Satellite Anomaly Detection and Explanation

1st Aniket Ghetla
Applied Data Science and Analytics
SRH Hochschule
Heidelberg, Germany

2nd Bhargav Deekshith Vasu
Applied Data Science and Analytics
SRH Hochschule
Heidelberg, Germany

3rd Kiran Mulawad
Applied Data Science and Analytics
SRH Hochschule
Heidelberg, Germany

4th Pooja Patil
Applied Data Science and Analytics
SRH Hochschule
Heidelberg, Germany

5th Spoorti Basarkod Math
Applied Data Science and Analytics
SRH Hochschule
Heidelberg, Germany

***Abstract*—The focus of this work is to develop a system that can recognize and analyze anomalies in the data returned from the planet by the Mars Express Orbiter (MEX). By examining three years' worth of Martian telemetry data, we want to estimate the power consumption of the spacecraft and identify any unusual patterns that might affect its mission. Three groups of anomalies are identified: natural phenomena like solar flares, unexplained occurrences, and operational issues. The behavior of the spacecrafts and the anomalies are identified by our solution. This technique may be applied to other dynamic systems like cloud computing and offers a fresh viewpoint on abnormalities in satellite data. Our ultimate goal is to manage and monitor satellites more effectively, hence increasing the reliability and effectiveness of space exploration.**

***Index Terms*—Satellite anomaly detection, Mars Express Orbiter, telemetry data, anomaly explanation, power consumption prediction, machine learning.**

I. Introduction

Mission success in space exploration depends on the spacecraft systems' dependability, and long-term operations critically depend on their health. After years of orbiting Mars, the Mars Express Orbiter (MEX) has amassed a vast amount of telemetry data that sheds light on the planet and the spacecraft's operational behavior. The orbiter's continued operation depends on the analysis of this data, particularly since a number of variables, including modifications to operations or outside events like solar flares, might affect power usage. The method for identifying and analyzing abnormalities in MEX's telemetry data is presented in this study, with a particular emphasis on predicting power consumption during a three-year period on Mars. The system recognizes any problems that might occur in the spacecraft's subsystems by looking at trends. Unexpected causes, natural occurrences, and operational issues are the three main categories into which the anomaly detection model divides anomalies. Large-scale preprocessing and feature engineering were applied to telemetry data sources, such as subsystem orders, event timelines, solar aspect angles, and long-term solar distance measurements, in order to improve anomaly identification. In order to adjust dynamically to operational changes in the spacecraft's behavior, we used a variety of machine learning models, choosing the best-performing strategy. The Behavioral Approach was especially successful at capturing the intricacies of MEX's operations since it used sliding temporal periods. This approach, which constantly adapts to changing operating states and environmental conditions, has promise for wider applications outside space missions, including cloud computing and other complex systems.

II. Literature review

A. Power Consumption Literature Study

According to the research "Data stream mining for predicting the thermal power consumption of the Mars Express spacecraft," the European Space Agency's (ESA) Mars Express mission has been supplying vital information about the planet for more than 18 years. Accurate thermal power consumption (TPC) forecasts are crucial to the spacecraft's operations because as it ages, its batteries deteriorate. This study investigates the use of telemetry data to estimate TPC in real-time over 33 power lines through the use of data stream mining algorithms. The study assesses the impact of time resolution on model performance by contrasting local and global approaches to multi-target regression. In order to adjust to an aging spacecraft, it is imperative to identify concept drift, or changes in the system's behavior over time. This study looks into how to detect these changes. The findings show that data stream mining is feasible for identifying abnormalities and predicting TPC, with possible implications for future space missions and dynamic systems.

Based on the study "Predicting thermal power consumption of the Mars Express satellite with machine learning" : Recently, there has been a growing application of machine learning (ML) approaches to forecast spacecraft operations, particularly in challenging environments such as Mars. A machine learning pipeline was presented by Breskvar et al. (2017) to forecast the thermal power consumption (TPC) of the Mars Express (MEX) probe. The study showed that ML models provide higher predicted accuracy and perform better than the European Space Agency's (ESA) manually built models. This increased accuracy makes it possible to use resources for science operations more effectively, which could extend MEX's mission duration. The authors used random forests and predictive clustering trees (PCTs) to evaluate telemetry data, taking use of characteristics like subsystem commands and energy influx. Their method also shed light on the thermal behavior of the spacecraft, showing how TPC is influenced by internal (subsystem activities) as well as external (solar energy) elements. This work expands the possibilities of using similar techniques to additional satellite subsystems and emphasizes the potential of ML to improve spacecraft performance.

B. Anomaly Detection Literature Study

Based on the paper, "Anomaly detection of satellite telemetry data based on extended dominant sets clustering": The research provides a clustering-based anomaly detection approach customized for satellite telemetry data. Because of the high dimensionality and complexity of spacecraft systems, unsupervised anomaly identification in complicated satellite telemetry is a common problem without labeled data. This approach is intended to address this issue. It presents an expanded approach for clustering dominant sets that can group data of any shape. The technique uses cluster analysis to find anomalies; data points that belong to small clusters or are not assigned to any cluster are regarded as anomalies. On the basis of relative similarity, anomalies within sizable clusters are also identified. By raising the Area Under the Curve (AUC) values by 3% to 10%, the suggested method improves anomaly detection accuracy. Experiments conducted on synthetic datasets and real telemetry data from the Tianping-2B satellite confirm the method's efficiency. It is able to identify anomalies in magnetometer data and produce "anomaly windows," which shed light on the condition of the satellite. By monitoring spacecraft in orbit, this clustering-based technique provides a reliable and strong strategy that enhances safety and dependability. All things considered, the method tackles the difficulty of unsupervised anomaly identification and offers a significant addition to the study of satellite telemetry data.

III. Methodology

A. Pre-processing and Feature Engineering

This section describes the feature engineering and preprocessing used to prepare the telemetry data from the Mars Express Orbiter for behavior pattern recognition and power consumption prediction. Context files (SAAF, DMOP, FTL, EVTF, LTDATA) and observation files (power data) make up the raw data; each one represents a distinct facet of the spacecraft's operations.

1) Data Source and Structure: The data files were temporally aligned and suitably prepared for model training using the preparation processes.

2) DMOP (Detailed Mission Operations Plan) Preprocessing: Timestamped UTC values were preprocessed into a standard datetime format in order to prepare this file. Removing the first four characters from each subsystem field allowed us to aggregate subsystem commands. To determine how often each subsystem occurred throughout these time frames, the data was then resampled into 15-minute intervals. Co-occurrence counts were also used to monitor interactions between predefined subsystem pairs. Thanks to this preprocessing, important operational patterns and relationships between the spacecraft's components could be found, offering important new perspectives on how the spacecraft behaved and used power over time.

3) EVTF (Event Timeline) Preprocessing: Several processes were required in the preprocessing of EVTF files, which log occurrences such as eclipses and signal dropouts. The first step was to extract event keywords, like "loss of signal" (LOS) and "acquisition of signal" (AOS), to produce binary features that showed whether these events were present. Next, the timestamps of the events were synchronized at 15-minute intervals. Finally, occurrences of each event category were resampled and totaled across these intervals to capture their frequency across time.

4) FTL (Flight Dynamics Timeline) Preprocessing: There were multiple phases involved in the preprocessing of FTL files, which list spaceship pointing events that alter its attitude. EARTH, SLEW, and NADIR were the first pointing types to be one-hot encoded into binary columns. Subsequently, a standard datetime format was employed to transform the utb_ms timestamps. Finally, the occurrences of each pointing event type were aggregated across 15-minute intervals to measure their frequency and impact on the spacecraft's activities.

5) SAAF (Solar Aspect Angles) Preprocessing: In order to appropriately reflect the spacecraft's placement, cosine transformations were applied to the solar aspect angles in the SAAF data, which indicate the spacecraft's orientation with respect to the Sun. In order to provide a consistent analysis over time, the data was then resampled and the mean of the angles was aggregated into 15-minute intervals.

6) LTDATA (Long Term Data) Preprocessing: LTDATA files, containing distances between celestial bodies, were resampled into 15-minute intervals and forward-filled to ensure alignment with other telemetry datasets for analysis.

7) Power Data Preprocessing: Power data, representing electric currents across thermal power lines, was resampled and interpolated into 15-minute intervals after timestamp conversion.

8) Feature Engineering: A key feature, *energy_received*, was derived using cosine-transformed solar aspect angles and the Sun-Mars distance to estimate the spacecraft's solar energy intake. This was calculated using the formula:

$$\text{energy\_received} = \frac{p_{\max} \times \cos(\text{angle})}{\text{sunmars\_km}^2}$$

Where $p_{\max}$ represents the maximum possible energy flux. This feature is critical for accurately modeling the spacecraft's power consumption and understanding its energy efficiency in varying operational states. Additionally, subsystem interactions and event-based features were engineered to capture the spacecraft's overall performance trends.

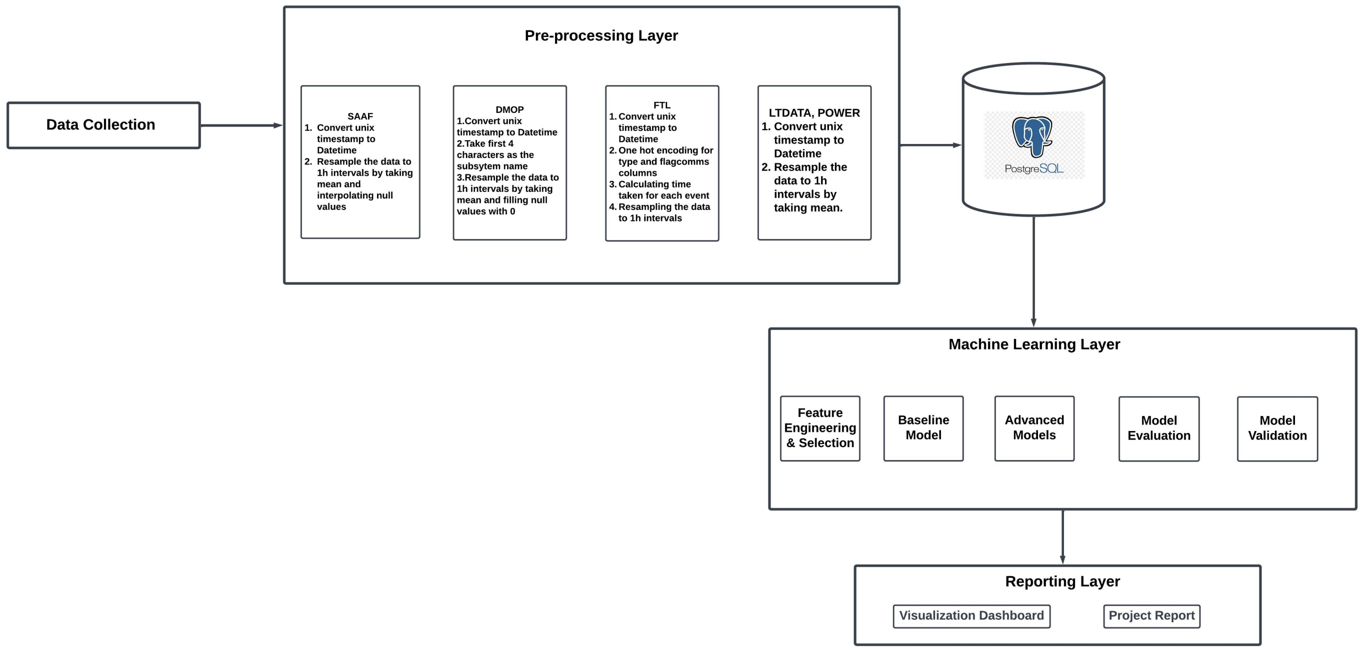

Fig. 1. Pipeline 2 Architecture

B. Machine Learning Models

The methods used to predict the Mars Express Orbiter's power usage and behavior are described in this section. After rigorous model testing, the behavioral approach that made use of time frames and model selection yielded the best results. Below is a synopsis of the methods that were tried and the reasoning for the chosen model.

1) Baseline Linear Regression Model: The first method used satellite telemetry and the top five most correlated characteristics for each line to create a basic linear regression model that predicted power usage for thermal power lines.

- Results: Average RMSE: Varies between power lines; some power lines demonstrated fair performance, but others resulted in low predictive power. R2 Score: The linear model was unable to capture complicated relationships within the data, as evidenced by the poor average R2 score over the majority of power lines.
- Conclusion: In summary, while linear regression provided a useful foundation, it proved to be overly simplistic to manage the intricacies of satellite operations.

2) Random Forest Model: Next, in order to enhance the linear regression model, we employed a Random Forest Regressor, a non-linear model. Capturing more intricate connections between the telemetry data and power consumption was the aim.

- Results: Best RMSE for Power Line NPWD2372: 0.0196. Best R^2 Score for Power Line NPWD2372: 0.699. While performance improved for some power lines, others still resulted in low or negative R^2 scores, suggesting further refinement was needed.
- Conclusion:Particularly for some power lines, the Random Forest model fared better than linear regression. To better represent the non-linearity in the data, more advanced modeling approaches were required, according to the overall average R2 score.

3) Ensemble Model Approach (XGBoost and Extra Trees): We looked into ensemble models, particularly XGBoost and Extra Trees, which are renowned for their capacity to capture both linear and non-linear interactions, in order to further improve predictive performance.

- Approach: We first identified the top 40 most important features using an Extra Trees Regressor for each power line. We then trained both XGBoost and Extra Trees on these selected features, using each power line as the target.
- Results: XGBoost: Average RMSE: 0.0447, Average R^2: 0.2028. Extra Trees: Average RMSE: 0.0451, Average R^2: 0.1121. XGBoost slightly outperformed Extra Trees in terms of RMSE and R^2, showing better predictive accuracy across power lines.
- Conclusion: With XGBoost and Extra Trees, the ensemble model technique demonstrated potential and greatly enhanced forecasts for specific power lines. The model's performance differed based on the specific power line and telemetry data, albeit not all power lines had strong R2 values.

4) Anomaly Detetion Technique: To identify anomalies in spacecraft telemetry data, we suggest an unsupervised anomaly detection algorithm that combines Extreme Value Theory (EVT) and Recurrent Neural Network (RNN) based on Gated Recurrent Units (GRUs). Prior to being divided into training

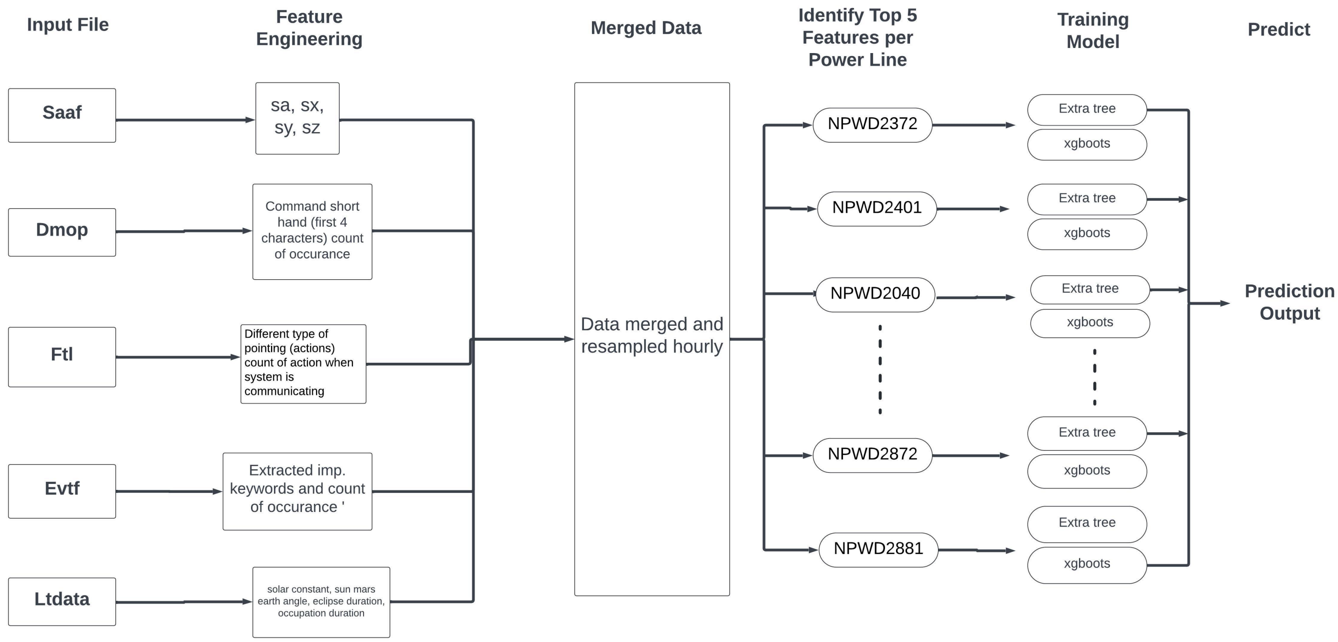

Fig. 2. Ensemble Model Approach (XGBoost and Extra Trees)

and test sets, the data is first normalized using Min-Max scaling. Sequences of ten time steps are then used for prediction. Three GRU-based predictors that learn typical telemetry behavior make up the model's prediction layer. Their outputs are averaged by a decision layer to improve accuracy. The Mean Squared Error (MSE) loss function is used to train the model, and the Adam optimizer is used to optimize it. Prediction errors are computed by comparing the real and predicted telemetry data after training. We use EVT, namely the Peak Over Threshold (POT) approach, concentrating on extreme deviations to identify abnormalities. By looking at values that are higher than a predetermined threshold—which is selected using a quantile of the prediction error distribution—the POT technique finds abnormalities. A Generalized Pareto Distribution is then used to model excesses, or deviations that exceed the threshold (GPD). Accurate anomaly detection is made possible by the shape, position, and scale parameters of the GPD, which aid in defining the distribution of these high values. This method improves the dynamic spacecraft systems' ability to detect uncommon, critical anomalies.

$$\epsilon_{\text{tr}} = \theta + \frac{\sigma}{\xi}\left(\left(\frac{q \cdot n}{N_\theta}\right)^{-\xi} - 1\right)$$

Where:

- θ: Initial threshold set by the quantile
- σ: Scale parameter of the GPD
- ξ: Shape parameter of the GPD
- n: Total number of data points (errors)
- N_θ: Number of excesses beyond the threshold θ
- q: Quantile value representing the rarity of the events we want to capture (e.g., $q = 1 \times 10^{-3}$ corresponds to the 0.1% rarest events)

This formula adjusts the threshold dynamically, considering the shape and scale of the extreme values, to ensure that anomalies are captured as accurately as possible. Every prediction error is compared to the dynamically changed threshold ϵ_{te} in order to identify abnormalities. When further data is processed, the threshold is adjusted using the fitted GPD. Errors over this threshold are reported as anomalies. With a test loss of 0.0123, the model demonstrated accuracy in its predictions under typical circumstances. However, without labeled anomaly data, evaluating the model's capacity to recognize actual anomalies remains problematic. Additional verification using tagged data is required for a thorough evaluation of the methodology.

5) Final Model Behavioral Approach with Windowed Modeling: Because the Behavioral Approach could take into consideration how spacecraft behaviors will change over time, it was chosen as the final model. Using this method, many models were trained for various time periods in order to capture fluctuations in satellite activities that affect power usage.

- Approach: We used sliding time windows of 7 hours with a 1-hour overlap to capture the changing behavior of the satellite. For each window, we trained five models: Random Forest, XGBoost, CatBoost, Support Vector Regressor and LSTM and selected the model with the lowest RMSE for that time window. Further we noticed SVR and LSTM were not fitting properly so we eliminated the two models (refer below image) and kept only three models. To increase our model efficiency, we tried to select top 10 features f and hyper tune the model parameters, but

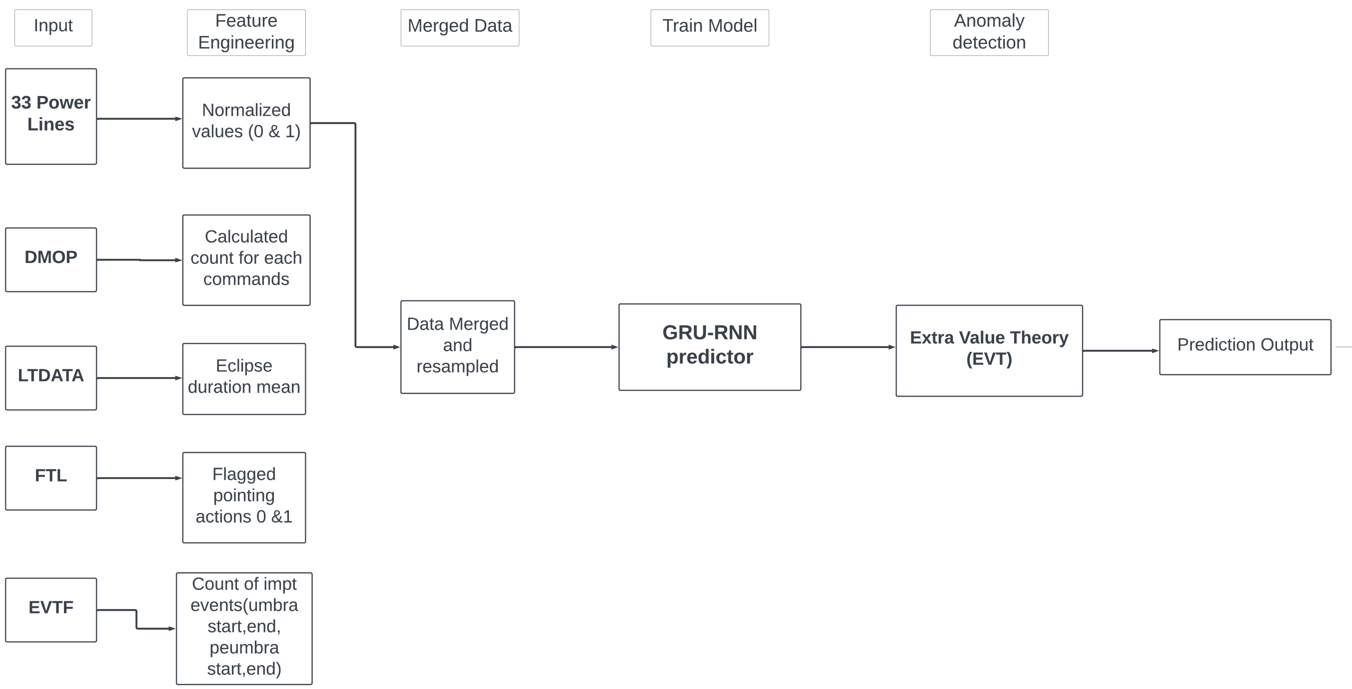

Fig. 3. Anomaly Detection Technique

it did not impact much on the model performance so we eliminated this approach. Models that performed well for continuous windows persisted, while new models were trained if there was a change in behavior (as indicated by RMSE scores).

- Results: RMSE for all windows combined: 0.0306. Top Models: Based on the best performance for each time, either RandomForest, CatBoost, or XGBoost was chosen, depending on the window. Smooth Transition: The spacecraft's usage of overlapping windows allowed for a seamless transition between its various operating phases.
- Conclusion: The model that performed the best, the Behavioral Approach, offered the adaptability needed to adjust to changing spaceship conditions. The most accurate model for predicting power consumption behaviors was the one with the lowest overall RMSE of all, chosen by dynamically correcting for behavior alterations and picking the optimal model for each window.

C. Data Visualization

The code offered creates a dashboard for data visualization utilizing Streamlit and telemetry data from the Mars Express Orbiter (MEX). Prior to processing and loading the telemetry and prediction data, Min-Max scaling and temporal alignment are used. Users can choose sun aspect angles and filter the data by date using a sidebar. Key metrics are displayed on the dashboard, including the energy collected by the spacecraft based on solar aspect angles.The top 10 operational commands over a period of time which can help detect anomalies like Loss of Signal, Acquistion of Signal at a particular time. A thorough chart of satellite model performance that shows the evolution of expected power values over time is also included in the dashboard. This is accomplished by identifying behavioral shifts in the satellite using machine learning models like XGBoost, Random Forest, and CatBoost. These visualizations are combined on the dashboard to provide an interactive experience that allows users to hover over plots to view detailed numbers, filter data by particular dates, and track the performance of various models over time.

D. Docker and Database

The automation of data intake, preprocessing, model training, testing, and Dockerization are the main goals of the Mars Express Orbiter Data Processing Pipeline project. Containerizing each component will ensure a scalable, reproducible, and modular solution.

1) Dockerizaton: Key objectives include isolating services, ensuring repeatability, scalability, and streamlining execution using Docker Compose. Services like the database, data intake, preprocessing, model training, and testing are defined in the docker-compose.yml file. Every service relies on the PostgreSQL database and operates within a separate container. Services are started in the right order thanks to environmental factors and health examinations.

2) Dockerfile Configurations: Services are containerized using lightweight Python 3.9-slim images. Every Dockerfile launches an entrypoint script, installs Python and system requirements, and configures the working directory. Entrypoint.sh makes ensuring that services wait to start operating on their own until the database and dependencies are available.

3) Database Implementation: Because of PostgreSQL's interoperability and resilience, it was selected. It is defined with environment variables for setup in docker-compose.yml. Environment variables are used to establish a connection between the database and application scripts. The insert_data.py script manages data ingestion by dynamically generating tables and adding data from CSV files.

4) Challenges and Solutions: Service coordination, volume control, and database preparedness are major obstacles. Docker Compose's depends_on directive, shared flag files, volume mounts, and wait-for-it.sh to manage dependencies and service readiness were used to address this.

5) Conclusion: For the Mars Express Orbiter Data Pipeline, Dockerization with PostgreSQL configuration guaranteed a scalable, repeatable, and modular system that successfully handled database coordination, dependency management, and service orchestration.

IV. Conclusion

The findings of this project are centered on finding abnormalities in the Mars Express Orbiter (MEX) telemetry data, with a particular emphasis on utilizing three years' worth of data to forecast power usage. The proposed method distinguishes between three primary types of anomalies: natural phenomena (such solar flares), unexplained events, and operational problems. The anomaly identification model makes use of a number of machine learning techniques. Among them, baseline models like random forest and linear regression underperformed against ensemble models like XGBoost and Extra Trees. XGBoost outperformed Extra Trees, which performed somewhat worse, with an average root mean square error (RMSE) of 0.0447 and a R^2 score of 0.2028 across various power lines. The system used a combination of Gated Recurrent Units (GRU) and Extreme Value Theory (EVT) for anomaly detection in order to find unusual behaviors in the telemetry data. The Generalized Pareto Distribution was utilized by the algorithm to identify uncommon, significant abnormalities, and it proved to be especially successful in identifying extreme deviations. The algorithm showed promise in spotting trends in power usage and pinpointing anomalies, but a thorough assessment of its efficacy is still difficult in the absence of labeled data for comparison. The strategy is optimized for satellite data, but the results show that it can adjust dynamically to changes in operations. The techniques also show promise for wider applications in other complex systems, such as cloud computing.

References

[1] Breskvar, Martin & Kocev, Dragi & Levatić, Jurica & Osojnik, Aljaž & Petković, Matej & Simidjievski, Nikola & Ženko, Bernard & Boumghar, Redouane & Lucas, Luke. (2017). Predicting Thermal Power Consumption of the Mars Express Satellite with Machine Learning. 10.1109/SMC-IT.2017.22.

[2] Bozhidar Stevanoski, Dragi Kocev, Aljaž Osojnik, Ivica Dimitrovski, Sašo Džeroski, Data stream mining for predicting the thermal power consumption of the Mars Express spacecraft, Acta Astronautica, Volume 210, 2023, Pages 411-427, ISSN 0094-5765, https://doi.org/10.1016/j.actaastro.2023.04.037. (https://www.sciencedirect.com/science/article/pii/S0094576523002102)

[3] Jin, Xin & Wang, Hui & Jin, Zhong. (2023). Anomaly detection of satellite telemetry data based on extended dominant sets clustering. Journal of Physics: Conference Series. 2489. 012036. 10.1088/1742-6596/2489/1/012036.