KDD CUP 2022

Application of BERT for Wind Power Forecasting

Teletraan

Longxing Tan, Hongying Yue

Introduction

Task

- A precise forecasting of wind power is beneficial to the sustainability and security of the power system
- Forecast wind power in 48 hours of every 10 minutes
- The performance is evaluated by average score of MAE and RMSE

Dataset

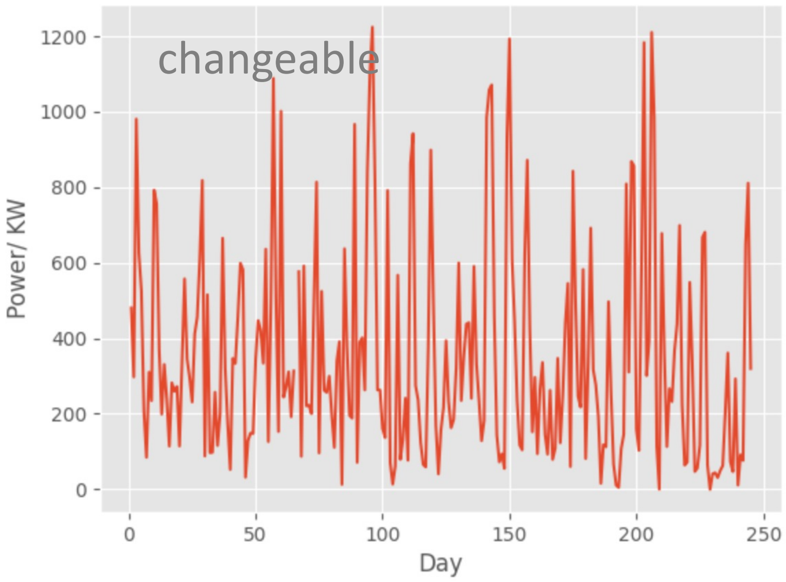
- 134 turbines of 245-day history data are provided
- In each timestamp, there are wind speed, direction, temperature, and so on, totally 10 parameters
- Spatial measurements of each turbine

Challenge

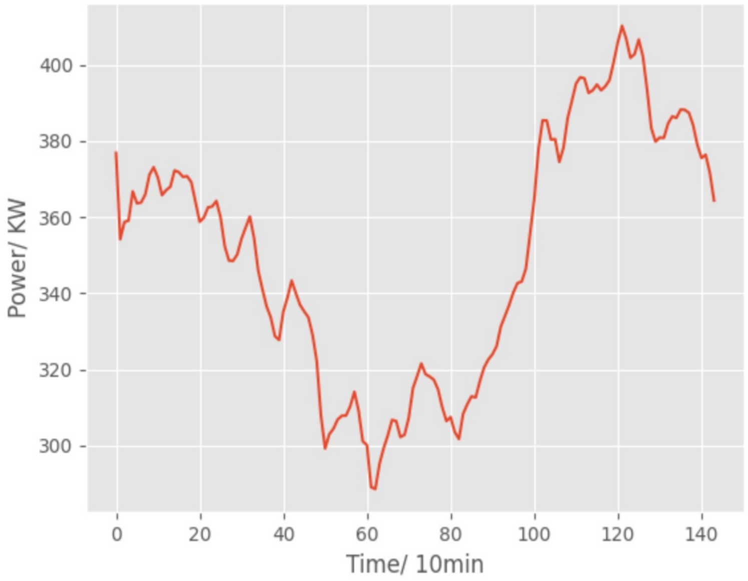
- Unpredictable nature, especially without meteorological forecast
- Long sequence prediction

A glimpse of data

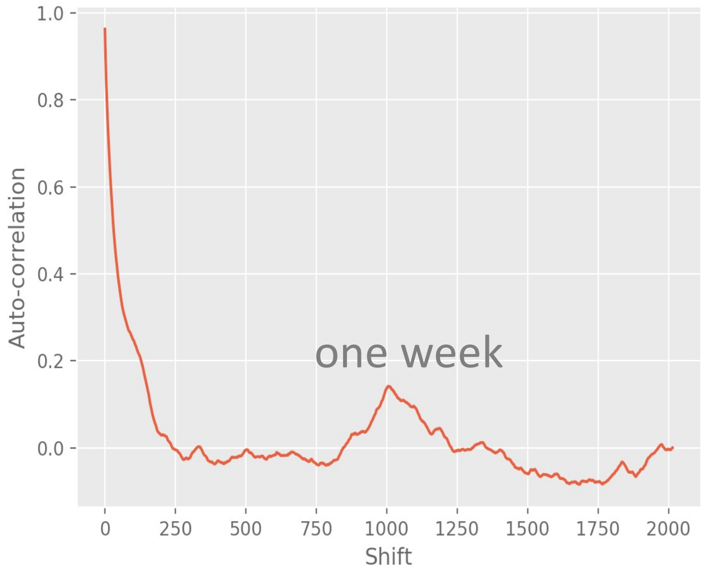
Overall trend



Daily period



Auto-correlation

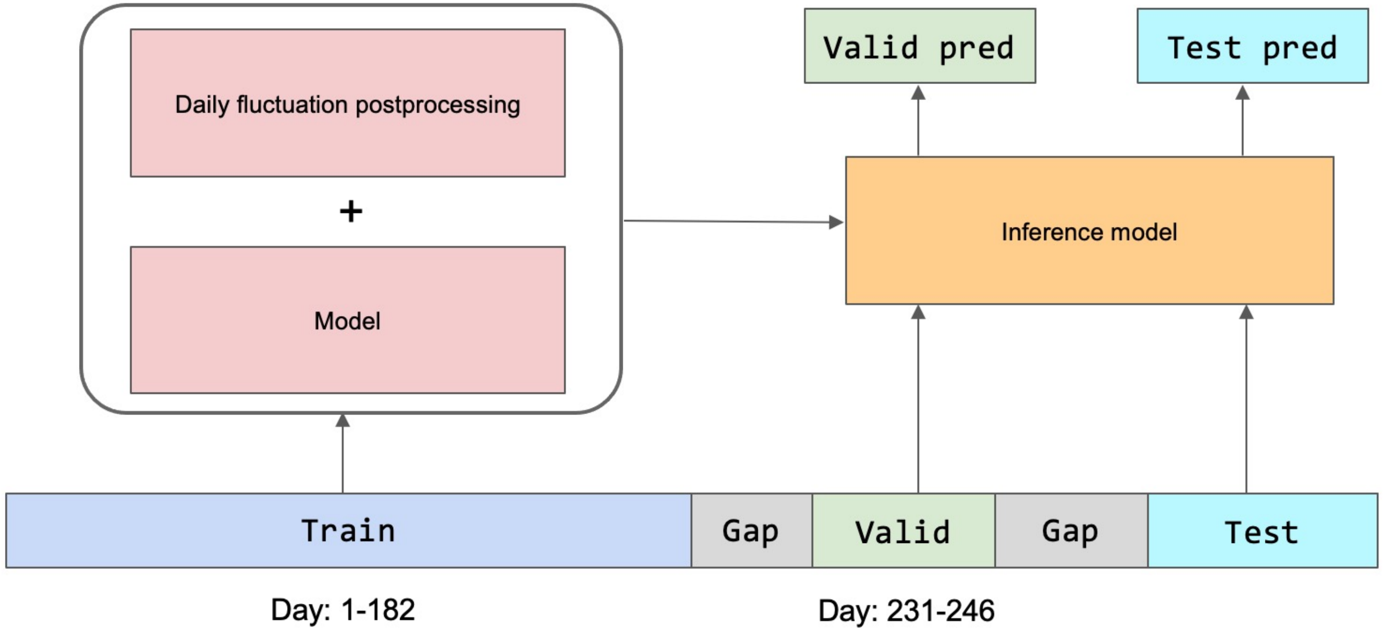


Solution architecture

Local validation strategy
- Split the data by time
- Reserve gap data to simulate the online scenario

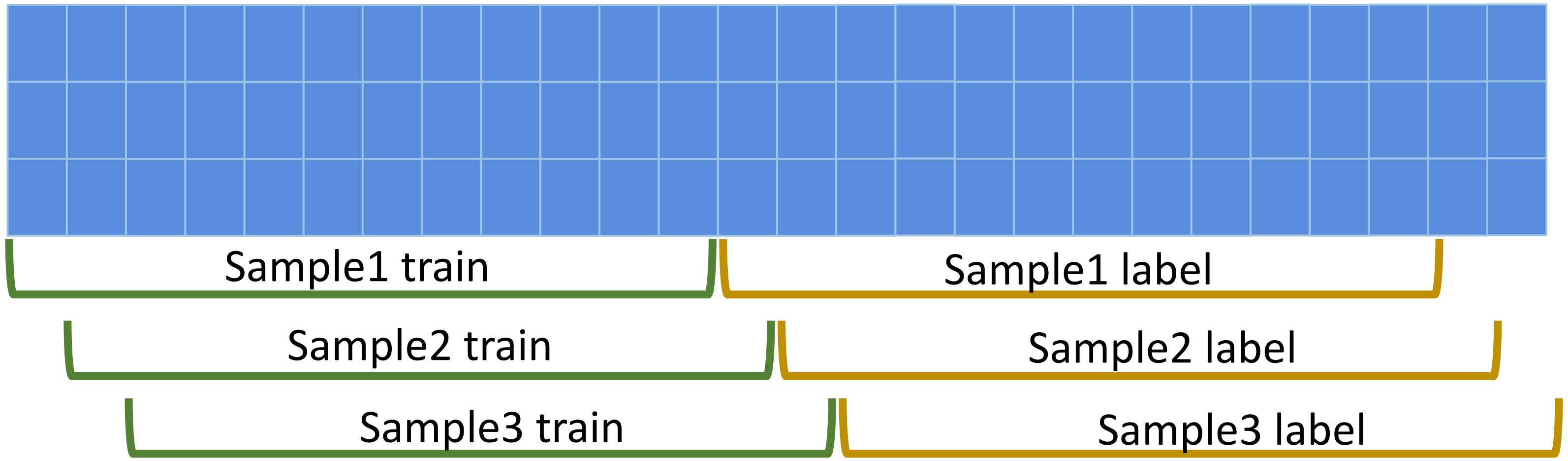
Model strategy
- Time series prediction model
- Spatial temporal prediction model

Multi-step strategy
- Multi-model
- Multi-ouptput
- Dynamic decoding
- One step decoding
- Forecasting decoding



Pre-process and features

Sample



Pre-process

- NAN values
  - Fill in with previous-period values
- Scaler
  - Min-Max scaler

Feature

- Wind speed
- Wind direction

Model selection

The most promising data-driven models are statistical models and deep leaning models. We can treat the task as time series prediction or spatial temporal prediction to apply related models. And we choose a single BERT as our final model.

Model	Local MAE	Local RMSE	Local score	Leaderboard score (phase II)
BERT	299	365	58.1	44.6
LSTM	305	369	58.9	44.8
TCN	310	371	59.4	45.1
KNN	316	368	60.6	unsubmit
LGB	319	386	61.4	unsubmit
Transformer	311	374	60.0	48.5
Seq2seq	296	366	57.6	47.1
Wavenet	306	370	59.1	47.9
GCN-LSTM	228	362	54.5	48.2

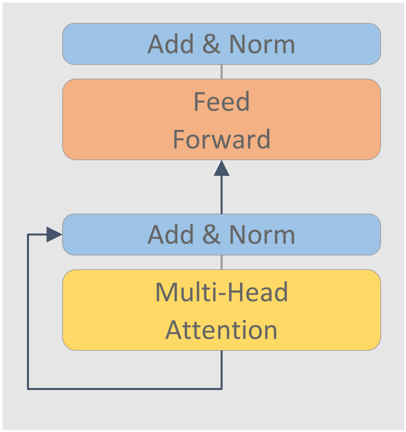
Model structure of BERT

Advantage of BERT model

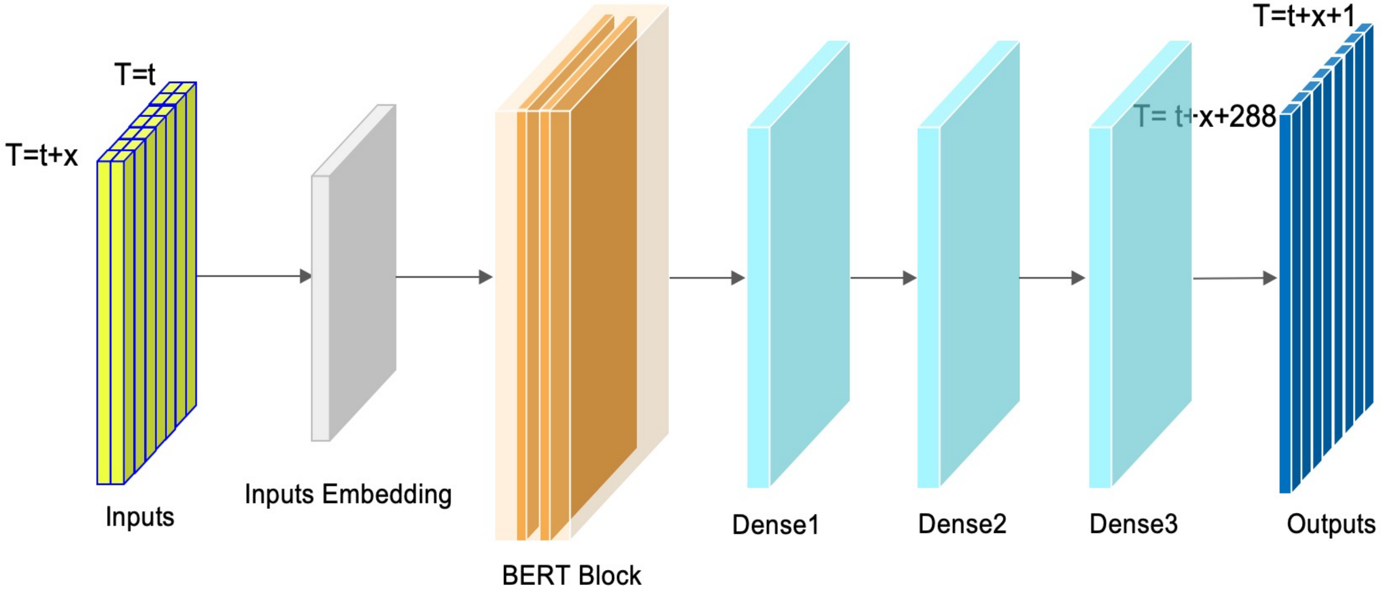
- Use attention to capture the long dependencies
- Use Dense layer as multi-output, without decoding to accelerate the training and inference

Special

- Without pre-training stage
- Without positional encoding
- A single BERT layer

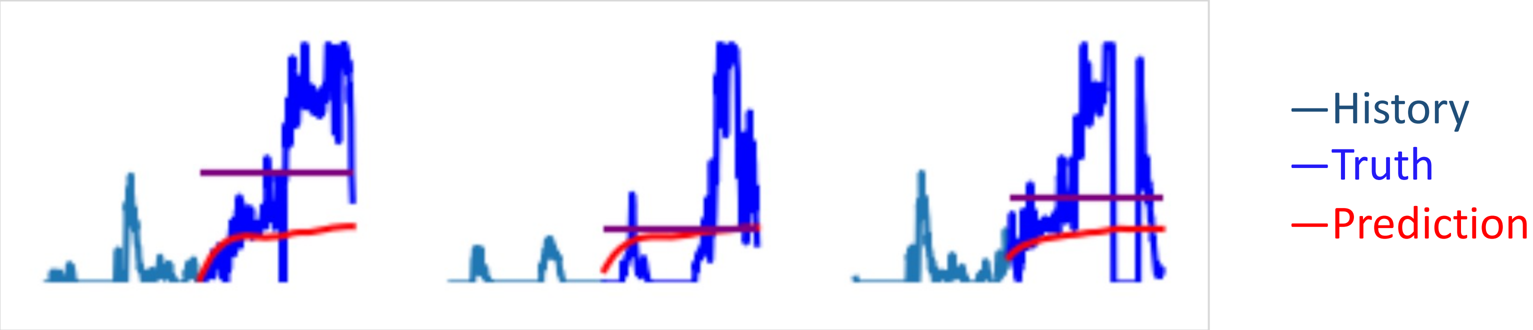


BERT Block



Post-process to add daily period

As mentioned, there is strong daily periodicity in history, but our prediction model could not reflect it



Postprocess

- Calculate the daily fluctuation for each timestamp
- Normalize the fluctuation into 0-1
- Magnify it into 0-36
- Shift it to match with the prediction's start time

Model optimization

- It's not helpful to add time information or time related statistical information as decoder feature

Conclusions

- It's both **simple and effective** for this wind power forecasting task to apply a **BERT** model, predicting the major trend.

- The **daily fluctuation** is added in post-processing to make the prediction effective with daily periodicity.

Full code is released at https://github.com/LongxingTan/KDDCup2022-Baidu

TELETRAAN
浙江天垂科技有限公司

Reference

[1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018).

[2] Haixu Wu, Jiehui Xu, Jianmin Wang, and Mingsheng Long. 2021. Autoformer: De-composition transformers with auto-correlation for long-term series forecasting. Advances in Neural Information Processing Systems 34 (2021), 22419–22430.

[3] JingboZhou,ShuangliLi,LiangHuang,HaoyiXiong,FanWang,TongXu,Hui Xiong, and Dejing Dou. 2020. Distance-aware molecule graph attention network for drug-target binding affinity prediction. arXiv preprint arXiv:2012.09624 (2020).

[4] HaoyiZhou,ShanghangZhang,JieqiPeng,ShuaiZhang,JianxinLi,HuiXiong, and Wancai Zhang. 2021. Informer: Beyond efficient transformer for long sequence time-series forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35. 11106–11115.

[5] Jingbo Zhou, Xinjiang Lu, Yixiong Xiao, Jiantao Su, Junfu Lyu, Yanjun Ma, and De- jing Dou. 2022. SDWPF: A Dataset for Spatial Dynamic Wind Power Forecasting Challenge at KDD Cup 2022. Techincal Report (2022).

[6] QingsongWen,TianZhou,ChaoliZhang,WeiqiChen,ZiqingMa,JunchiYan, and Liang Sun. 2022. Transformers in time series: A survey. arXiv preprint arXiv:2202.07125 (2022).

Thank You

Teletraan, Hangzhou, China