

Совместное применение рекуррентных нейронных сетей и статистических методов для увеличения точности прогнозирования океанологических данных

В. Ю. Кузьмин¹

Данная работа посвящена вопросам применения статистических моделей для увеличения точности прогнозирования океанологических данных. Исходный ряд наблюдений описывается с помощью смесей конечных нормальных распределений, статистические характеристики смесей используются для предварительной инициализации слоёв рекуррентной нейронной сети. Прогнозы, построенные с использованием статистических моделей, сравниваются с прогнозами, построенными для оригинальных данных. Продемонстрировано, что для всех проанализированных рядов наблюдается значительное улучшение точности.

Ключевые слова: EM, MSM, LSTM, нейронные сети, машинное обучение, расширение признакового пространства, смеси конечных нормальных распределений.

1. Введение

Нейронные сети могут быть эффективно использованы в задачах построения прогнозов и предсказания эволюции самых различных процессов. В частности, они применяются при предсказании поведения финансовых индексов [1], [2], интенсивности коммуникационного трафика [3], поведения турбулентной плазмы [4] и метеорологических данных [5].

Повышение точности прогнозов чаще всего требует либо усложнения обучаемой модели, либо сбора дополнительных данных. Оба варианта приводят к дополнительным затратам времени и ресурсов. При этом сбор дополнительных данных может быть невозможен при работе с историческими данными.

В данной работе будет обсуждено повышение точности прогнозирования с применением статистических характеристик исследуемых данных. Исходный нестационарный ряд разбивается на набор окон с применением

¹ Кузьмин Виктор Юрьевич — программист ф-та космических исследований МГУ, e-mail: shadesilent@yandex.ru

Kuzmin V.Y. — programmer, Lomonosov Moscow State University, Faculty of Space Research

метода скользящего среднего, на каждом окне создаётся статистическая модель с помощью алгоритмов EM-типа [6],[7].

2. Методология

В качестве исходных данных выбраны шесть рядов океанологических данных. Каждый ряд содержит наблюдения за 10 лет, интервал между наблюдениями составляет 6 часов, всего в ряду 14600 наблюдений. Исходные нестационарные ряды с помощью метода скользящего среднего разбиваются на набор окон. Ширина окна выбрана в 200 наблюдений или 50 дней. Размер окна является компромиссом между необходимым количеством данных для построения статистической модели и затуханием влияния более ранних наблюдений на предсказываемый временной период.

На каждом окне с помощью метода максимизации функции правдоподобия [6] строится статистическая модель – конечная смесь нормальных распределений. В ходе исследования были рассмотрены смеси из трёх, четырех и пяти компонент. Поскольку большую часть времени в смеси выделены одна либо две основные компоненты, в качестве итогового решения была выбрана трёхкомпонентная смесь. Для статистического обогащения данных будут применены первые четыре момента построенной модели – математическое ожидание, дисперсия, коэффициенты асимметрии и эксцесса.

Для каждого ряда рассматривается два типа прогнозов – краткосрочный и среднесрочный. Краткосрочный прогноз – прогнозирование 12 наблюдений по исходной выборке в 200 наблюдений. Среднесрочный прогноз – прогнозирование 12 наблюдений по выборке в 200 наблюдений с пропуском 28 наблюдений между концом выборки и началом прогноза.

Для каждого прогноза создаётся и оптимизируется две нейронные сети – сеть с входом из исходных данных и сеть с входом из исходных данных и инициализацией скрытого состояния сети с помощью моментов статистической модели [8]. В качестве архитектуры сети выбрана рекуррентная нейронная LSTM-сеть. В качестве метрики ошибок используется RMSE. Количество и размер слоёв, dropout, оптимизатор и иные параметры сети подбираются с помощью сеточного метода оптимизации гиперпараметров. Для построения нейронной сети использовался пакет Tensorflow для языка программирования Python, вычисления производились на гибридном высокопроизводительном вычислительном кластере.

3. Результаты

В таблице приведено относительное уменьшение ошибки в метрике RMSE для прогнозов, построенных с помощью статистических моделей по сравнению с прогнозами, построенными традиционным образом.

	1	2	3	4	5	6
Краткосрочный	15%	6%	10%	7%	1%	5%
Среднесрочный	17%	7%	8%	6%	1%	3%

Таблица 1. Результаты

Характерная ошибка составляет 0.07 по метрике RMSE для краткосрочных прогнозов и 0.09 для среднесрочных. Применение статистического обогащения позволяет уменьшить ошибку в среднем на 6%. При этом статистическое обогащение не приводит к значительному замедлению обучения модели и не приводит к ухудшению точности ни в одном из рассмотренных случаев.

4. Выводы

Применение метода статистического обогащения приводит к заметному улучшению точности прогнозирования проанализированных рядов. При этом построение статистической модели не является задачей с высокой вычислительной сложностью и не требует сбора дополнительных данных. Поскольку построенный набор моделей является локальным, то появление новых данных не приводит к необходимости перестраивания всего массива построенных моделей, что позволяет применять подход при обработке real-time задач.

Исследование выполнено при поддержке Междисциплинарной научно-образовательной школы Московского университета «Мозг, когнитивные системы, искусственный интеллект»

Список литературы

- [1] Cao, Jian, Zhi Li, and Jian Li., “Financial time series forecasting model based on CEEMDAN and LSTM”, *Physica A: Statistical Mechanics and its Applications*, 2019, № 519, 127-139.
- [2] Yan, Binbin, and Memon Aasma., “A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM.”, *Expert systems with applications*, 2020, 159.
- [3] Du, Xunsheng, et al., “Stacked LSTM deep learning model for traffic prediction in vehicle-to-vehicle communication”, *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017.

- [4] Meneghini O., Luna C.J., Smith S.P., Lao L.L., “Modeling of transport phenomenon in tokamak plasmas with neural networks”, *Physics of Plasmas*, **21:6** (2014), 060702-0607024.
- [5] Gorshenin A. K., Kuzmin V. Y., “Neural network forecasting of precipitation volumes using patterns”, *Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications.*, **28:3** (2018), 450–461.
- [6] Королев В.Ю., *Вероятностно-статистические методы декомпозиции волатильности хаотических процессов*, Издательство Московского университета, Москва, 2011, 512 pp.
- [7] Gorshenin A. K., Kuzmin V. Y., “A machine learning approach to the vector prediction of moments of finite normal mixtures”, *Advances in Intelligent Systems and Computing*, 2020, № 1127, 307–314
- [8] Karpathy, Andrej, and Li Fei-Fei., “Deep visual-semantic alignments for generating image descriptions”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39:4** (2017), 664–676

Combined application of recurrent neural networks and statistical methods for improved oceanographic data forecasting accuracy

Kuzmin V.Y.

This paper is devoted to the application of statistical models to increase the prediction accuracy of oceanological data. The initial time series are modelled with mixtures of finite normal distributions. Statistical characteristics of constructed mixtures are used to pre-initialize the layers of a recurrent neural network. Forecasts made with the application of statistical models are compared with forecasts made using only the original data. It is demonstrated that a significant improvement in accuracy is observed for all analyzed series.

Keywords: EM, MSM, LSTM, neural networks, machine learning, feature enrichment, finite normal mixtures.

References

- [1] Cao, Jian, Zhi Li, and Jian Li., “Financial time series forecasting model based on CEEMDAN and LSTM”, *Physica A: Statistical Mechanics and its Applications*, 2019, № 519, 127-139
- [2] Yan, Binbin, and Memon Aasma., “A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM.”, *Expert systems with applications*, 2020, 159
- [3] Du, Xunsheng, et al., “Stacked LSTM deep learning model for traffic prediction in vehicle-to-vehicle communication”, *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017
- [4] Meneghini O., Luna C.J., Smith S.P., Lao L.L., “Modeling of transport phenomenon in tokamak plasmas with neural networks”, *Physics of Plasmas*, **21:6** (2014), 060702-0607024

- [5] Gorshenin A. K., Kuzmin V. Y., “Neural network forecasting of precipitation volumes using patterns”, *Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications.*, **28**:3 (2018), 450–461
- [6] Korolev V.Y., *Probabilistic and statistical methods for the decomposition of the volatility of chaotic processes*, Moscow University Press, Moscow, 2011 (In Russian), 512 pp.
- [7] Gorshenin A. K., Kuzmin V. Y., “A machine learning approach to the vector prediction of moments of finite normal mixtures”, *Advances in Intelligent Systems and Computing*, 2020, №1127, 307–314
- [8] Karpathy, Andrej, and Li Fei-Fei., “Deep visual-semantic alignments for generating image descriptions”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**:4 (2017), 664–676