переобучений, оптимизаторами и функциями потерь. В исследовании сравнивались результаты комбинации широко используемых функций потерь (MSE, MAE, MAPE, SMAPE, RMSLE) и современных оптимизаторов (SGD, ADAM, ADAGRAD). Следует также отметить, что эти комбинации использовались в рекуррентных сетях, таких как RNN, LSTM и GRU.

Рекуррентные нейронные сети

Рекуррентные нейронные сети — это тип архитектуры нейронной сети, который в основном используется для последовательной обработки данных. Основное различие между сетями RNN и обычными нейронными сетями заключается в том, что в традиционных нейронных сетях входом является только x_t , а в рекуррентных сетях используется не только x_t , но и x_{t-1} . Несколько скрытых слоев объединяются в один блок.

Исследования

В качестве объекта исследования была взята цена иностранной валюты на финансовом рынке по сравнению с национальной валютой. Записывается ежедневная цена иностранной валюты с 2010 по 2023 год. 70 % собранных данных используются для обучения нейронной сети, а оставшиеся 30% — для тестирования.

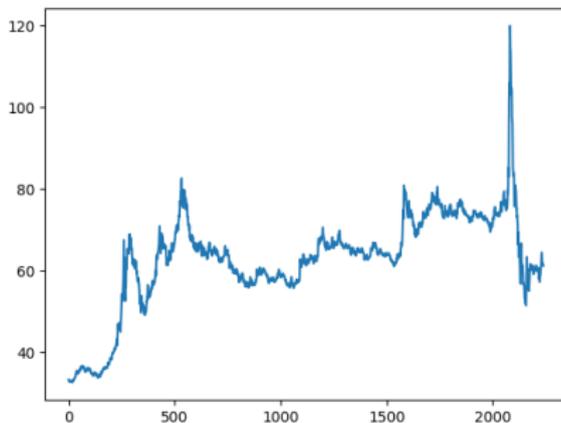


Рис.: График зависимости иностранной валюты от национальной валюты.

RNN-MAE-ADAGRAD

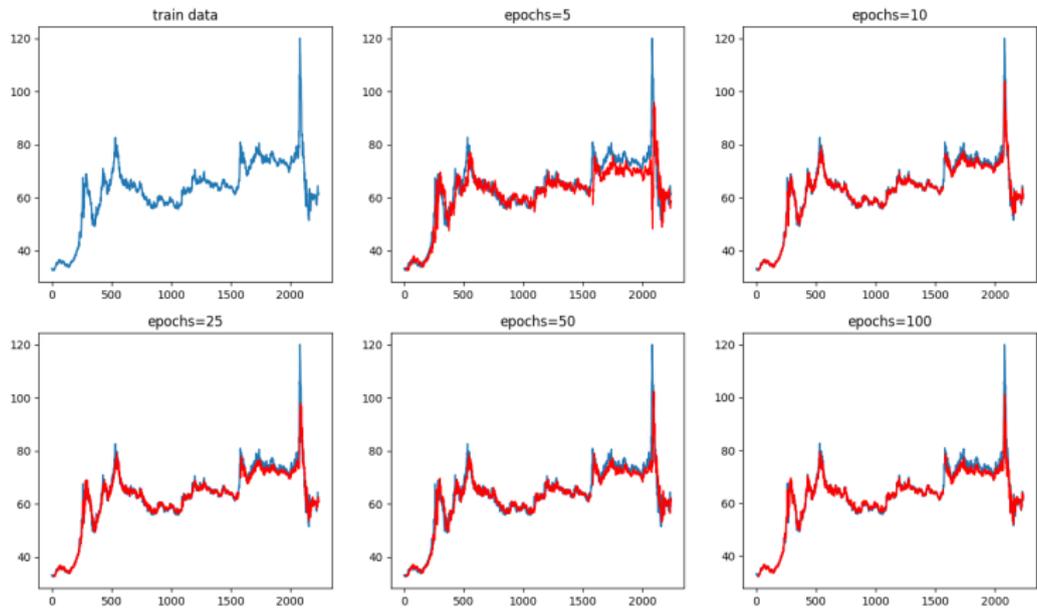


Рис.: Прогнозирование временных рядов с помощью нейронной сети RNN, оптимизатора ADAGRAD и различного количества переобучений.

RNN-MAE-ADAGRAD

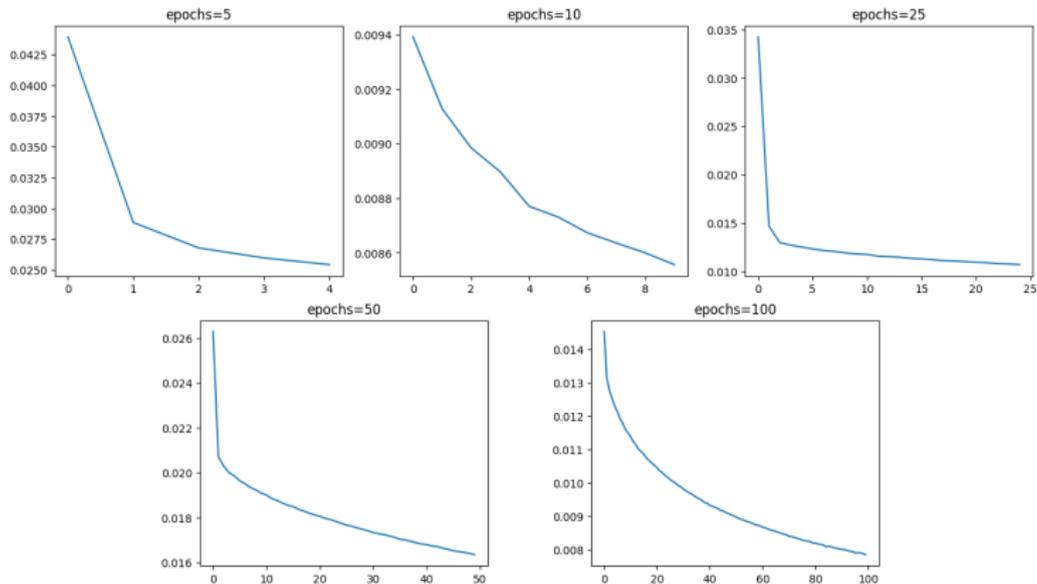


Рис.: График состояния функций потерь.

LSTM-MAPE-ADAGRAD

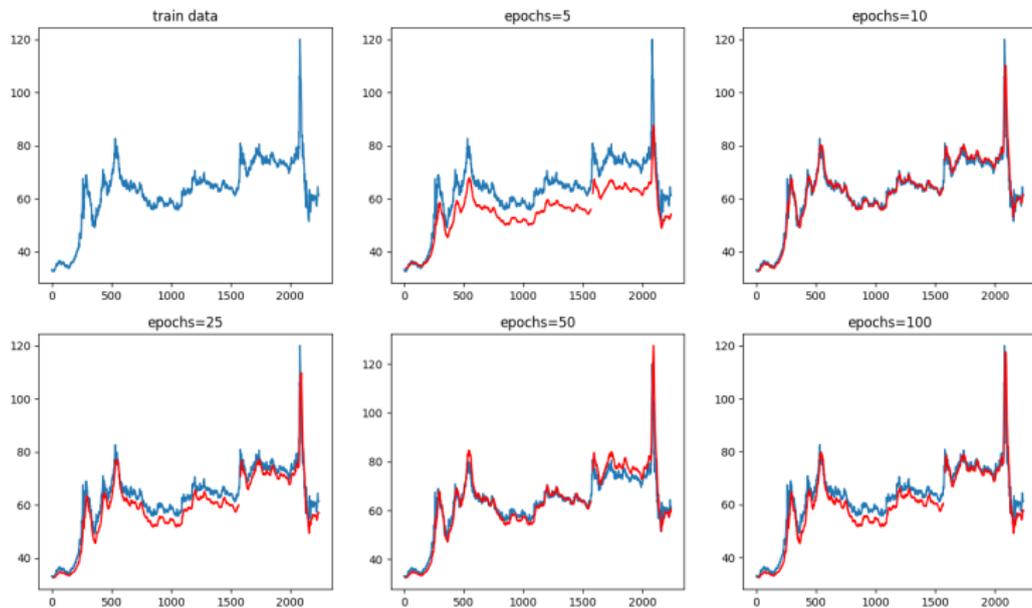


Рис.: Прогнозирование временных рядов с помощью нейронной сети LSTM, оптимизатора ADAGRAD и различного количества переобучений.

LSTM-MAPE-ADAGRAD

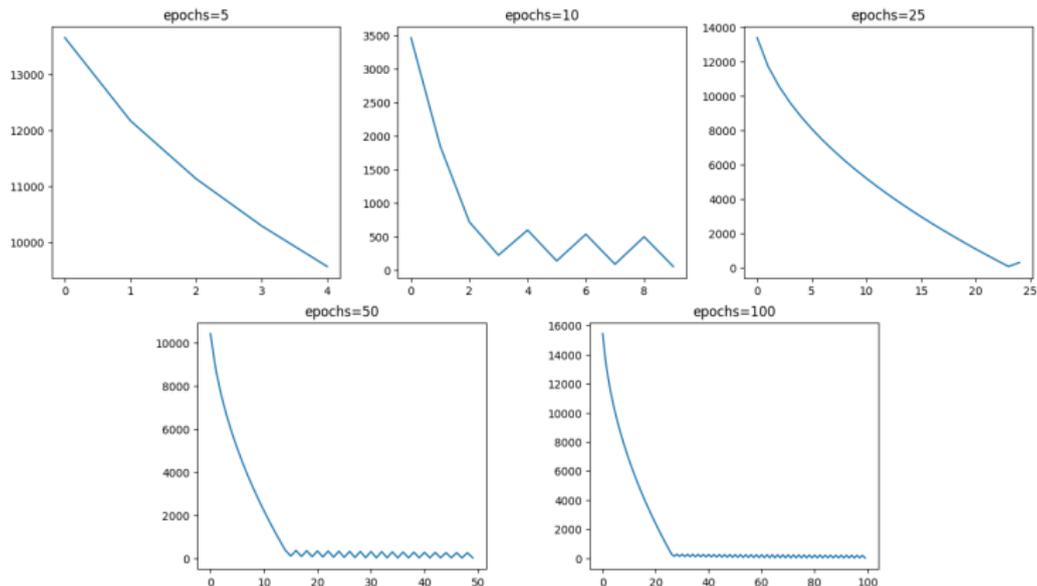


Рис.: График состояния функций потерь.

LSTM-MAPE-ADAM

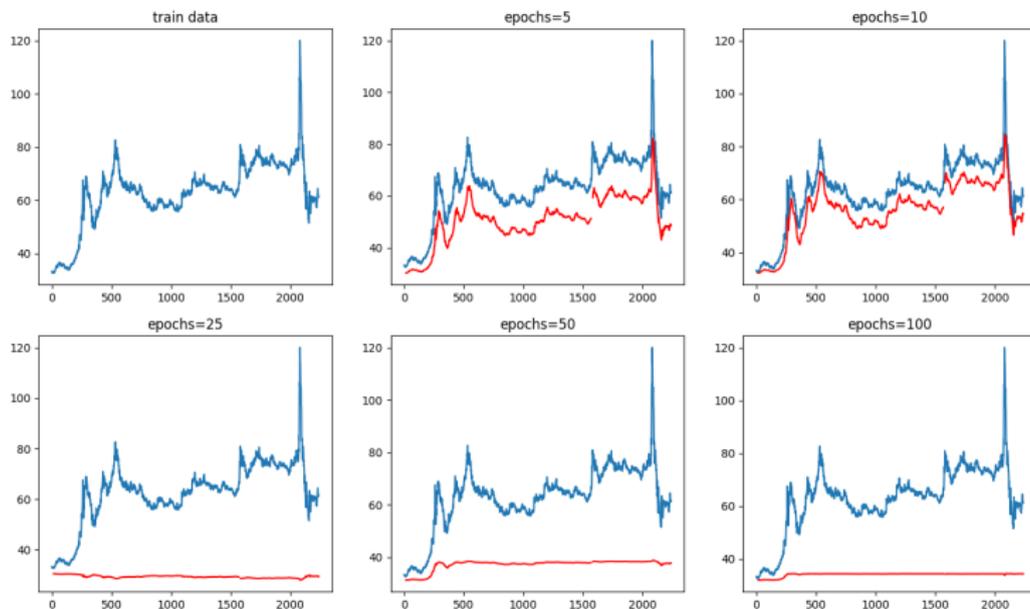


Рис.: Прогнозирование временных рядов с помощью нейронной сети LSTM, оптимизатора ADAM и различного количества переобучений.

LSTM-MAPE-ADAM

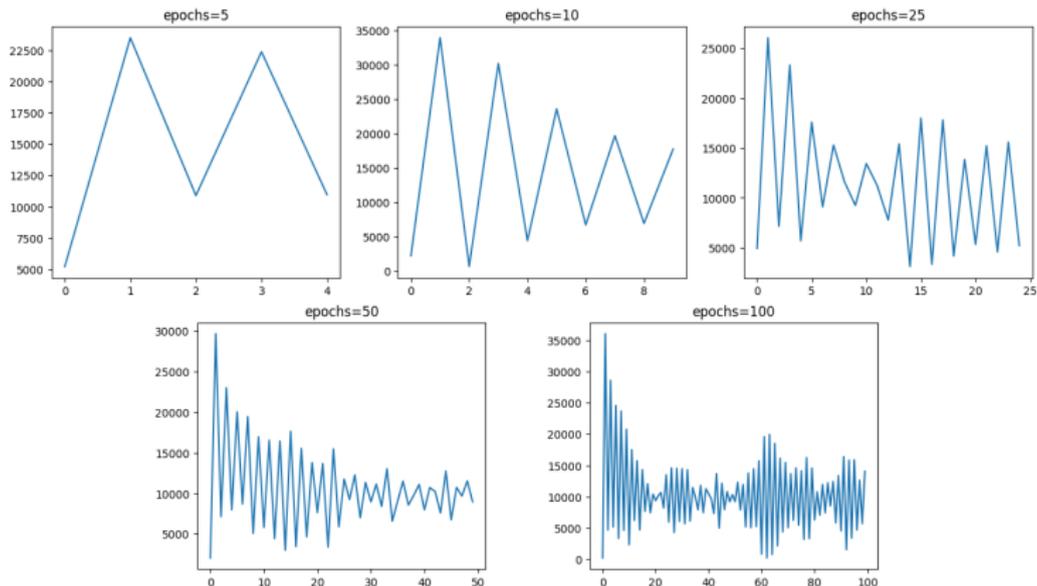


Рис.: График состояния функций потерь.

Заключение

Полученные результаты показывают, что прогнозы зависят не только от функций потерь и оптимизаторов, но и от количества переобучений. Как видно из графика, в некоторых моделях результаты улучшались по мере увеличения количества переобучений, а в некоторых моделях лучшие результаты достигались при определенном количестве переобучений.

Sulandari, W., Suhartono, Subanar & Rodrigues, P. Exponential smoothing on modeling and forecasting multiple seasonal time series: An overview. *Fluctuation And Noise Letters*. **20**, 2130003 (2021)

Ariyo, A., Adewumi, A. & Ayo, C. Stock price prediction using the ARIMA model. *2014 UKSim-AMSS 16th International Conference On Computer Modelling And Simulation*. pp. 106-112 (2014)

Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Computation*. **9**, 1735-1780 (1997)