

EnWeather Data Science: Recap

Envision Digital

DS Team

2020-03-27

Machine learning workshop material

First session

- EnWeather Data Science overview (1hour): [link](#)
- Use machine learning model for weather forecast (45min): [link](#)
- **Lab session:**
 - Use XGBoost model for weather forecast (1hour): [link](#)

Second session

- Use deep learning model for weather forecast (1hour+): [link](#)
- Lab session
 - Use Pytorch for weather forecast (1.5hours): [link](#)

Machine learning in weather forecast

- Challenges
 - From forecast to forecast
 - Information from different sources
 - NWP: no long term winner
 - Weather always change
- Solutions
 - Domain knowledge, data analysis, Feature selection techniques, etc.
 - Data fusion
 - Merge NWP rather than select NWP
 - Retrain and select data

Machine Learning Researches in Weather Forecast

Weather forecast groups

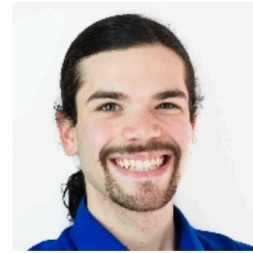
- <https://www.climatechange.ai/>
- Advisors
 - Yoshua Bengio: 2018 ACM A.M. Turing Award

Open and public dataset

- Github: <https://github.com/pangeo-data/WeatherBench>
 - EC reanalysis data for more than 30 years
 - For global weather pattern analysis

Deep learning courses/tutorials/references

- Tensorflow: <https://www.tensorflow.org/tutorials>
- Pytorch: <https://pytorch.org/tutorials/>
- DL: <https://d2l.ai/index.html>



David Rolnick
University of Pennsylvania
Chair



Priya L. Donti
Carnegie Mellon
Co-Chair



Lynn H. Kaack
ETH Zürich
Co-Chair



Yoshua Bengio
Mila, U. de Montréal

How to use XGBoost in Weather Forecast

Use XGBoost for wind speed forecast

- Use cases
 - Optimize wind speed forecast for each wind turbine -> help to estimate wind power
- Stacking of single layer machine learning models
 - use multiple NWP
 - reduce the bias for each NWP

Use XGBoost for rainfall forecast

- Multi-layer single machine learning models
 - First layer rainfall probability
 - Second layer rainfall volume
- Each NWP have very different rainfall pattern
 - EC tends to have more rain
 - IBM tends to have less rain
 - Merge them together may get inconsistent results

Thank
You