Cluj Napoca Temperature Prediction

Stăncioiu Laurențiu

Intelligent Modelling

What did we try to achieve?

Data colection

- Hourly data from January 1st 2008 up until May 1st 2023.
- 13 years of training data, 2 years of validation data.

	January	February	March	April	May	June	July	August	September	October	November	December
2008	-1.92	1.5	5.44	10.01	14.67	18.79	18.93	19.75	13.48	9.83	3.71	1.09
2009	-2.07	-0.82	3.56	12.51	15.39	18.13	20.46	20.12	16.62	9.5	5.92	-0.13
2010	-2.72	1.01	4.28	10.14	14.96	18.61	20.36	20.33	13.81	6.62	6.7	-2.09
2011	-3.31	-2.7	4.72	10.47	14.71	18.63	19.62	20.14	17.28	7.63	0.02	1.07
2012	-2.99	-7.15	4.07	11.21	15.42	19.89	23.2	21.42	17.59	10.31	4.58	-2.42
2013	-2.13	2.0	3.01	11.43	16.16	18.96	20.33	21.38	13.49	9.73	6.73	-2.06
2014	0.34	3.12	8.26	11.51	15.1	18.22	20.44	19.74	16.12	10.51	4.62	1.34
2015	-0.55	0.26	5.46	9.18	15.54	19.13	21.81	21.92	17.19	9.46	5.74	1.51
2016	-2.51	5.14	6.09	12.48	14.22	19.96	20.56	19.38	16.55	8.52	2.84	-2.77
2017	-6.52	1.73	8.32	9.46	15.56	20.24	20.8	22.19	15.55	9.84	4.96	1.49
2018	0.24	-0.17	3.32	14.94	18.16	19.35	20.28	21.98	16.3	12.03	5.37	-0.25
2019	-1.77	1.72	6.92	11.6	14.09	21.36	20.35	21.97	16.57	10.6	8.8	0.69
2020	-2.66	2.31	5.96	9.76	13.35	19.09	19.86	21.24	17.3	11.21	3.36	3.25
2021	-0.39	1.46	3.15	7.88	13.8	19.38	21.99	19.04	14.15	7.72	3.94	1.26
2022	-1.44	1.66	3.2	8.44	15.34	20.15	21.88	21.67	14.07	10.3	5.61	1.93
2023	3.08	0.53	5.64	8.05	12.75	-	-	-	-	-	-	-

Fig. 1. Average temperatures by month from January 2008 to May 2023

Data Analysis

- 29 features from which we selected 15
- Temperature set as the target
- After doing a Correlation Matrix dropped all Temperature related features.

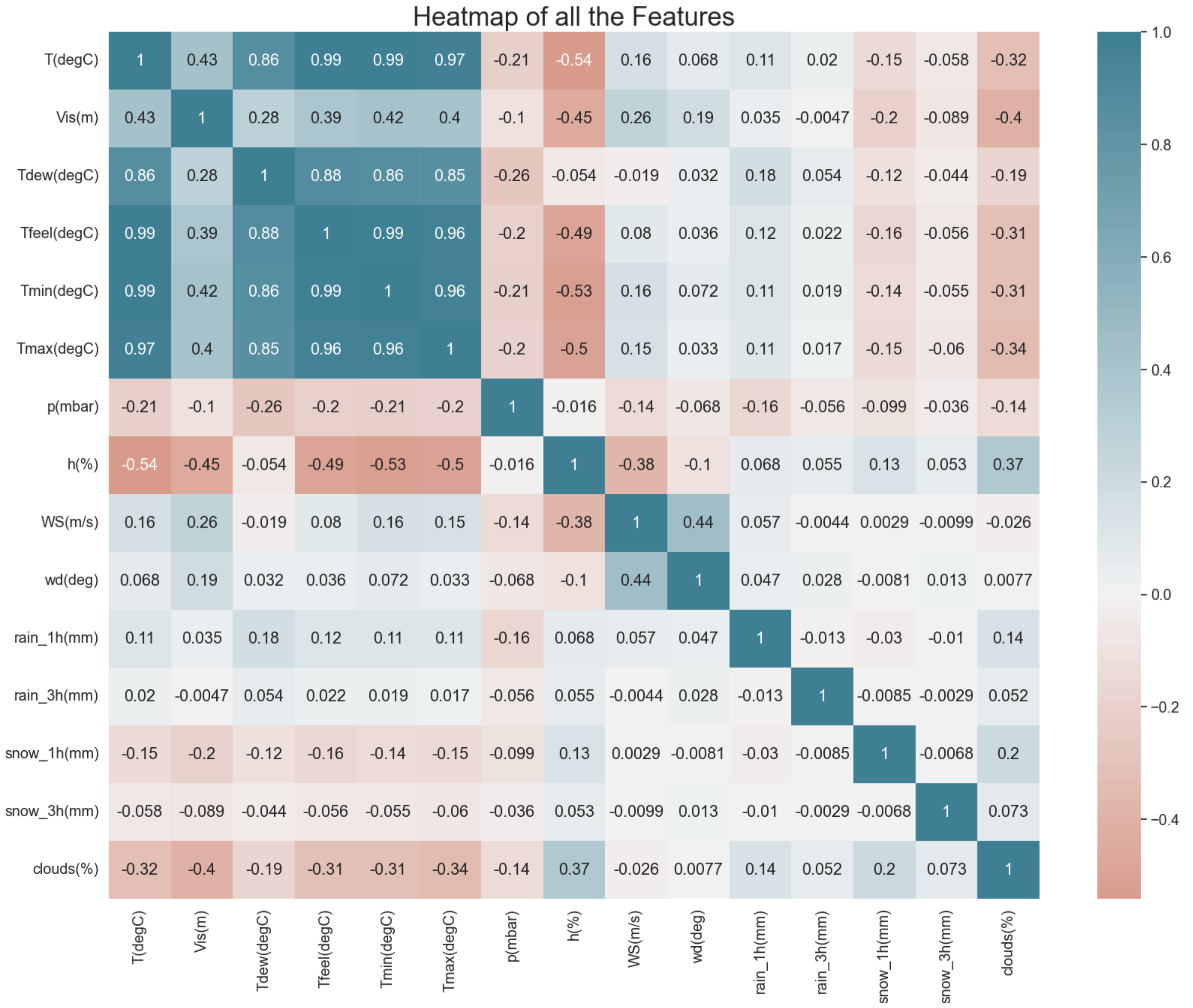

Fig. 2. Correlation matrix of selected features

Model Selection

Selected the transformer architecture Autoformer and a Long-Short Term Memory (LSTM) as the comparison models.

Wanted to see which model has a better short-term temperature prediction.

After deciding the hyperparameters of the model, trained and made comparisons for a 24 hours prediction length and a 12 hours prediction length.

The metrics used were Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE).

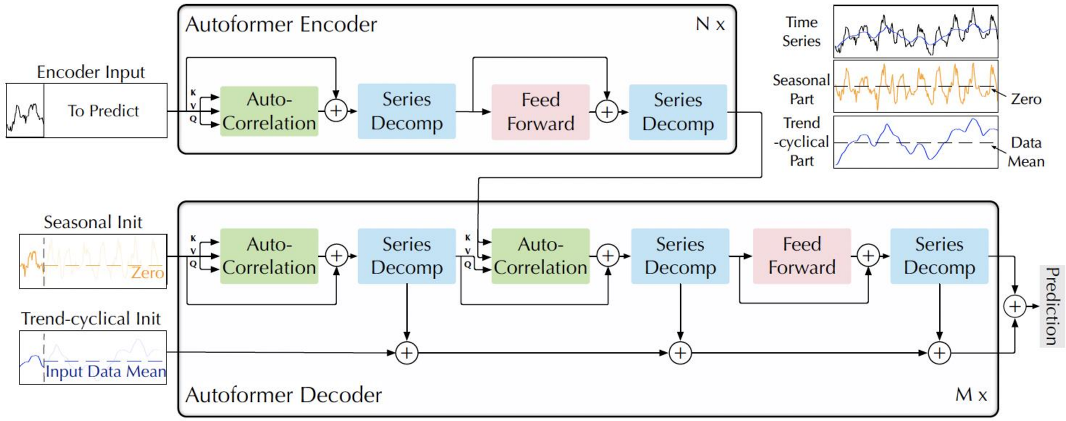

Fig. 3. Autoformer architecture

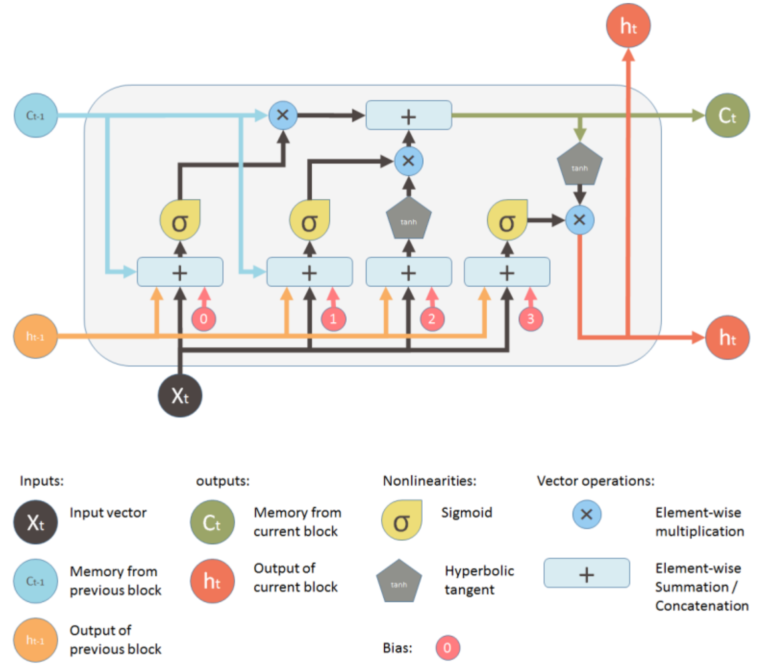

Fig. 4. Long-Short Term Memory block architecture

LSTM results

- Here are presented the results for a prediction of 12 hours taking into consideration 120 hours of past data for training.
- RMSE = 2.39, MAE = 1.85

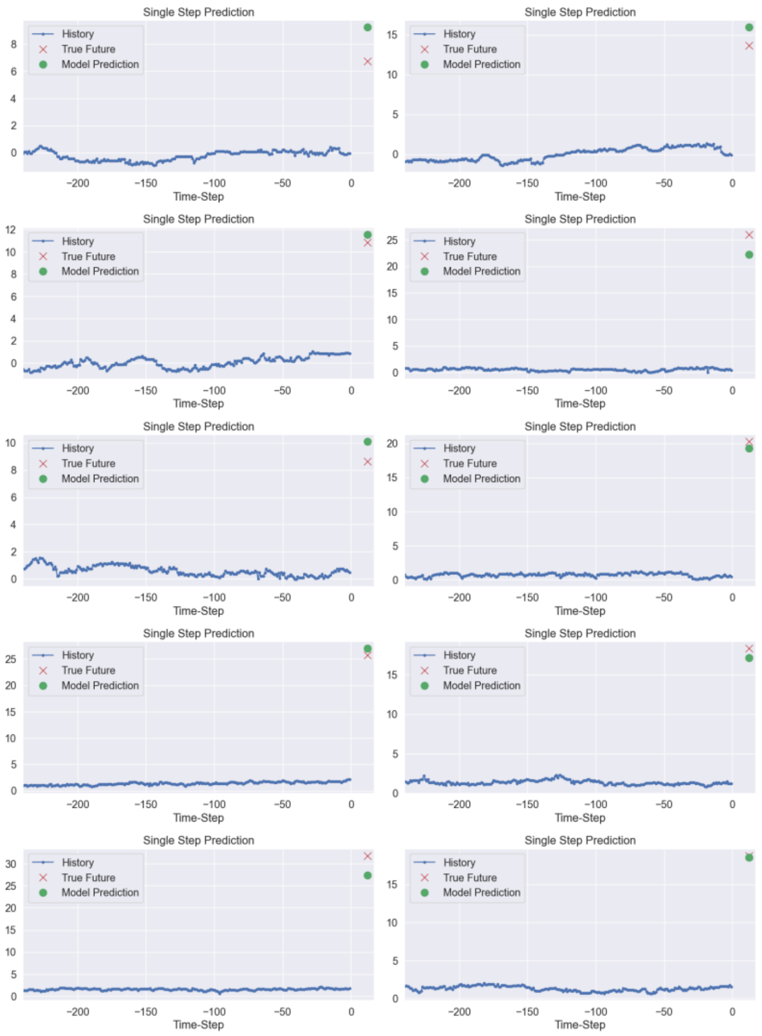

Fig. 5. Single Step Prediction for LSTM 120 hours of training 12 hours of prediction

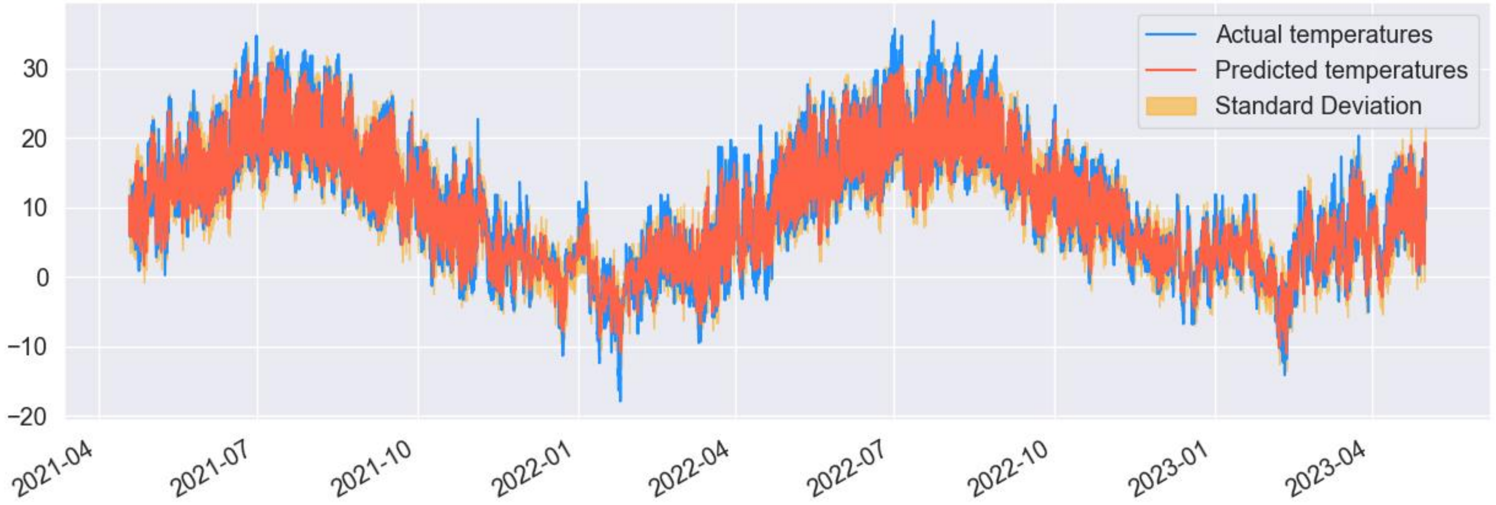

Fig. 6. Actual vs. Predicted temperatures for test data from april 2021 until May 2023

LSTM results

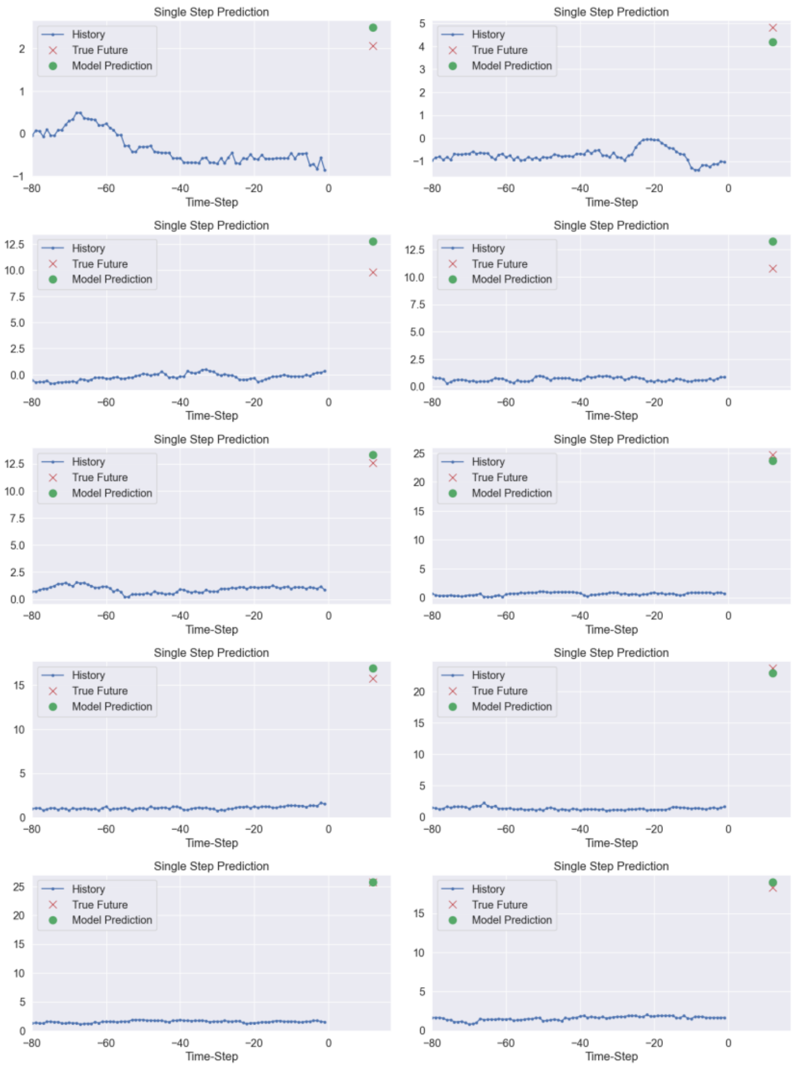

Fig. 7. Single Step Prediction for LSTM 96 hours of training 24 hours of prediction

- Here are presented the results for a prediction of 24 hours taking into consideration 96 hours of past data for training.
- RMSE = 3.01, MAE = 2.39

Fig. 8. Actual vs. Predicted temperatures for test data from april 2021 until May 2023

Autoformer

Here are presented the results for a prediction of 12 hours taking into consideration 120 hours of past data for training.

RMSE = 2.75, MAE = 2.08

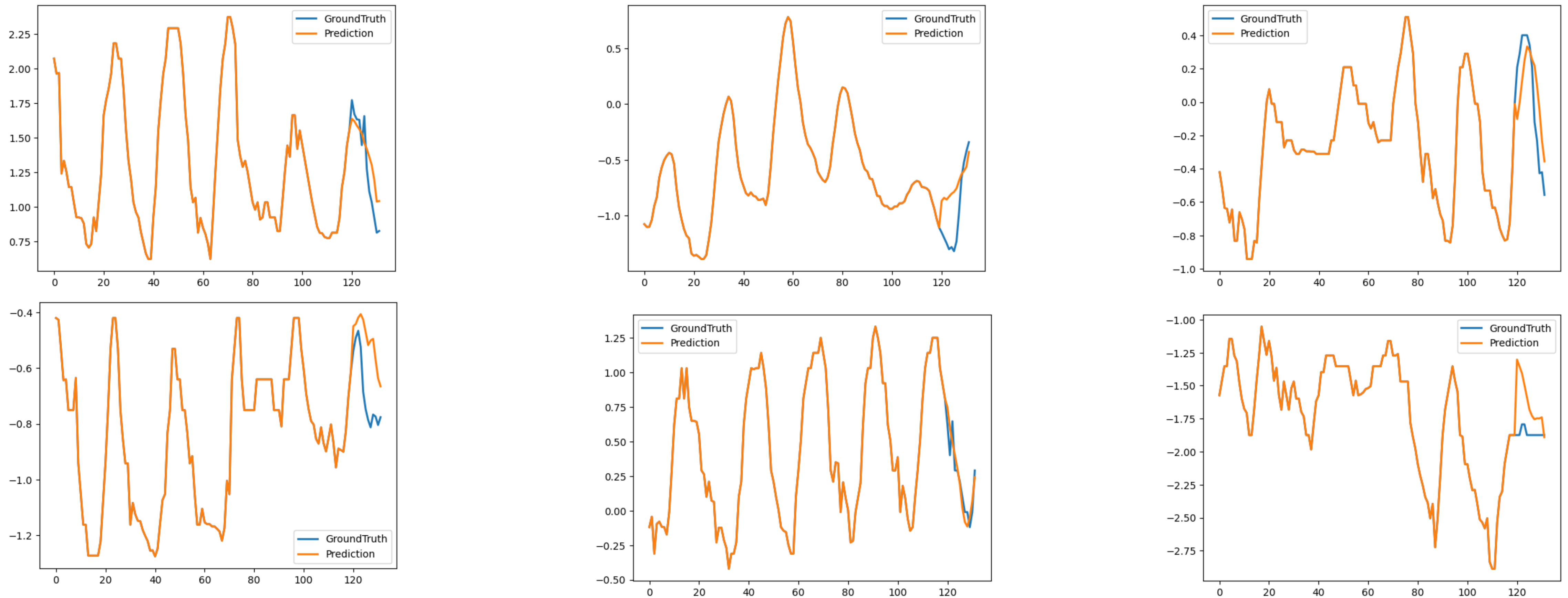

Fig. 9. Multi Step 12 hour Prediction of Autoformer for 120 hours of training.

Autoformer

Here are presented the results for a prediction of 24 hours taking into consideration 96 hours of past data for training.
RMSE = 3.25, MAE = 2.23

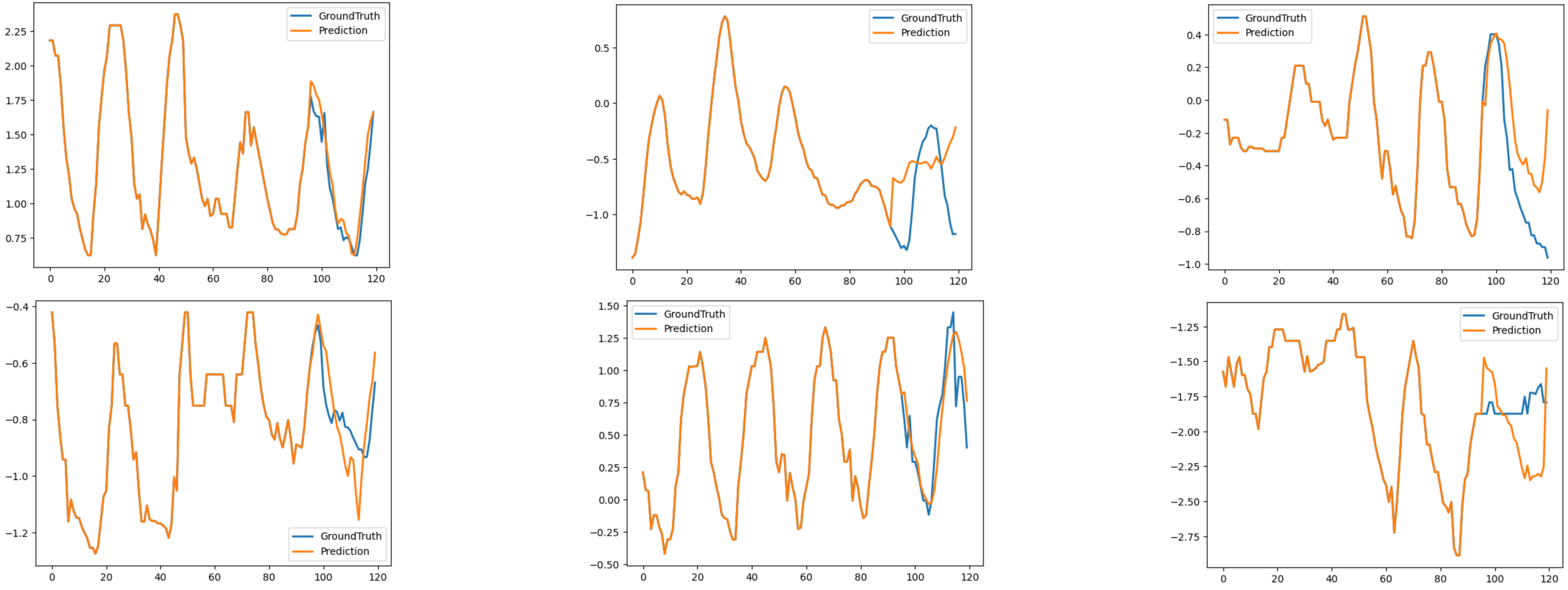

Fig. 10. Multi Step 24 hours Prediction for 96 hours of training data.

SWOT Analysis

Strengths

- We used two of the best models available LSTM and Autoformer.
- We had thorough data analysis before feeding the data into the models.
- We used metrics which are explainable for the feature we wanted to predict.
- We obtained results which can be usable in the real world.
- We verified the *No Free Lunch* theorem

Weaknesses

- Only the LSTM model was given a proper hyperparameter tuning. The Autoformer lasted too long in order to try different approaches.
- We did not try simpler models, which might be better at predicting short-term dependencies (ARIMA).
- We only did the predictions on temperatures, instead of having a multi-variate model.

Opportunities

- We could decrease training times by using cloud GPU.
- More recent transformer architectures which are said to be better at predictions.
- Training the current model more and tweaking some hyperparameters might lead to better results.

Threats

- Both models didn't do well at predicting extreme temperatures. This could present a lot of not so good consequences if we deploy the model.
- High level of abstraction when it comes to explaining a prediction
- Computationally expensive especially for the transformer model.

Bibliography

- "Autoformer: Decomposition Transformers with Auto-Correlation for Long-Term Series Forecasting" (NeurIPS 2021), https://arxiv.org/abs/2106.13008
- Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. Neural computation, 9(8):1735–1780, 1997.
- https://github.com/LaurentiuStancioiu/Weather_project/