**PROBLEM STATEMENT**: When operating reliably, the National Oceanic and Atmospheric Administration’s (NOAA’s) space weather station, the Deep Space Climate Observatory (DSCOVR), can measure the strength and speed of the solar wind in space, which enables us to predict geomagnetic storms that can severely impact important systems like GPS and electrical power grids on Earth. DSCOVR, however, continues to operate past its expected lifetime and produces occasional faults that may themselves be indicators of space weather. Your challenge is to use the "raw" data from DSCOVR—faults and all—to predict geomagnetic storms on Earth.

In this project, a Long Short-Term Memory (LSTM) model was chosen. LSTM is a type of Recurrent Neural Network (RNN).

The objective of this project was to construct a machine learning model capable of predicting geomagnetic storms using the Kp index data. The Kp index is a widely used measure of geomagnetic activity in research.

Recurrent Neural Networks (RNNs) are a type of neural network designed to recognize patterns in sequences of data, such as text, genomes, handwriting, or the time series data used in this project. Unlike feedforward neural networks, RNNs can use their internal state (memory) to process sequences of inputs. This makes them ideal for tasks where the order of the data points is important.

However, traditional RNNs have a problem known as the vanishing gradient problem, which makes it difficult for them to learn and tune the model parameters during training, especially when dealing with long sequences of data.

Long Short-Term Memory (LSTM) networks are a type of RNN that are designed to avoid the vanishing gradient problem. LSTMs are explicitly designed to avoid the long-term dependency problem. They can remember information for long periods of time, which is a major advantage for many tasks, including our task of predicting geomagnetic storms based on historical data.

The LSTM model was defined using the Keras library. The model was set up with one LSTM layer with 50 units and one dense output layer. The model was compiled with the Adam optimizer and the Mean Absolute Error (MAE) loss function.

The model was then trained on the training data for 50 epochs with a batch size of 72. An epoch is one complete pass through the entire training dataset, and a batch is a subset of the training data. The batch size determines how many samples the model should “see” before updating its weights. The number of epochs and the batch size are hyperparameters that you can tune to find the best values for your specific problem.

During training, the model learns to map the input data (the Kp index from the previous hours) to the output data (the Kp index at the current hour). It does this by adjusting its weights based on the error of its predictions, as calculated by the loss function. The optimizer determines how the weights are updated.

After training, the trained model’s performance was evaluated on the test data using the Root Mean Squared Error (RMSE) metric. The model’s predictions were compared to the actual values, and the RMSE was calculated to quantify the model’s prediction error. The RMSE was found to be 0.051, which is a good value indicating that the model’s predictions are close to the actual values.

The model was used to make predictions on new data. The new data was preprocessed in the same way as the training and test data, and the model’s predictions were inverse transformed to get them back to the original scale.

Throughout the project, there were some issues with data types and reshaping data, but these were resolved by converting the data to the appropriate type and shape.

The project was successful in building a machine learning model with a good performance (low RMSE) on the test data. This model can now be used to predict geomagnetic storms using the Kp index data, which can be very useful in space weather forecasting.

In this project, a Long Short-Term Memory (LSTM) model was chosen. LSTM is a type of Recurrent Neural Network (RNN).

The objective of this project was to construct a machine learning model capable of predicting geomagnetic storms using the Kp index data. The Kp index is a widely used measure of geomagnetic activity in research.

Geomagnetic storms that can severely impact important systems like GPS and electrical power grids on Earth. Deep Space Climate Observatory (DSCOVR) can measure the strength and speed of the solar wind in space. Right now, these predictions are not available to the general public. We aim to use the raw data from DSCOVR to predict Geomagnetic Storms on Earth. Geomagnetic storms can severely impact important systems like GPS and electrical power grids on Earth. Our prediction model will significantly help in safeguarding these systems by providing a precautionary measure beforehand.

It will also help small-scale organizations and start-ups related to rocketry/telecommunications /power etc. For example, if someone wants to launch a rocket, they need to make sure that optimized conditions are satisfied beforehand. Our model will inform them about the geomagnetic disturbances in space ahead by 2 hours. This will prevent the failure of rocket launch due to unfavourable environmental conditions, thus, saving resources and capital.

Our mission is to empower society with the knowledge and tools to prepare for geomagnetic storms. We are committed to making our predictions accessible, user-friendly, and open to all who seek to understand and safeguard against the effects of these natural phenomena. By providing reliable forecasts, we aim to mitigate potential damages and ensure a more resilient, connected, and prepared world.