README

This README outlines the process for extracting, preprocessing, and combining data for the SAMPLEX satellite, as well as executing predictive modeling and subsequent calculations.

Data preparation

Follow these steps in sequence to prepare the data:

1. HASDM_preprocess.py
    - Fetches data from 2000-2004 at the altitude of the SAMPLEX satellite.
    - Converts GSM_X/Y/Z coordinates to Altitude/Longitude/Latitude.
2. Space_data_preprocess.py
    - Preprocess the fetched data for 4 years, focusing on the space weather around the SAMPLEX satellite.
3. Space_data_combine.sh
    - Combines all CSV files of the preprocessed space data into a single file.
4. Space_data_merge_preprocess.py
    - Converts millisecond data to hourly data.
    - Renames the combined data of 4 years for consistency.
5. RHO_prediction_combiner.py (1st and 4th step data)
    - Combines RHO values with space data based on the satellite's position and timestamp.

Prediction Model

To run the prediction models, follow these instructions:

1. RHO_Regression_Model.py (2000-2004)
    - Applies z-score normalization and PCA to reduce dimensionality of the data.
2. Predict_RHO.py: Fetches the 1999 space weather dataset (excluding RHO), preprocesses it, and runs the prediction model to add a new column predicted_RHO.
3. Satellite_drag_calculation.py (1999)
    - Calculates drag using predicted_RHO based on values referenced from MDPI article.

Additional Predictive Modeling

For further analysis:

1. Merges communication metric (signal strength and latency)with the 1999 space weather dataset that includes predicted_RHO and satellite drag calculations.
2. communication_metric_regression_model.py
    - Fetches communication metric dataset for 1999 for the SAMPLEX satellite.
    - Pre-process it.
    - Runs regression models to predict signal latency based on the combined datasets.
3. Calculate_fuel_when_needed.py
    - Calculates fuel requirements when signal strength is below a safe threshold or latency exceeds a safe limit.