End-to-End Timer Series Analysis

December 5, 2020

End-to-End Learning

E2E learning refers to training a possibly complex learning system represented by a single model (specifically a Deep Neural Network) that represents the complete target system, bypassing the intermediate layers usually present in traditional pipeline designs.

Deep Learning Models for E2E Time Series Analysis

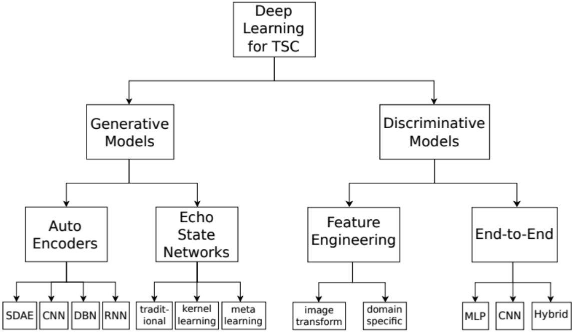

Figure: An overview of the different deep learning approaches for time series classification. Image by Fawaz et al. [3].

Deep Learning Models for E2E Time Series Analysis

- Multilayer Perceptrons
- CNN
- RNN & LSTM
- All kinds of combinations of the above ones

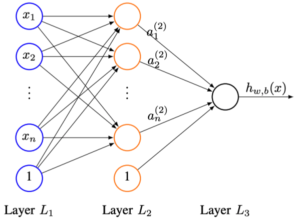

Figure: Schematic of a Neural Network Model. Image by Qiu et al. [7].

Recurrent Neural Network

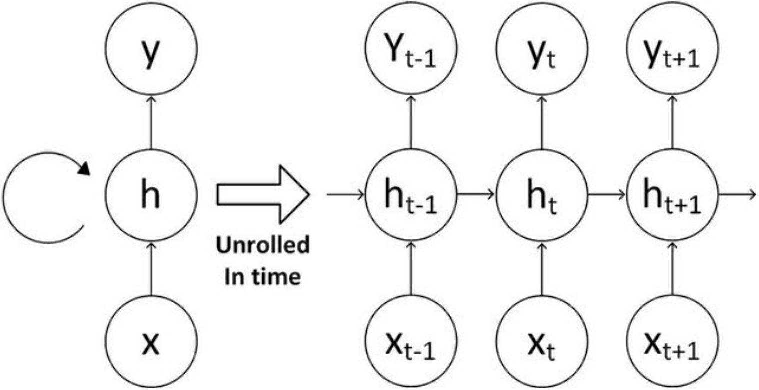

Figure: A conceptual visualization of the recurrent nature of an RNN. Image by Varsamopoulos et al. [8]

Long Short Term Memory

- Inheriting most of the characteristics of RNN models
- Solving gradient vanishing/explosion caused by the gradient BP process
- Suitable for handling problems that are highly related to time series

Long Short Term Memory

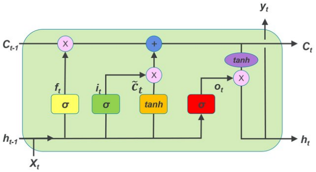

Figure: Schematic of a Neural Network Model. Image by Ismail et al. [4].

$$
\begin{aligned}
f^{(t)} &= \sigma\left(W_f h^{(t-1)} + U_f x^{(t)} + b_f\right)\\
i^{(t)} &= \sigma\left(W_i h^{(t-1)} + U_i x^{(t)} + b_i\right)\\
a^{(t)} &= \tanh\left(W_a h^{(t-1)} + U_a x^{(t)} + b_a\right)\\
C^{(t)} &= C^{(t-1)} \odot f^{(t)} + i^{(t)} \odot a^{(t)}\\
o^{(t)} &= \sigma\left(W_o h^{(t-1)} + U_o x^{(t)} + b_o\right)\\
h^{(t)} &= o^{(t)} \odot \tanh\left(C^{(t)}\right)\\
\hat{y}^{(t)} &= \sigma\left(V h^{(t)} + c\right)
\end{aligned}
$$

Time Series Analysis

- Time Series Forecasting: Stock predicting, inventory study
- Time Series Classification: Human activity recognition, acoustic scene classification
- Anomaly Detection: Holiday trips, Attack detection

Time Series Forecasting

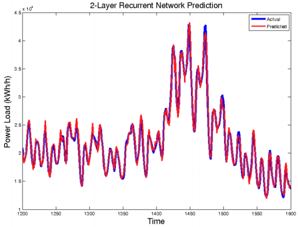

Figure: A prediction with recurrent neural network. Image by Busseti et al. [2].

We can use the time series data generated in the past to predict the future. In travel applications, historical data is used to predict the prices of future hotels and air tickets, so as to recommend the lowest purchase points for users (for example: prompting users to buy cheaper after five days). This small feature is enough to acquire a large number of loyal users.

Time Series Classification

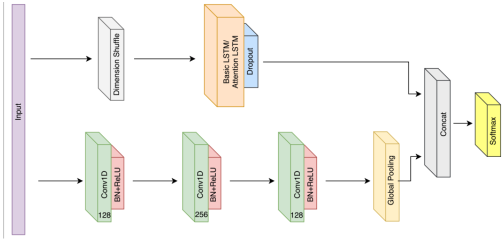

Figure: The LSTM-FCN architecture. LSTM cells can be replaced by Attention LSTM cells toconstruct the ALSTM-FCN architecture. Image by Karim et al. [5].

After revolutionizing the performances of several tasks in computer vision like object classifications, face verifications and etc., CNN has been introduced to this topic to address TSC problems because of its success in representation learning.

Anomaly Detection

Reliable temporal uncertainty estimation is a critical industrial demand. Before the wide application of deep learning, the community relied on traditional models, e.g. EGADS. Later, some end-to-end deep learning based model have been proposed.

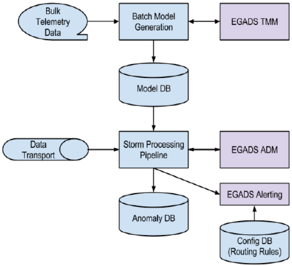

Figure: The architecture for EGADS-YMS. Image by Laptev et al. [6].

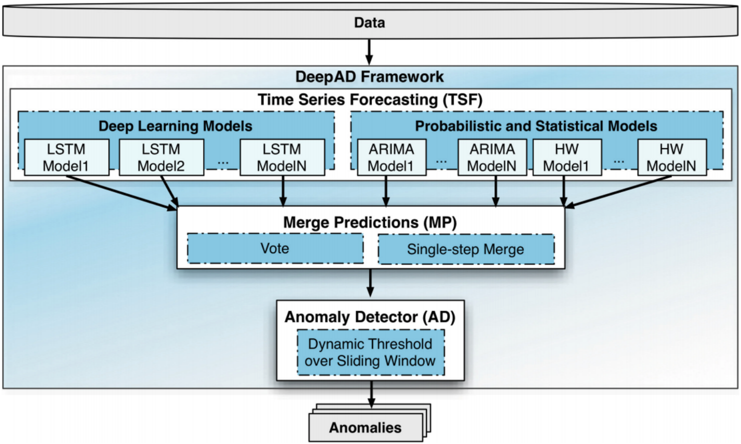

Figure: The framework overview of DeepAD. Image by Buda et al. [1].

References I

[1] Teodora Sandra Buda, Bora Caglayan, and Haytham Assem, *Deepad: A generic framework based on deep learning for time series anomaly detection*, Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer, 2018, pp. 577–588.

[2] Enzo Busseti, Ian Osband, and Scott Wong, *Deep learning for time series modeling*, Technical report, Stanford University (2012), 1–5.

[3] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller, *Deep learning for time series classification: a review*, Data Mining and Knowledge Discovery **33** (2019), no. 4, 917–963.

References II

[4] Aya Abdelsalam Ismail, Timothy Wood, and Héctor Corrada Bravo, *Improving long-horizon forecasts with expectation-biased lstm networks*, arXiv preprint arXiv:1804.06776 (2018).

[5] Fazle Karim, Somshubra Majumdar, Houshang Darabi, and Shun Chen, *Lstm fully convolutional networks for time series classification*, IEEE access **6** (2017), 1662–1669.

[6] Nikolay Laptev, Saeed Amizadeh, and Ian Flint, *Generic and scalable framework for automated time-series anomaly detection*, Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining, 2015, pp. 1939–1947.

References III

[7] Xueheng Qiu, Le Zhang, Ye Ren, Ponnuthurai N Suganthan, and Gehan Amaratunga, *Ensemble deep learning for regression and time series forecasting*, 2014 IEEE symposium on computational intelligence in ensemble learning (CIEL), IEEE, 2014, pp. 1–6.

[8] Savvas Varsamopoulos, Koen Bertels, and Carmen G Almudever, *Designing neural network based decoders for surface codes*, arXiv preprint arXiv:1811.12456 (2018).