Collective Design Lab

Optimizing Earth Science Observations: Developing Reinforcement Learning Techniques for Autonomously Determining Priority Observations in a Dynamic Environment

Shashwat Raj, Computer Systems Engineering
Mentor: Paul Grogan, Associate Professor
School of Computing and Augmented Intelligence, Fulton School of Engineering

Introduction

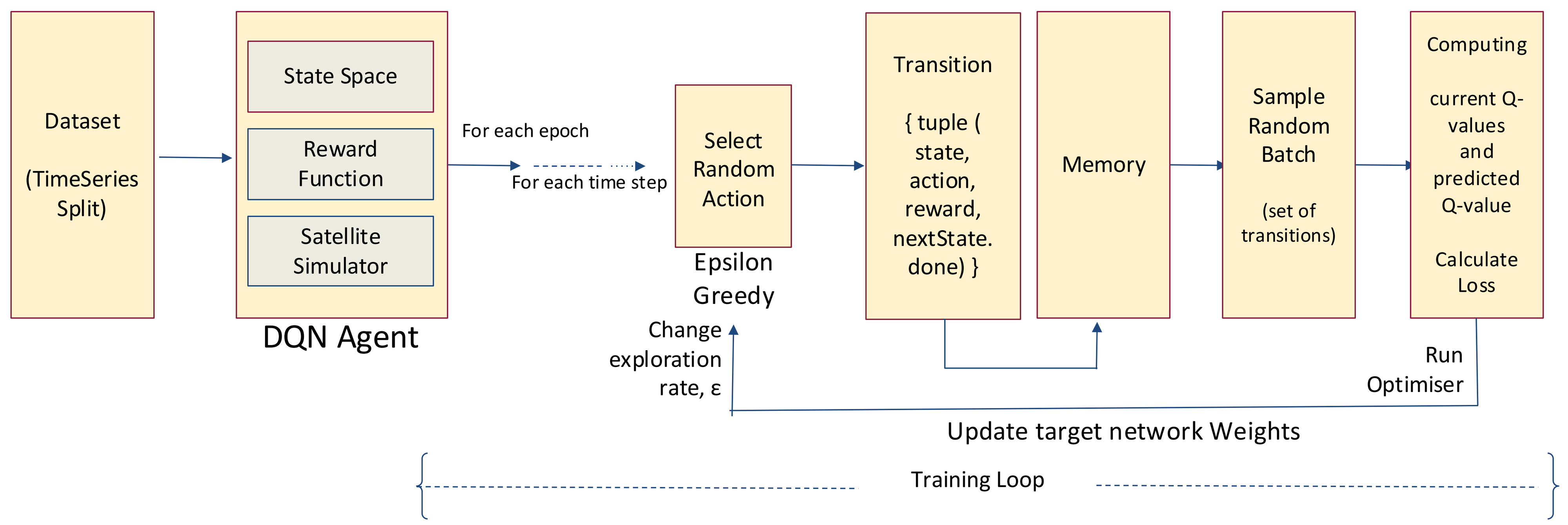
In an attempt to make the conventional system of scheduling satellite for earth science observations autonomous, dynamic targeting has been a highly discussed topic of interest. Helping satellites to autonomously enable the priority observations in space reduces costs, labor and improves efficiency of Earth Science Missions. A reinforcement learning technique is proposed to predict valid observations in a dynamic environment like space. The research question the project addresses is how to devise a near optimal way to prioritize areas with high atmospheric activity, precisely observing convective precipitation, resulting in enhanced environmental monitoring and providing deeper insights to Earth's atmospheric dynamics.

Tools and Technologies

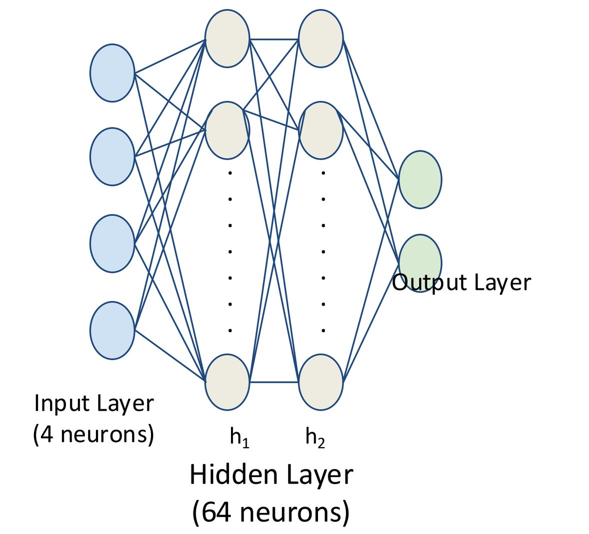
- Tradespace Analysis Tool for Constellations (TAT-C), for early stage trade space exploration and analysis.
- PyTorch and Scikit Learn, a python libraries used to create deep neural networks and analyze data.
- NASA GEOS5 G5NR Dataset for preparing training dataset with satellite orbit, instrument view geometry, temporal aggregation, geolocation, geophysical feature i.e CNPRCP (Convective Precipitation)
- GeoPandas, open source library to use pandas objects with geospatial data types.
- Matplotlib and Cartopy, python library for data visualization.

Methodology

Model Flowchart and Structure



Neural Network Architecture



DQN Agent (Setting up Environment)

State Space : Longitude, Latitude, Solar Hour, Elapsed Time, isGround

Action Space : 0 - decides not to observe; 1 - prioritized observation

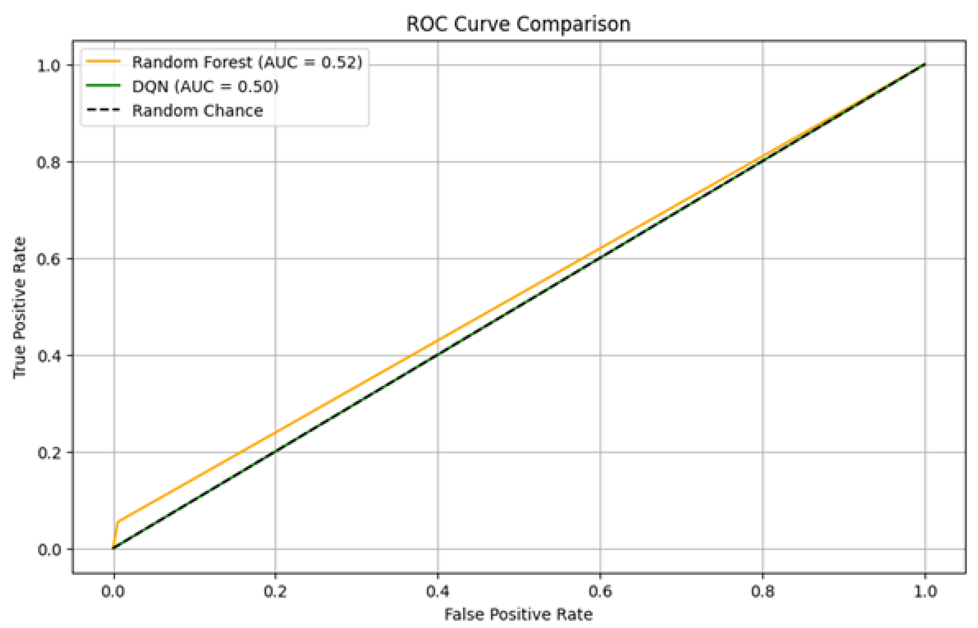
Reward Function :
+1 if Action == 1 and CNPRCP > 0
-1 if action == 1 and CNPRCP == 0
-0.5 if action == 0 and CNPRCP >1
0 if action == 0 and CNPRCP ==0

Simulator : Timestamp in the training dataset is used to emulate movement of satellite and locate the state variables at each time step.

Bellman Equation (Decision Markov Process):

$$V(s) = \max_a \left(R(s,a) + \gamma \sum_{s'} P(s,a,s') V(s') \right)$$

Receiver Operating Characteristic Curve (ROC)

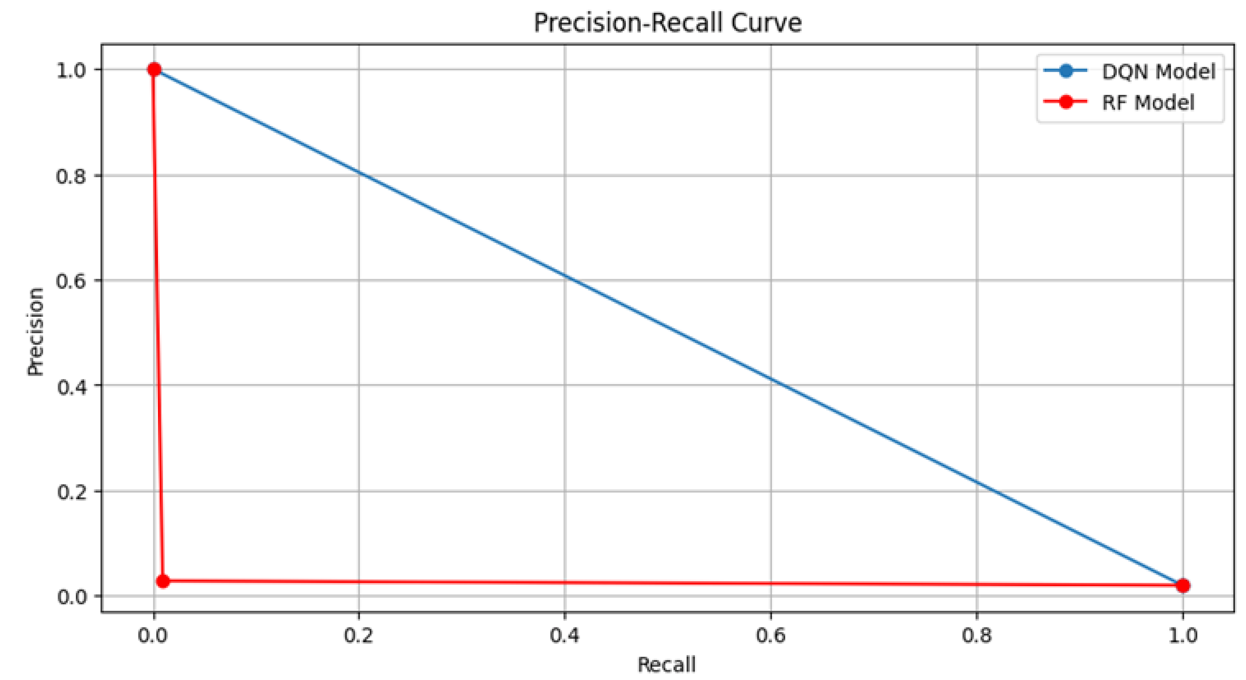


Assessing model's ability to balance recall against all thresholds.

- DQN line almost overlaps with the diagonal line, indicating poor performance and random guessing.
- RandomForest performs much better as shown in the plot.

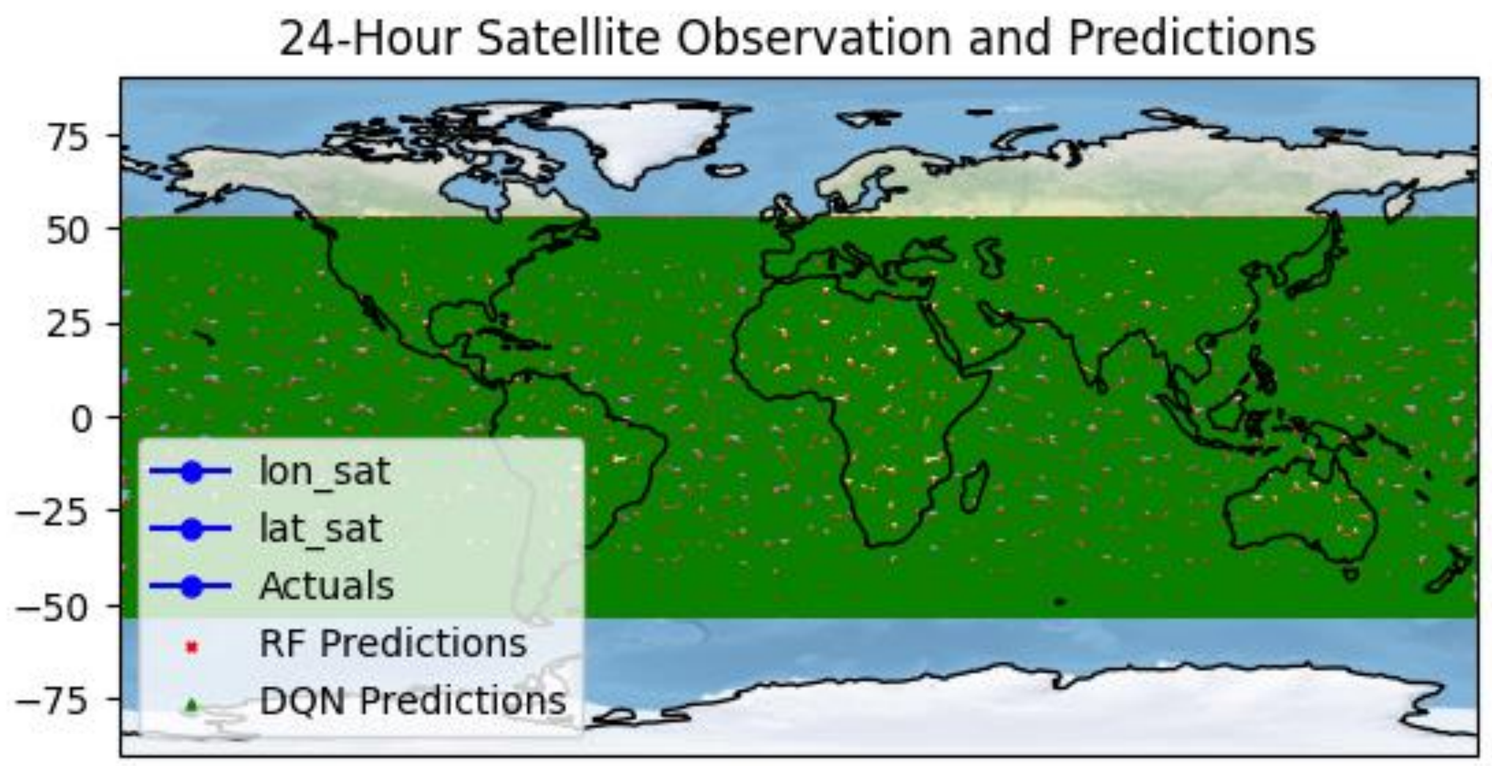
Precision-Recall Curve

Plotting Accuracy of Positive Predictions (true positives) against recall



Monitoring the model's predictions

- In the following plot, DQN model is showing to largely overfit, indicating very high negative rewards.
- On the contrary, the Random Forest Classifier model is able to predict certain convective precipitation events accurately.



Conclusion

The reinforcement learning model's performance is measured by its precision and recall. However, the current version of the machine learning model has a precision of only 4%. The random forest classifier model that was initially used as a comparison model showed 34% precision. There could be several reasons for the ineffective testing performance of the model including re-evaluating reward function, initial random action selection via epsilon greedy algorithm and the hyperparameters

Future Work

Having a dynamic decay rate for epsilon while selecting the action will result in lesser total negative reward. The tradeoff between exploration and exploitation while choosing the action to be taken by the satellite is going to be the primary. Furthermore, future work revolves around addition of state space features like battery cycle as well as discretisation of the action space.

FURI

Acknowledgements

Expressing immense gratitude to Dr. Paul Grogan's mentorship and the support provided by members in his CoDe lab. Special thanks to Mr. Suvan Kumar and Dr. Ismael Tapia Tamayo for their constant guidance

Ira A. Fulton Schools of Engineering
Arizona State University