WEATHER FORECASTING IN CANADA

BY

OLUYOMI ALABI

Oluyomy

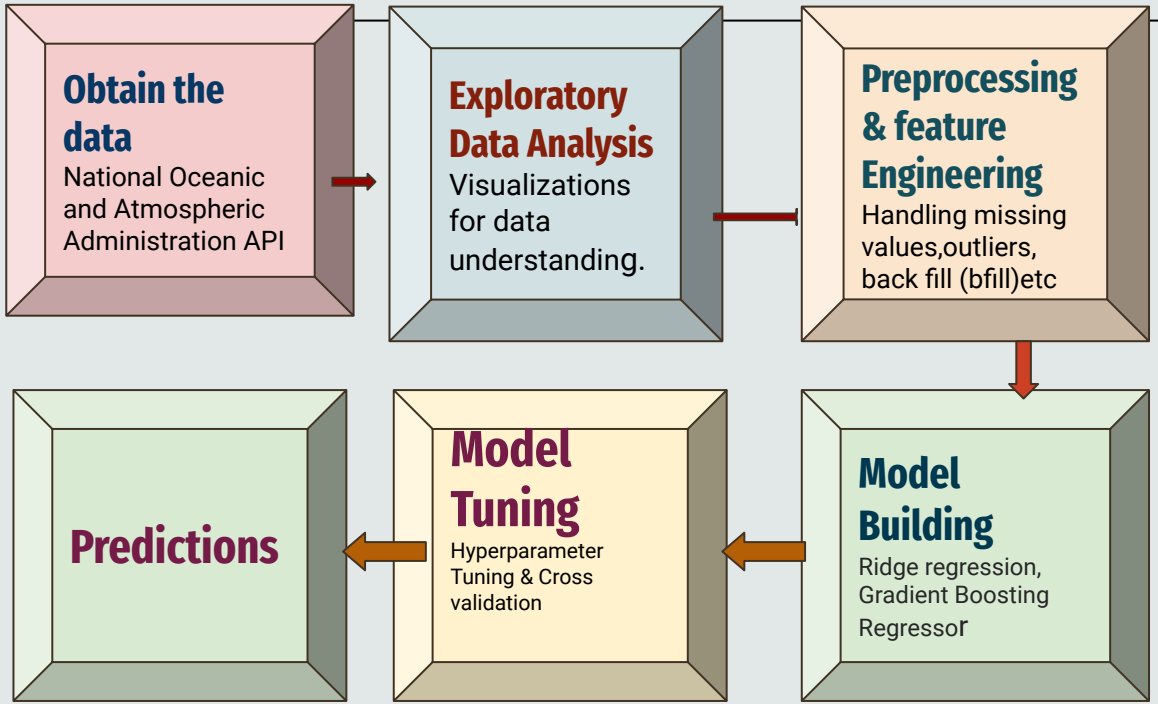
INTRODUCTION

weather forecasting has evolved significantly over the years with the utilization of Machine Learning (ML) to enhance the accuracy and efficacy of predictions.

Its effectiveness is largely derived by training diverse data sources including weather station data, satellite imagery and radar data. For this project,a decade of Calgary airport base station data was acquired from NOAA API achieving 94% predictive accuracy.

- To feed models with data that allowed us to understand the relationships between weather variables.
- Identify the best predictor model.
- Contrast predicted outcomes with the real-world weather conditions to ascertain improvements.

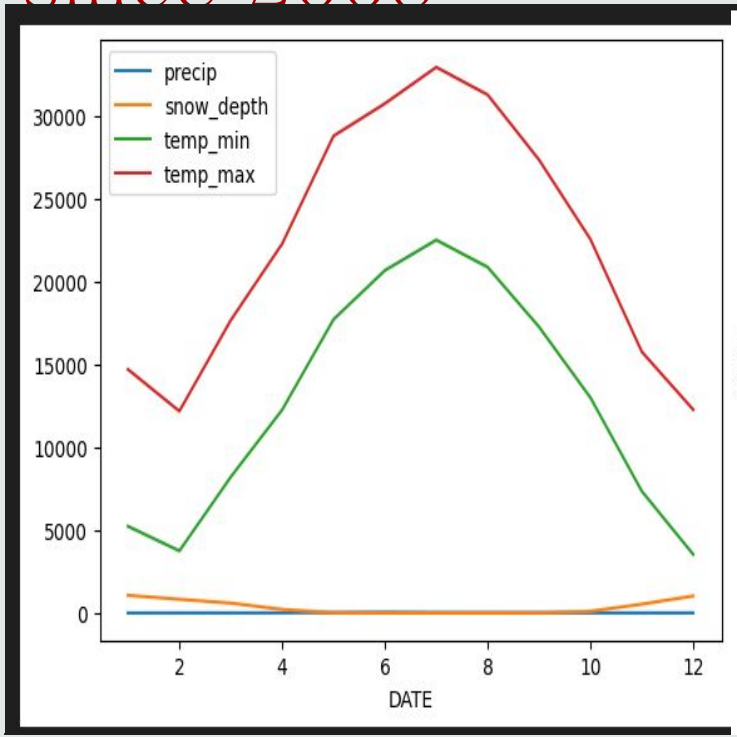
Process

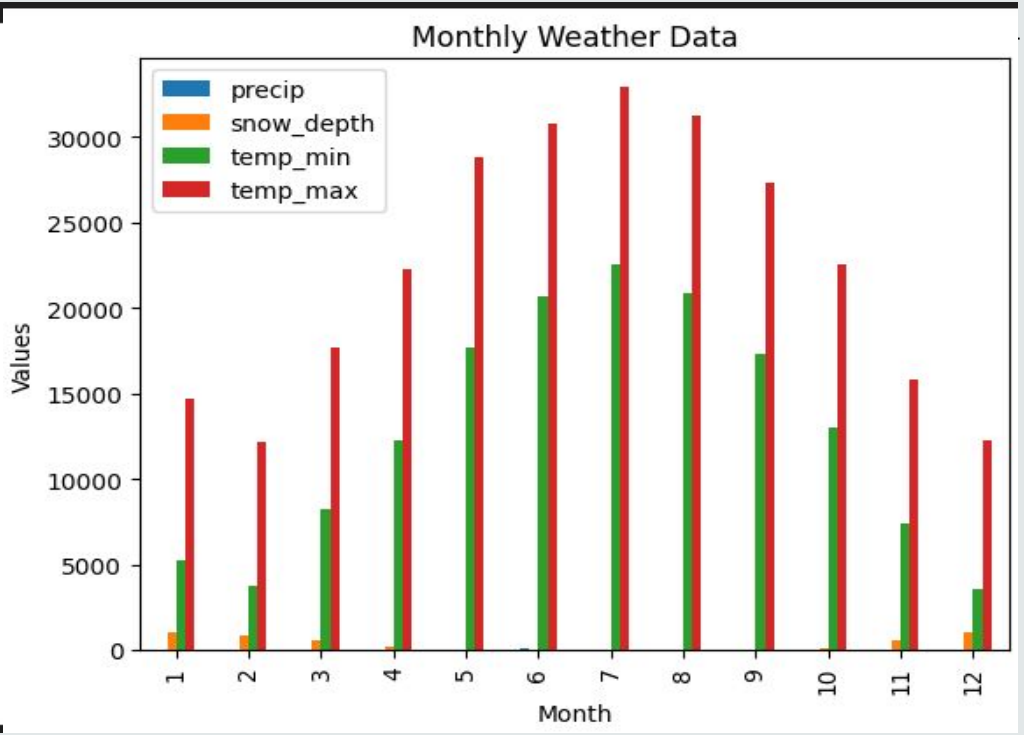


TOOLS

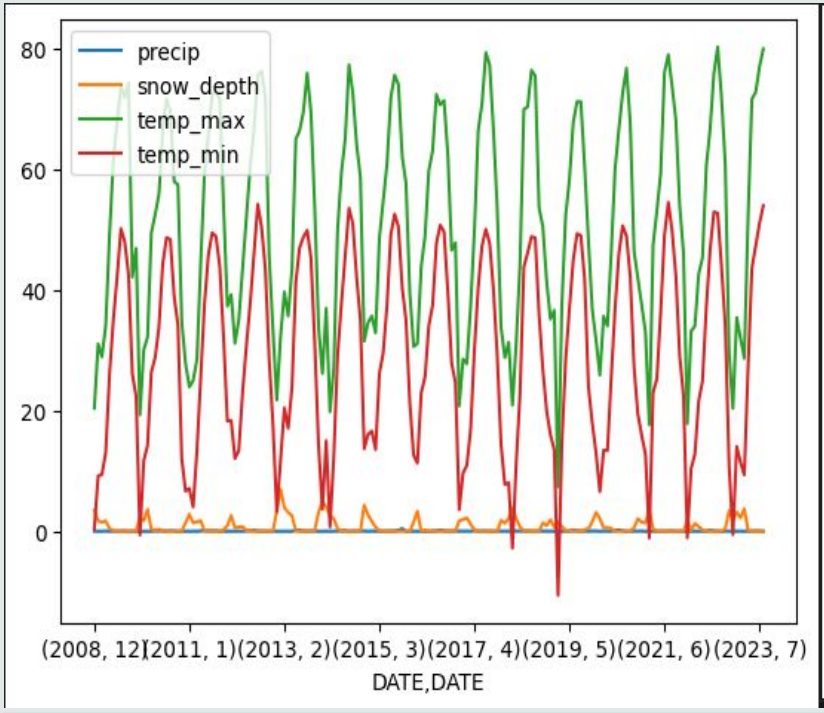
- Calgary International CS, Ca weather data.
- Jupyter Notebook
- Pandas
- Numpy
- Seaborn
- Matplotlib
- Statsmodels
- Scikit Learn

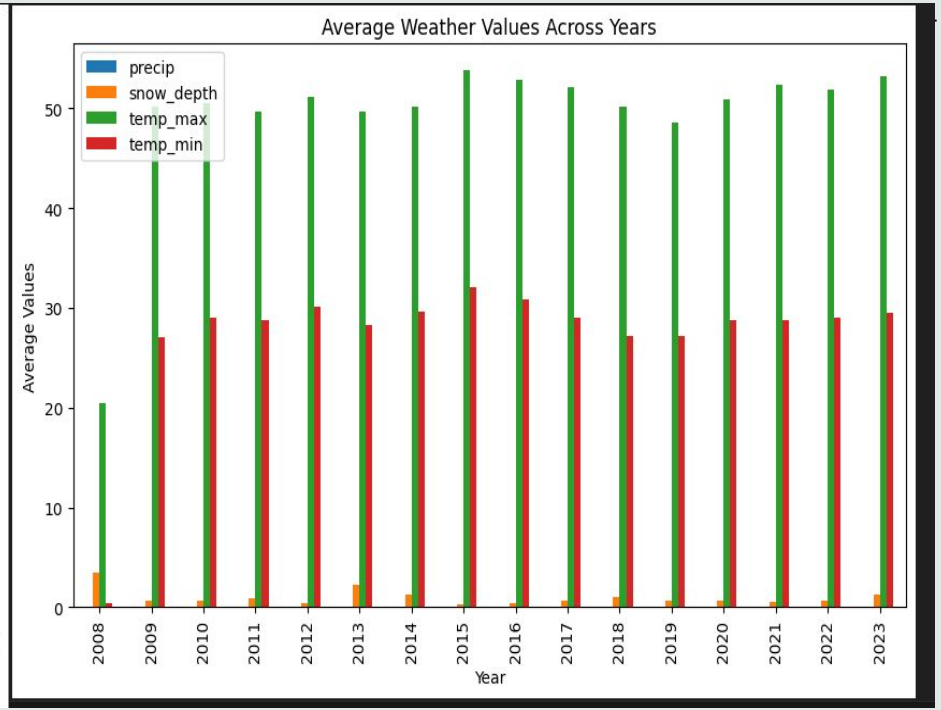
Monthly weather predictors distribution since 2008



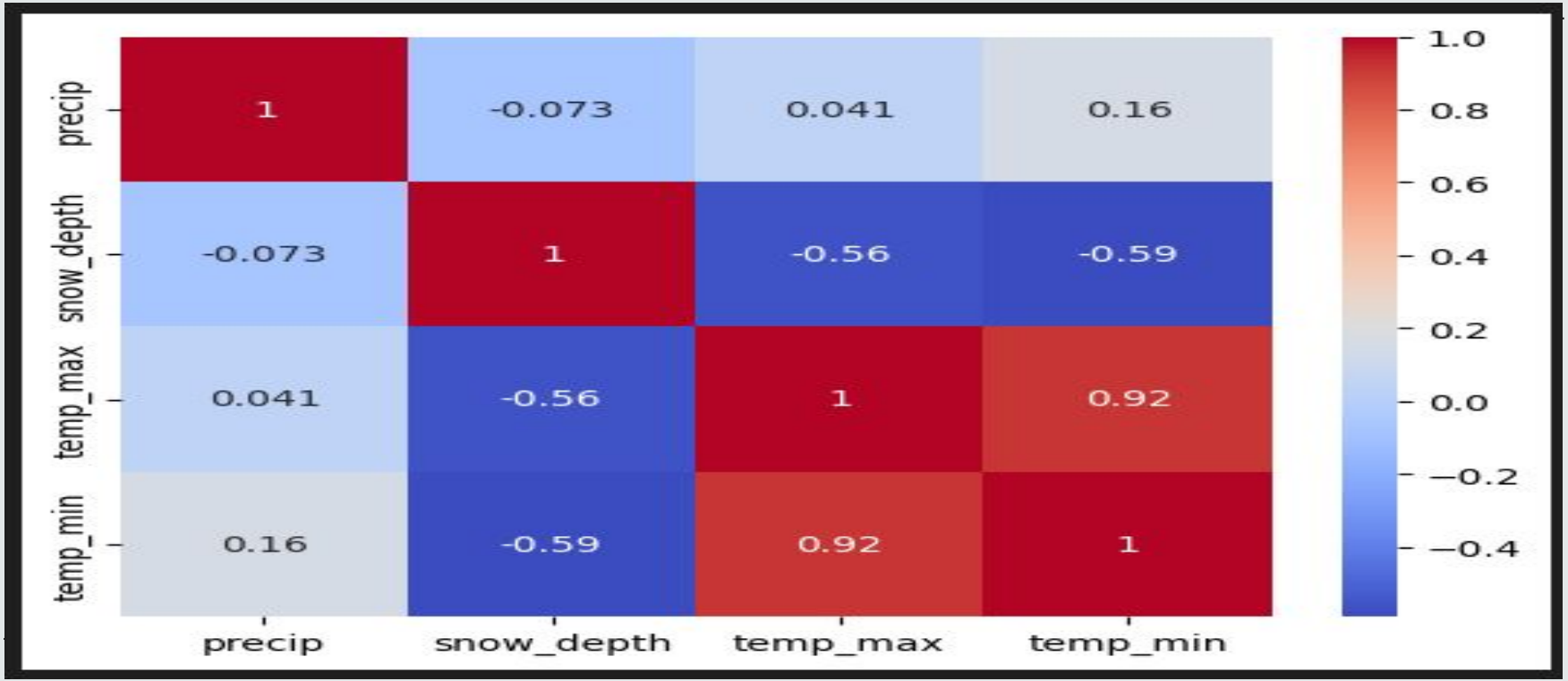


Average weather predictor values across the years

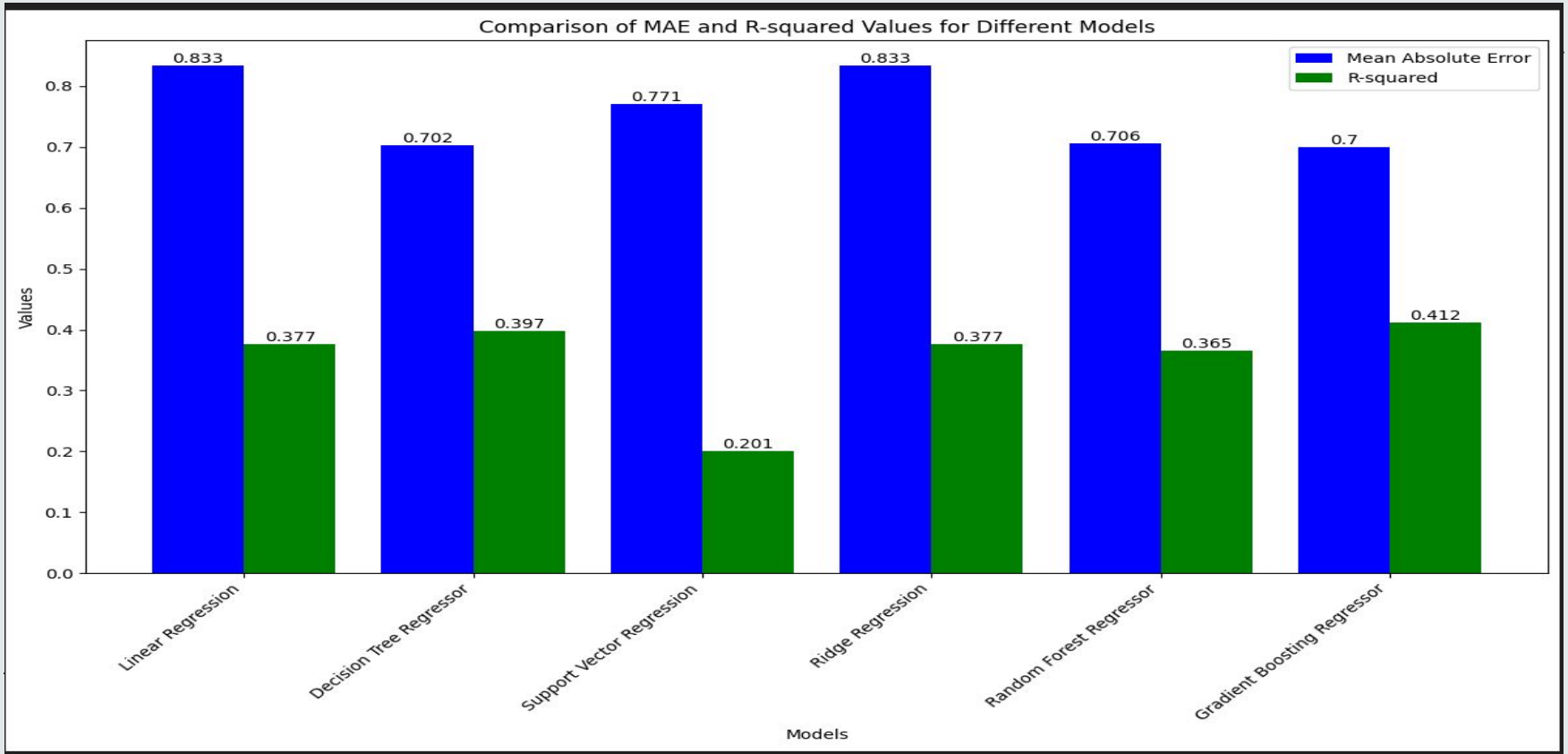




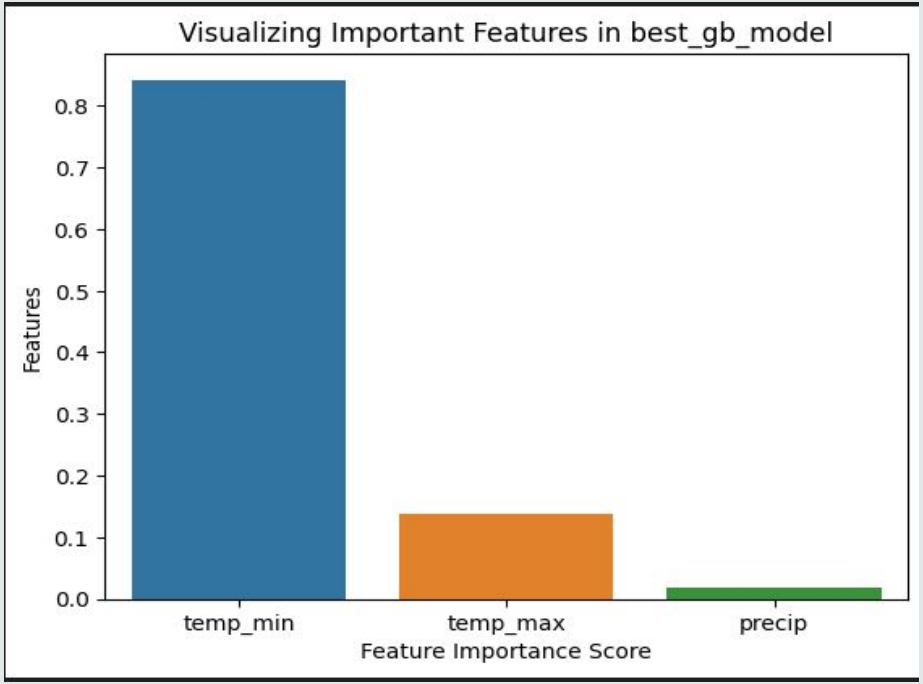
Correlation Matrix showing relationship between variables.



Model Selection Results



Feature scores showing the most influential variable for snow prediction



Temp_min (minimum temperature)=0.842534

Temp_max (maximum temperature)= 0.138820.

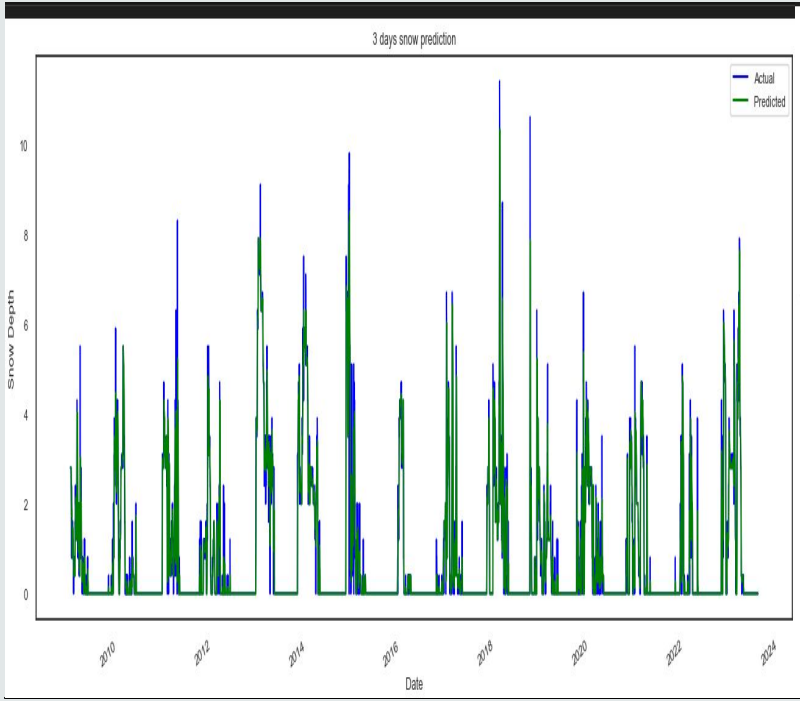
Precip (precipitation)=0.018646.

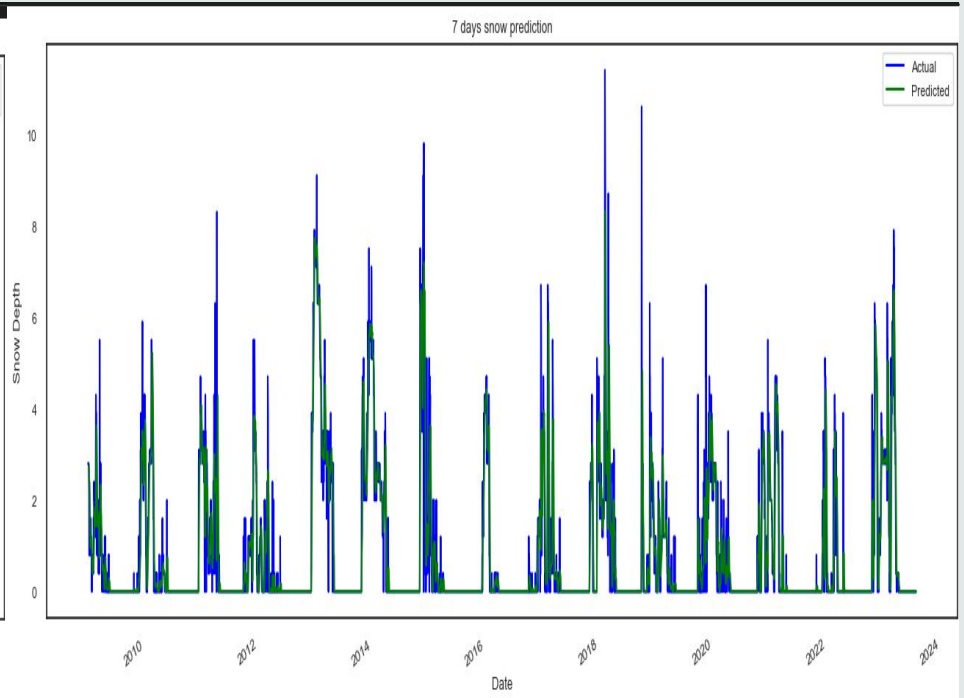
Temperature is often a key factor in weather-related phenomena.

Broader context

I further explored feature engineering by creating new predictors, hyperparameter tuning, which led to improved optimization of the model before making a final decision. Predictions were made for tomorrow's snow_depth, 3 , 7&14 days.

Predictions for the next 3 & 7 days





Conclusion

Accurate weather prediction impacts agriculture, transportation , disaster management and daily planning. Harnessing precise weather predictions transforms challenges into opportunity, enhancing safety, efficiency, and preparedness across crucial domains.

Future goals

- Explore time series prediction using SARIMA (Seasonal ARIMA).
- Exploring Recurrent neural network (RNN) which models extended time dependencies.

Challenges

- API data retrieval complexity.
- Prediction complexity from interactive variables.

THANK YOU