HAIL RISK PREDICTION VIA GRAPH NEURAL NETWORKS

A PREPRINT

Ivan Lukyanenko
Department of Control and Applied Mathematics
Moscow Institute of Physics and Technologies
Moscow
lukianenko.ia@phystech.edu

Yuriy Maximov
Research Center in Artificial Intelligence
Skoltech

ABSTRACT

Geo-spatial time series prediction is an open area with great potential for theoretical and practical work. In particular, hail risk assessment is necessary to predict the probability of damage (agriculture, animal husbandry). The aim of our study is to build a hail forecasting model based on graph neural networks. Forecasting has been carried out in the short-term range based on the values of climate variables since 1991. The key features of the problem are: 1) rare events - over the past 30 years there have been less than 700 hail events throughout Russia, 2) the spatial structure of the series data. We are expecting to improve quality of solving such problems by combining methods from [1] and [2].

Keywords Hail risk prediction · GNN · spatial time series

1 Introduction

The main goal of our research is to introduce new GNN architecture by combining state-of-art results in the right way. Our object of research is geo-spatial time series. In common time series is a series of values of a quantity obtained at successive times, often with equal intervals between them. But we want to pay our attention to more extended concept – spatial time series. Spatial time series are almost the same as ordinary time series, but instead of values we observe some spatial objects(i.e photos, heat maps and etc.). In particular our research is aimed to predict extreme events. We are going to provide a solution to the hail forecasting problem. Hails are extreme events, because over the past 30 years there have been less than 700 hail events throughout Russia. The assessment of hail risk is necessary because of its environment damage. According to Verisk's 2021 report, in 2020 year losses due to hails reached $14.2 billion in the USA. In the research we set the goal to improve state-of-art methods by combining them in our new architecture. The producing architecture is based on works [1] and [2]. [2] introduce **StrGNN** that solves the anomaly detection problem in dynamic graph, this architecture solves big class of problems well, but it should be adapted to narrower tasks. In our work we will re-implement it for our purpose in combination with GNN-network from [1]. [1] based on the model presented in [3]. We are going to use climate data from Google Earth Data, CMIP5, NOAA Storm Events Database, Severe Weather Dataset for training and evaluating our architecture.

2 Problem Statement

As mentioned above the architecture that we are going to use is described in [2]. This is our parametric family of function to map design space to target space. The criterion will be log-loss function. We consider Hail Prediction problem as classification problem. Our goal is to classify months in the prediction horizon whether hail will appear or not. Mathematically our problem summarized as following:

$$w^* = \arg\min_w \left(-\frac{1}{N} \sum_{t=1}^{N} (y^t \log(g(w, h^t)) + (1 - y^t) \log(1 - g(w, h^t))) \right) \tag{1}$$

, where:

- w^* is optimal vector of weights of our architecture;
- w is vector of weights of our architecture;
- t is timestamp from the beginning of the prediction horizon;
- N is number of timestamps in prediction horizon;
- y is target class of the timestamp;
- g is fully-connected network;
- h is output of re-implemented version of **StrGNN** [2]

We will compare different impementations by such metrics as:

- Precision
- Recall
- F1
- ROC-AUC
- PR-AUC

References

[1] S. R. Cachay, E. Erickson, A. F. C. Bucker, E. Pokropek, W. Potosnak, S. Osei, and B. Lütjens. Graph neural networks for improved el niño forecasting. *CoRR*, abs/2012.01598, 2020.

[2] L. Cai, Z. Chen, C. Luo, J. Gui, J. Ni, D. Li, and H. Chen. Structural temporal graph neural networks for anomaly detection in dynamic graphs. *CoRR*, abs/2005.07427, 2020.

[3] Z. Wu, S. Pan, G. Long, J. Jiang, X. Chang, and C. Zhang. Connecting the dots: Multivariate time series forecasting with graph neural networks, 2020.