Combining Global and Sequential Patterns for Multivariate Time Series Forecasting - Summary

Roni Ben Dom (207576463)
roni.bendom@post.runi.ac.il
Reichman University
Israel

ABSTRACT

Multivariate series forecasting is an important, widely used tool in many applications. According to "Combining Global and Sequential Patterns for Multivariate Time Series Forecasting" by Zhaoxi Li, Jun He, Hongyan Liu and Xiaoyong Du, published in 2020 IEEE International Conference on Big Data [3], existing methods has one of three shortcomings, either global trends across multivariate time series are ignored, times series and covariates cannot both be taken into consideration or they are lacking of interpretability. Aiming to resolve these issues, the authors propose a new model, named TEDGE, which demonstrates superior performance over existing models while maintaining interpretability of the results.

ACM Reference Format:
Roni Ben Dom (207576463). 2022. Combining Global and Sequential Patterns for Multivariate Time Series Forecasting - Summary. In *Big Data Platform course 2022: HW 1 - paper summary.* ACM, New York, NY, USA, 2 pages.

1 SUMMARY

Many real world applications are in-fact multivariate time series forecasting problems, such as stocks prices and traffic load forecasting. In these applications often exist global trends among different time-series within the same field who share the same timeline. Furthermore, covariates of each time series may hold important information about the time series which can help in forecasting. This leads us to the problem the authors tried to solve: forecasting of M time series given past values and covariates for each of the M time series, while taking into account global trends and maintaining interpretable and explainable results.
Aspiring to solve this problem the authors developed a new tensorized deep recurrent encoder-decoder framework with attention mechanism to capture long- and short-term temporal patterns in series with better interpretability, alongside autoregression based matrix factorization method to extract global trends across multivariate times series. The encoder in the suggested framework is tensorized LSTM (TLSTM). LSTM is a frequently used model in time series prediction even though it shows low interpretability of the output of the model due to blending all covariates information together. Trying to maintain explainability of the model, The authors chose to use TLSTM model, a LSTM based model which solves the interpretability problem as suggested in [1]. The TLSTM model is designed to predict only one step forward, therefore, the authors designed a decoder with dual attention in order to use both historical and future time covariates. The decoder shares the parameters of TLSTM in order to get the desired input for the dual attention which in turn provides variable and temporal importance metrics for each of the time series. This framework, named TED, can predict multiple future steps for each time series based on its covariates and historical time series values, still not taking into account global trends among different time series. To detect global trend the authors utilized a matrix factorization method with temporal regularization (TRMF) as suggested in [6]. The authors combined the output of TRMF model with other covariates, and fed it to the TED model as input. By combining these two models, the authors designed an explainable model which takes into account global trends and covariates when performaing multivarite time series forecasting, the TEDGE model.

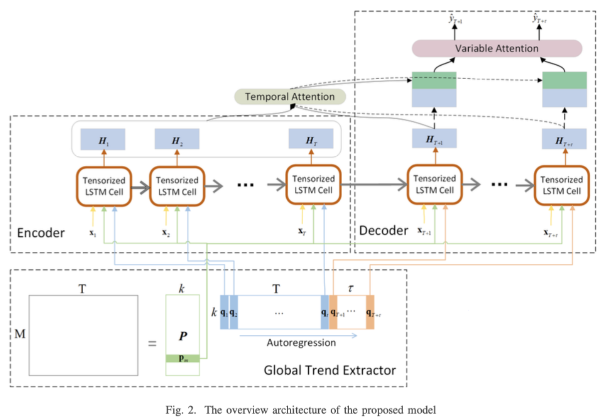

Figure 1: model architecture [3]

The model was evaluated against 8 different models with 3 different baseline methods. The first group of models TEDGE was tested against, referred by the authors as global models, do not take covariates into consideration. The second group, which include TED model, do take covariates into consideration but ignore global patterns, and the last group, referred as mixed models, take into consideration both global trend and local properties of the time series. The models were tested on 3 datasets from different fields with different properties: multivariate forecasting with covariates (Live600 dataset, provided by a live streaming platform), multivariate forecasting without covariates (Electricity dataset, also used in [2]) and single variable forecasting with covariates (PM2.5 dataset, a public dataset provided by [4]). The model's performance on these datasets were measured by two metrics: Root Relative Squared Error (RRSE) and Normalized Mean Absolute Error (NMAE), scaled versions of RMSE and MAE, since comparing varying value ranges with RMSE and MAE can distort the results.

TABLE III
PERFORMANCE COMPARISON WITH BASELINES ON THREE DATASETS.

Method		Live600		PM2.5		Electricity	
		NMAE	RRSE	NMAE	RRSE	NMAE	RRSE
Global models	TRMF	0.4538	2.5910	–	–	0.0884	0.0152
	TCN-MF	0.4520	3.0365	–	–	0.2785	0.0755
	LSTNet	0.3834	1.8422	–	–	0.0997	0.0166
	DSANet	0.4111	1.9615	–	–	0.1167	0.0175
Local models	ARIMAX	0.4853	3.8637	0.8337	7.8195	0.5102	2.5531
	DSSM	0.2994	0.5515	**0.1175**	**0.1919**	**0.0508**	0.0162
	IMV-LSTM	0.4164	1.6058	0.4836	2.2092	0.0951	0.0182
	TED	0.3077	0.6654	**0.1177**	**0.1939**	0.0623	0.0085
Mixed models	DeepGLO	0.4111	1.9615	0.3470	1.7998	0.0998	0.0412
	TEDGE	**0.2324**	**0.4105**	–	–	0.0570	**0.0083**

Figure 2: model performance [3]

All and all, TEDGE demonstrated best performance on Live600 dataset, it performed similarly to DSSM model [5] on the Electricity dataset, as well as similar results for TED model and DSSM PM2.5 dataset. Although showing similar results to DSSM model, the strength of this model is the explainability using both variable and temporal importance metrics. Global models and TEDGE model were not suitable for single time series forecasting, therefore the lack of results in the table.

Furthermore, The effect of the attention component in the model was tested, meaning TEDGE model was tested on Live600 dataset against itself with some changes - once omitting the attention mechanism related to temporal importance and once the one related to variable importance. Third version tested against the full TEDGE model was without attention at all. All 3 versions achieved lesser performance than the full model.

Finally, the model proposed by the authors achieves their goal of designing an accurate and interpretable multivariate time series forecasting model which takes into account both global trends and covarites, while showing superior performance over existing models.

REFERENCES

[1] Tian Guo, Tao Lin, and Nino Antulov-Fantulin. 2019. Exploring Interpretable LSTM Neural Networks over Multi-Variable Data. https://doi.org/10.48550/ARXIV.1905.12034

[2] Guokun Lai, Wei-Cheng Chang, Yiming Yang, and Hanxiao Liu. 2017. Modeling Long- and Short-Term Temporal Patterns with Deep Neural Networks. *CoRR* abs/1703.07015 (2017). arXiv:1703.07015 http://arxiv.org/abs/1703.07015

[3] Zhaoxi Li, Jun He, Hongyan Liu, and Xiaoyong Du. 2020. Combining Global and Sequential Patterns for Multivariate Time Series Forecasting. In *2020 IEEE International Conference on Big Data (Big Data)*. 180–187. https://doi.org/10.1109/BigData50022.2020.9378380

[4] Xuan Liang, Tao Zou, Bin Guo, Shuo Li, Haozhe Zhang, Shuyi Zhang, Hui Huang, and Song Xi Chen. 2015. Assessing Beijing's PM2. 5 pollution: severity, weather impact, APEC and winter heating. *Proc. R. Soc. A* 471, 2182 (2015), 20150257.

[5] Syama Sundar Rangapuram, Matthias W Seeger, Jan Gasthaus, Lorenzo Stella, Yuyang Wang, and Tim Januschowski. 2018. Deep State Space Models for Time Series Forecasting. In *Advances in Neural Information Processing Systems*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Eds.), Vol. 31. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2018/file/5cf68969fb67aa6082363a6d4e6468e2-Paper.pdf

[6] Hsiang-Fu Yu, Nikhil Rao, and Inderjit S Dhillon. 2016. Temporal Regularized Matrix Factorization for High-dimensional Time Series Prediction. In *Advances in Neural Information Processing Systems*, D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (Eds.), Vol. 29. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2016/file/85422afb467e9456013a2a51d4dff702-Paper.pdf