Twins of The Winds

Team: Cosmic Energies

Project Report

Initial Preprocessing
The training data contained missing values and was arranged in a random order, so it was required to combine the date, month and year columns to get the final date in dd/mm/yyyy format and arrange the data in chronological order for ease of viewing the data.

Preprocessing and filling up NaN Values
cKDTree was used to fill up the missing values because it allows us to get the data from nearby locations on the same date which are closer to the real data as compared to taking the mean or last observation carried forward method.

Training and Testing
Used the Random Forest Regressor to train the model and cKDTree again to impute any missing values in testing files. The model was finally used on testing files to produce the predictions of sea surface temperature with minimal mean square error of 0.0253616….

Contribution Report

Raghuveer Verma
- Cleaning the training file and sorting the data by time.
- Suggestions for improving preprocessing of the data
- Training the AI Model

Yashaswini L

- Filling missing data of given files using the cKDTree imputer
- Reduced the error of the trained model
- Implementing the trained AI Model to get the final predictions