Project Proposal - Graph Neural Networks for Global Weather Prediction

1 Literature, Problem Statement and Proposed Solution

Despite the success of deep learning methods in problems spanning across various domains, rigorous studies of their application in weather prediction are scarce. Presently, purely physical, numerical models are used for state-of-the-art weather prediction. Since current numerical weather prediction (NWP) requires building large ensembles of models solving governing physical equations on a global grid, their use is limited not only due to the large computational cost, but occasionally also systematically, as parametrizations in physical models may not capture the complexity of specific climatological phenomena.[1, 2, 3] Data-driven, deep learning methods might inherently address the aforementioned shortcomings due to their proven ability to capture high-dimensional dependencies in such complex, non-linear systems at reasonable computational cost. [4]

In current weather prediction pipelines, deep learning methods are commonly used for post-processing the numerical simulation ensembles.[5, 6] In recent years however, several studies attempt to show their feasibility for full medium-range predictions, i.e. predicting specific climatological properties globally with a lead time of a few hours based only on past reanalysis data gathered from satellite observations using e.g. encoder-decoder CNNs or pre-trained ResNets. [2, 4, 7, 8] To improve comparability between such studies, the WeatherBench dataset was recently published, alongside baselines from purely data-driven as well as numerical approaches. [9] WeatherBench will act as the main point of reference to assess the feasibility our proposed methods.

The objective of this project is the development of a GNN-based weather prediction model. Weather prediction is a timeseries forecasting problem, which requires simultaneously finding a suitable representation of the spatial dependencies between measurement points, as well as their temporal dynamics. Because the spatial dependencies are non-uniform caused by uneven spacing from mapping a sphere to a flat image and by local features such as terrain, we will attempt to use graph-based representations to model such dependencies. The focus lies on experimenting with different graph representations and GNN architectures. The final model is then evaluated on the WeatherBench dataset and compared to the given baselines.

2 Methods and Experiments

The WeatherBench data in it's coarsest resolution of 32×64 sample points contains 300 GB of data spanned over 40 years and 115 features. A selection of features will be used to predict the geopotential and temperature like in the baseline models.

Since we are leveraging the spatial dependencies that are inherently natural to a graph (the one-hop neighbourhood of a vertex), Graph Attention Networks [10] can be used as a substitute for Graph Convolutional Networks [11] in a GNN-based prediction approach. GATs improve the learning capacity of the model using the attention mechanism, which assigns different importance to each neighbour's contribution, compared to the GCN where the normalized sum is computed, therefore potentially yielding better overall results.

To model the timeseries forecasting we will try out multiple approaches such as WaveNet [12], which utilizes time dilated convolutions to widen it's receptive field. Alternatively, spatio-temporal features can also be extracted simultaneously from the graph-structure, such as in STGCNs.[13]

3 Computing Resources

In order to perform hyperparameter tuning over the architectures and pipelines we study, we require either multiple-CPU computation nodes with sufficient RAM, or a CPU-computation node with a GPU, having both enough RAM and VRAM. Moreover, in order to store the dataset, we also require a physical storage with at least 500 GB of memory.

Author Information

Blaga Cristian-Ioan (cblaga@ethz.ch)
Chan Robin (chanr@ethz.ch)
Gavrilas Gabriel (ggabriel@ethz.ch)
Vogelsanger Christopher (cvogelsa@ethz.ch)

References

[1] E. Kalnay. *Athmospheric modeling, data assimilation, and predictability*, volume 54. Cambridge University Press, 2003.

[2] Jonathan A. Weyn, Dale Richard Durran, and Rich Caruana. Can machines learn to predict weather? using deep learning to predict gridded 500-hpa geopotential height from historical weather data. *Journal of Advances in Modeling Earth Systems*, 2019.

[3] Peter Vogel, Peter Knippertz, Andreas H. Fink, Andreas Schlueter, and Tilmann Gneiting. Skill of global raw and postprocessed ensemble predictions of rainfall over northern tropical africa. *Weather and Forecasting*, 33 (2):369 – 388, 2018. doi: 10.1175/WAF-D-17-0127.1. URL `https://journals.ametsoc.org/view/journals/wefo/33/2/waf-d-17-0127_1.xml`.

[4] P. D. Dueben and P. Bauer. Challenges and design choices for global weather and climate models based on machine learning. *Geoscientific Model Development*, 11(10):3999–4009, 2018. doi: 10.5194/gmd-11-3999-2018. URL `https://gmd.copernicus.org/articles/11/3999/2018/`.

[5] Stephan Rasp and Sebastian Lerch. Neural networks for postprocessing ensemble weather forecasts. *Monthly Weather Review*, 146(11):3885–3900, Oct 2018. ISSN 1520-0493. doi: 10.1175/mwr-d-18-0187.1. URL `http://dx.doi.org/10.1175/MWR-D-18-0187.1`.

[6] Peter Grönquist, Chengyuan Yao, Tal Ben-Nun, Nikoli Dryden, Peter Dueben, Shigang Li, and Torsten Hoefler. Deep learning for post-processing ensemble weather forecasts. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194):20200092, Feb 2021. ISSN 1471-2962. doi: 10.1098/rsta.2020.0092. URL `http://dx.doi.org/10.1098/rsta.2020.0092`.

[7] Sebastian Scher. Toward data-driven weather and climate forecasting: Approximating a simple general circulation model with deep learning. *Geophysical Research Letters*, 45, 11 2018. doi: 10.1029/2018GL080704.

[8] Sebastian Scher and Gabriele Messori. Predicting weather forecast uncertainty with machine learning. *Quarterly Journal of the Royal Meteorological Society*, 144, 10 2018. doi: 10.1002/qj.3410.

[9] Stephan Rasp, Peter D. Dueben, Sebastian Scher, Jonathan A. Weyn, Soukayna Mouatadid, and Nils Thuerey. Weatherbench: A benchmark data set for data-driven weather forecasting. *Journal of Advances in Modeling Earth Systems*, 12(11), Nov 2020. ISSN 1942-2466. doi: 10.1029/2020ms002203. URL `http://dx.doi.org/10.1029/2020MS002203`.

[10] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph Attention Networks. *International Conference on Learning Representations*, 2018. URL `https://openreview.net/forum?id=rJXMpikCZ`.

[11] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *CoRR*, abs/1609.02907, 2016. URL `http://arxiv.org/abs/1609.02907`.

[12] Aäron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew W. Senior, and Koray Kavukcuoglu. Wavenet: A generative model for raw audio. *CoRR*, abs/1609.03499, 2016. URL `http://arxiv.org/abs/1609.03499`.

[13] Bing Yu, Haoteng Yin, and Zhanxing Zhu. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, Jul 2018. doi: 10.24963/ijcai.2018/505. URL `http://dx.doi.org/10.24963/ijcai.2018/505`.