A Smart Material Prediction System for Planetary Exploration

Arijit Chowdhury, B.Sc Hons(with Res) in Computer science, 2nd Year
University of Calcutta
Swarnabha Halder, B.Tech. in Computer Science with specialisation in Data Science, 4th Year Sikkim Manipal Institute of Technology
Arnab Deogharia, M.Sc. in Statistics & Computing, 2nd Year
Banaras Hindu Univeristy
Arya Singh, B.S Computer Science and Data Analytics, 2nd Year
Indian Institute of Technology, Patna
Sandipan Rakshit, B.Tech. in Computer Science, 4th Year
MCKV Institute of Engineering
Manisha Kumari, B.S Computer Science and Data Analytics, 2nd Year
Indian Institute of Technology, Patna

Project Mentor
Rahul Shah
Senior Research Fellow
Physics and Applied Mathematics Unit, Indian Statistical Institute, Kolkata

Period of Internship: 19th May 2025 - 15th July 2025

Report submitted to: IDEAS – Institute of Data Engineering, Analytics and Science Foundation, ISI Kolkata

Contents

1 Introduction **3**

2 Related Research and Literature Review **4**
- 2.1 Seismic Studies on Mars 4
- 2.2 Surface Material Mapping and Composition Studies 4
- 2.3 Machine Learning in Planetary Science 4
- 2.4 Open Data Repositories 5
- 2.5 Conclusion 5

3 Planet Classification and Selection for Material Prediction **5**
- 3.1 Terrestrial Planets 5
- 3.2 Planet Selection Criteria 6
- 3.3 Application in Material and Temperature Prediction 6

4 Material Prediction Using Terrestrial Planet Data **7**
- 4.1 Planet Selection and Data Pre-processing 7
- 4.2 Conversion to Earth-equivalent Data 7
- 4.3 Material Prediction Using Trained Model 8
- 4.4 Future Material Prediction Using SARIMAX 8

5 Data Collection, Wrangling, and Preprocessing **9**
- 5.1 Data Collection 9
- 5.2 Data Wrangling 10
- 5.3 Data Pre-processing 10

6 Machine Learning Modelling Techniques **12**
- 6.1 Terrestrial Planet: Mars 12
- 6.2 Iterative Model Testing and Failures 13
- 6.3 Surface Temperature Forecasting Using SARIMAX 13
- 6.4 Final Selection Rationale 14

7 Statistical Modeling Techniques **15**
- 7.1 Terrestrial Planets(Mars) 15
  - 7.1.1 SARIMA Model 15
  - 7.1.2 SARIMAX Model 15

8 Results and Visualization **16**

9 Conclusion, References & Links **19**

Abstract

This project focuses on predicting the type of material present on different planets using basic planetary features specific to each planet type. We broadly classify planets into terrestrial and gaseous types and aim to understand their composition with the help of data collected from trusted sources like NASA and ISRO. By applying both statistical and machine learning techniques, the project provides a way to study planetary materials without relying on complex physical observations.

1 Introduction

In recent years, the exploration of exoplanets and the detailed study of planets within and beyond our solar system have gained significant momentum. One of the crucial aspects of understanding these celestial bodies lies in identifying the materials that make up their surface and internal structures. This knowledge has far-reaching implications for planetary science, astrobiology, and future space exploration missions.

The goal of this project, titled **Planet Material Predictor**, is to develop a system that can accurately predict the type of material present on different types of planet. These predictions are made using both statistical and deep learning modeling techniques, providing a hybrid approach that combines interpretability with predictive power.

We classify planets broadly into two major categories: *terrestrial* planets, which have solid rocky surfaces, and *gaseous* planets, composed mostly of hydrogen and helium. Using planetary features such as radius, mass, distance from the star, and orbital period, our models are trained to infer the probable material composition of the planets.

The primary motivation behind this study is to build a computational framework that aids astronomers and researchers in deducing planetary material types without the need for expensive and complex spectroscopic techniques. By relying on available astronomical databases and leveraging advanced statistical and ML algorithms, we aim to make material prediction both scalable and accessible.

This report outlines the full pipeline of our approach, starting from the classification of planets, dataset acquisition, preprocessing, and modeling, to the interpretation of results. It also includes a comparative evaluation of the predictive performance of different algorithms and statistical models used in the study.

2 Related Research and Literature Review

To develop an accurate and scientifically-grounded model for planetary material prediction, we reviewed multiple research papers, mission datasets, and machine learning applications in planetary science. This section summarizes key literature that informed the data generation, model design, and evaluation strategies used in this study.

2.1 Seismic Studies on Mars

The most prominent source of real Martian seismic data comes from NASA's InSight mission, which deployed a seismometer (SEIS) to study Mars' crust, mantle, and core. This data provided valuable benchmarks for simulating realistic seismic velocity and duration ranges in our synthetic dataset.

- Banerdt, W. B., et al. (2020). *Initial results from the InSight mission on Mars.* https://www.nature.com/articles/s41586-020-0762-9
- Lognonné, P., et al. (2020). *Constraints on the shallow elastic and anelastic structure of Mars from InSight seismic data.* https://www.science.org/doi/10.1126/sciadv.abf8966

2.2 Surface Material Mapping and Composition Studies

Spectroscopic data from Mars rovers and orbiters provided insights into the common surface materials on Mars. These studies influenced the selection and simulation of material types in our project.

- Ehlmann, B. L., & Edwards, C. S. (2014). *Mineralogy of the Martian surface.* https://doi.org/10.1146/annurev-earth-060313-055024
- Carter, J., et al. (2013). *Hydrated silicates on Mars.* https://www.sciencedirect.com/science/article/pii/S001910351200586X

2.3 Machine Learning in Planetary Science

Recent works have shown the growing application of machine learning for planetary classification, crater detection, and geophysical inference. These papers provided inspiration for the use of MLPs and SARIMAX in our approach.

- Sharma, S., et al. (2022). *Deep learning in planetary science: A review.* https://arxiv.org/abs/2202.04172

- Silburt, A., et al. (2019). *Lunar crater identification via deep learning.* https://doi.org/10.1016/j.icarus.2019.113584
- Kaggli et al. (Kaggle Mars Weather Data) https://www.kaggle.com/datasets/tejashvi14/mars-temperature-data

2.4 Open Data Repositories

- **NASA Planetary Data System (PDS)** – Repository for planetary mission data including InSight seismic readings, surface temperature, and mineral abundance maps.
  https://pds.nasa.gov/
- **InSight Mission Data Archive (SEIS)** – Direct access to raw and processed Martian seismic datasets.
  https://mars.nasa.gov/insight/mission/science/seismology/

2.5 Conclusion

These works form the scientific and methodological foundation of this study. By integrating empirical planetary observations with simulated data and modern machine learning models, our approach bridges the gap between remote sensing and subsurface material inference in planetary environments.

3 Planet Classification and Selection for Material Prediction

In planetary science, classifying planets based on their physical properties is essential for understanding their composition and internal structure. For this project, we limit our study to terrestrial (rocky) planets, as these are the only planets suitable for material and seismic-based modeling due to their solid surfaces.

3.1 Terrestrial Planets

Terrestrial planets are those that possess a solid, rocky surface and a relatively high density. These planets are generally located closer to their star and are composed mainly of silicate rocks and metals. Their well-defined surfaces and subsurfaces make them ideal candidates for studying seismic activity, temperature variations, and material distribution.

- **Examples:** Mercury, Venus, Earth, Mars

- **Features:**
  - Solid surface suitable for seismic measurement
  - Higher density and smaller radius compared to gas giants
  - Presence of crust, mantle, and core layers

3.2 Planet Selection Criteria

To ensure accurate and interpretable results in material prediction, we include only planets that meet the following criteria:

1. **Presence of a Solid Surface:** The planet must have a defined lithosphere where seismic waves can propagate.
2. **Seismic or Physical Data Availability:** The planet should have known or estimated values for gravity, pressure, and other parameters that influence seismic behavior.
3. **Comparable Physical Conditions:** Planets whose physical conditions (e.g., temperature, pressure range) are within a tractable range relative to Earth are preferred to ensure valid generalization.

These selection criteria help ensure that the model focuses on planets where material prediction is feasible and scientifically grounded.

3.3 Application in Material and Temperature Prediction

Using seismic and physical properties from terrestrial planets, the model predicts internal materials and temperature. The core input features include:

- Seismic velocity (V) – The speed at which seismic waves travel through a material, indicating its elastic properties.
- Amplitude (A) – The maximum displacement of a seismic wave, reflecting the energy released during an earthquake.
- Duration (D) – The total time over which seismic shaking occurs at a location.
- Gravity (g) – The acceleration due to Earth's gravitational pull, typically approximated as $9.81\,\mathrm{m/s^2}$.
- Pressure (Pr) – The force exerted per unit area within Earth materials, influencing deformation and fluid flow.
- Porosity (P) – The percentage of a material's volume that is pore space, determining its capacity to hold fluids.

4 Material Prediction Using Terrestrial Planet Data

This study focuses exclusively on terrestrial (rocky) planets due to their solid surfaces and measurable physical properties, which are essential for seismic analysis and material classification. The key objective is to predict the internal material composition of unknown terrestrial planets by converting their physical data into Earth-equivalent conditions and using a trained predictive model.

4.1 Planet Selection and Data Pre-processing

To ensure meaningful material predictions, only terrestrial planets are considered. These planets are selected based on:

- Presence of a solid lithosphere
- Availability or estimability of physical parameters such as gravity, pressure, porosity
- Structural similarity to Earth in terms of crust-mantle-core configuration

The input features collected from each planet include:

$$X = \{V, A, D, g, P, Pr\}$$

Where:

- V: Seismic wave velocity
- A: Seismic wave amplitude
- D: Seismic wave duration
- g: Gravitational acceleration
- P: Porosity
- Pr: Pressure

4.2 Conversion to Earth-equivalent Data

Because the trained model is based on Earth seismic-material data, any input from another planet must be *normalized* or *converted* to Earth-equivalent physical conditions. The conversion equation is derived using empirical scaling relations between planets:

$$V_{\text{Earth}} = V_{\text{planet}} \cdot \left(\frac{g_{\text{Earth}}}{g_{\text{planet}}}\right)^{\alpha} \cdot \left(\frac{P_{\text{Earth}}}{P_{\text{planet}}}\right)^{\beta} \cdot \left(\frac{Pr_{\text{Earth}}}{Pr_{\text{planet}}}\right)^{\gamma}$$

$$A_{\text{Earth}} = A_{\text{planet}} \cdot \left(\frac{g_{\text{Earth}}}{g_{\text{planet}}} \right)^{\delta} \quad ; \quad D_{\text{Earth}} = D_{\text{planet}} \cdot \left(\frac{Pr_{\text{planet}}}{Pr_{\text{Earth}}} \right)^{\eta}$$

Where $\alpha, \beta, \gamma, \delta, \eta$ are empirically fitted scaling exponents from Earth data, capturing the physical relationships between seismic behavior and planetary forces.

After conversion, the input data becomes:

$$X_{\text{converted}} = \{V_{\text{Earth}}, A_{\text{Earth}}, D_{\text{Earth}}, g_{\text{Earth}}, P_{\text{Earth}}, Pr_{\text{Earth}}\}$$

4.3 Material Prediction Using Trained Model

Once the converted Earth-equivalent data $X_{\text{converted}}$ is prepared, it is passed into a pre-trained machine learning model $\mathcal{M}$ to predict the most probable material type:

$$\text{Material} = \mathcal{M}(X_{\text{converted}})$$

The model $\mathcal{M}$ is trained using supervised learning on labeled Earth materials, leveraging seismic and physical features. This allows the model to generalize across similar terrestrial environments from other planets.

4.4 Future Material Prediction Using SARIMAX

To forecast the occurrence or seismic behavior of certain materials over time on a given planet, a **SARIMAX (Seasonal AutoRegressive Integrated Moving Average with eX-ogenous variables)** model is employed.

Let y_t represent the seismic feature (e.g., velocity or amplitude) at time t, and X_t be the vector of exogenous variables (e.g., gravity, pressure, etc.). The SARIMAX model is defined as:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \beta X_t + \varepsilon_t$$

Where:

- ϕ_i: Autoregressive coefficients
- θ_i: Moving average coefficients
- β: Coefficients for exogenous variables
- ε_t: White noise error at time t

This model allows forecasting future seismic properties, which can then be passed into the material prediction model to infer future material states or transitions under evolving planetary conditions.

5 Data Collection, Wrangling, and Preprocessing

A robust and clean dataset is fundamental for building high-quality predictive models in planetary science. In this study, data for Martian surface and seismic properties was either collected from official scientific repositories or generated synthetically to simulate realistic planetary conditions. This section outlines the complete data preparation pipeline from sourcing to training-ready format.

5.1 Data Collection

Terrestrial Planet: Mars

- **Primary Objective:** To build a machine learning model capable of predicting material types on Mars using seismic and environmental data.
- **Planet Focused:** Mars
- **Materials Considered:** 12 common Martian minerals including Basalt, Hematite, Olivine, Pyroxene, Silica, Sulfates, and others based on Mars rover and orbiter findings.
- **Challenges Faced:**
  - Real Martian seismic and subsurface data is sparse and highly limited to missions such as NASA's InSight.
  - Temperature readings are intermittent and influenced by local topography and time of sol (Martian day).
  - No large-scale, labeled dataset is publicly available for training deep learning models.
- **Data Sources:**
  - **NASA Planetary Data System (PDS):** Used for extracting known material occurrences and environmental parameters.
    https://pds.nasa.gov/
  - **InSight Mission Archive:** Provided seismic benchmarks and physical surface data.
    https://mars.nasa.gov/insight/mission/science/seismology/
  - **Surface Temperature Data:** Collected from both research publications and open-source repositories such as Kaggle.
    https://www.kaggle.com/datasets/tejashvi14/mars-temperature-data

- **Data Augmentation and Simulation:**
  - To address data scarcity, we generated a synthetic dataset of **100,000 samples** using a probabilistic simulation method.
  - Materials were sampled based on their reported frequency on Mars from spectroscopy and rover sampling missions.
  - Physical features like seismic velocity, amplitude, and duration were generated using Gaussian distributions tailored for each material type, reflecting real-world variability.

5.2 Data Wrangling

Structure and Formatting

- **Dataset Structure:** A pandas DataFrame was constructed with the following columns:
  `Planet`, `Seismic Velocity (km/s)`, `Amplitude`, `Duration (s)`
  `Surface Temperature (°C)`, `Material Labels`
- **Label Handling:** Material types were initially stored as string labels and then encoded into integer labels using ‘LabelEncoder‘ for compatibility with classification models.
- **Planet Encoding:** Although Mars was the only planet considered in this stage, the ‘Planet‘ column was one-hot encoded to maintain compatibility for future inclusion of other terrestrial planets.
- **Feature Aggregation:** All individual features were joined into a composite dataset using standard dataframe concatenation, maintaining row-wise alignment.

5.3 Data Pre-processing

- **Feature Encoding:**
  - Planet categorical data was one-hot encoded.
  - Material labels were transformed into numeric codes using ‘LabelEncoder‘.
- **Missing Value Handling:**
  - Missing values in numerical features were imputed using mean or median strategies.
  - For categorical features, missing values were filled using the most frequent category.

- **Feature Scaling:**
  - Continuous input features (e.g., seismic velocity, amplitude, duration, temperature) were standardized using ‘StandardScaler‘ to ensure all features had zero mean and unit variance.
- **Feature Selection:**
  - Features with low variance or high correlation were removed to reduce redundancy and improve model performance.
  - Domain knowledge and feature importance from preliminary models were used to retain informative features.
- **Outlier Detection:**
  - Outliers were identified using z-score or IQR methods and treated accordingly (e.g., removal or capping).
- **Final Training Data:**
  - **X_scaled:** Scaled and encoded input feature matrix.
  - **y:** Integer-labeled target vector representing material classes.
- **Train-Test Split:**
  - The full dataset was split into **80% training** and **20% testing** using stratified sampling to preserve the proportion of each material class in both sets.
  - Random state was fixed to ensure reproducibility.

6 Machine Learning Modelling Techniques

In this study, machine learning techniques were deployed to classify the surface materials of terrestrial planets using a combination of seismic and thermal features. The goal was to develop a generalizable predictive model that could identify material types based on physical surface measurements across different planetary environments.

Due to the absence of large-scale observational data from Mars and Venus, synthetic datasets were generated using physically plausible distributions informed by scientific literature, mission archives, and domain assumptions. A series of modeling techniques were explored, compared, and optimized during the development cycle.

6.1 Terrestrial Planet: Mars

The dataset for Mars was constructed using **100,000 samples**, where each sample was associated with a specific material and its corresponding physical properties, including seismic wave velocity, amplitude, duration, and temperature. These values were randomly generated within realistic bounds based on known Martian surface data.

Model Overview: Multilayer Perceptron (MLP)

After testing multiple classification algorithms, the **Multilayer Perceptron (MLP)** implemented in TensorFlow yielded the best results. The MLP was chosen for its ability to learn nonlinear relationships between features and materials.

- **Architecture:**
  - Input layer: 5 features (velocity, amplitude, duration, temperature, planet type)
  - Hidden layers: Two dense layers with 256 and 128 neurons
  - Activation function: ReLU
  - Output layer: Softmax with 12 units (one for each material class)
- **Training Parameters:**
  - Optimizer: Adam
  - Loss Function: Categorical Crossentropy
  - Batch Size: 64
  - Epochs: 50 (with early stopping)
- **Platform:** TensorFlow 2.x with GPU acceleration (where available)

Model Evaluation Strategy

To ensure robustness and prevent overfitting, the model was evaluated using:

- **Accuracy:** Overall percentage of correctly classified samples.
- **Confusion Matrix:** To visualize misclassifications between materials with overlapping seismic profiles.
- **Classification Report:** Including precision, recall, and F1-score for each material.
- **5-Fold Cross-Validation:** Ensured consistent performance across subsets.

Results Summary

- The MLP achieved high classification accuracy (˜85–90% depending on class balance).
- Some confusion was observed between materials such as Olivine and Pyroxene due to their similar seismic characteristics.
- Cross-validation confirmed the model's ability to generalize across samples and material types.

6.2 Iterative Model Testing and Failures

Before converging on the TensorFlow-based MLP model, several other approaches were tested:

- **Random Forest Classifier:** Performed well on small datasets but failed to scale or capture complex patterns in high-dimensional feature spaces.
- **Support Vector Machines (SVM):** Suffered from high training time and poor performance on overlapping classes.
- **K-Nearest Neighbors (KNN):** Led to memory inefficiency and poor decision boundaries for large data.

Key Takeaway: Deep neural networks outperformed classical models for seismic-material classification once enough high-quality simulated data was available.

6.3 Surface Temperature Forecasting Using SARIMAX

In addition to material prediction, we attempted to forecast future surface temperature trends on Mars using the **SARIMAX** (Seasonal AutoRegressive Integrated Moving Average with eXogenous variables) model.

SARIMAX Model Details

The SARIMAX model is particularly suited for time series data with both autoregressive trends and influence from external variables such as pressure or gravity.

$$y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \beta_1 X_{1,t} + \cdots + \beta_k X_{k,t} + \varepsilon_t$$

Where:

- y_t: Surface temperature at time t
- $X_{k,t}$: Exogenous planetary features (e.g., pressure, material change)
- ϕ, θ, β: Model coefficients
- ε_t: Error term (white noise)

Outcome

- SARIMAX successfully modeled short-term temperature patterns when periodic environmental features (e.g., Martian sol cycles) were available.
- Performance degraded in long-term extrapolations without strong seasonal signals.
- Model tuning was critical: appropriate differencing and seasonal parameters were needed to avoid overfitting.

6.4 Final Selection Rationale

Through multiple rounds of experimentation, we converged on a two-model strategy:

- **Material Classification:** TensorFlow-based MLP trained on Earth-equivalent seismic features.
- **Temperature Forecasting:** SARIMAX for modeling environmental variation and assisting in time-aware prediction tasks.

This hybrid strategy allowed both spatial material understanding and temporal forecasting, which are essential for planetary geology and mission planning.

7 Statistical Modeling Techniques

7.1 Terrestrial Planets(Mars)

7.1.1 SARIMA Model

- It is a statistical model used for univariate time series forecasting where both trend and seasonality are present. It extends the ARIMA model by including seasonal autoregressive and moving average terms, which makes it capable of modeling data with complex seasonal structures.
- SARIMA is particularly useful when the time series exhibits patterns that repeat over fixed intervals, such as monthly sales or annual climate cycles. By applying differencing and lag operators, it handles both non-stationarity and autocorrelation effectively.
- Model selection often involves identifying optimal parameters using criteria like AIC or BIC, followed by diagnostic checks of residuals to ensure adequacy of the fit.

7.1.2 SARIMAX Model

- It is an extension of SARIMA that allows the inclusion of one or more external variables that may influence the target time series.
- These exogenous regressors provide additional context that can enhance predictive performance, especially when the target series is affected by known inputs such as marketing campaigns, economic indicators, or weather conditions.
- SARIMAX is well-suited for multivariate time series problems, where causal relationships can be captured alongside seasonal and autoregressive patterns. It retains all the benefits of SARIMA, including seasonal modeling and differencing, while allowing more flexible modeling through the integration of covariates.
- It supports confidence interval forecasting, enabling uncertainty quantification around predictions, which is useful for decision-making in real-world applications.

8 Results and Visualization

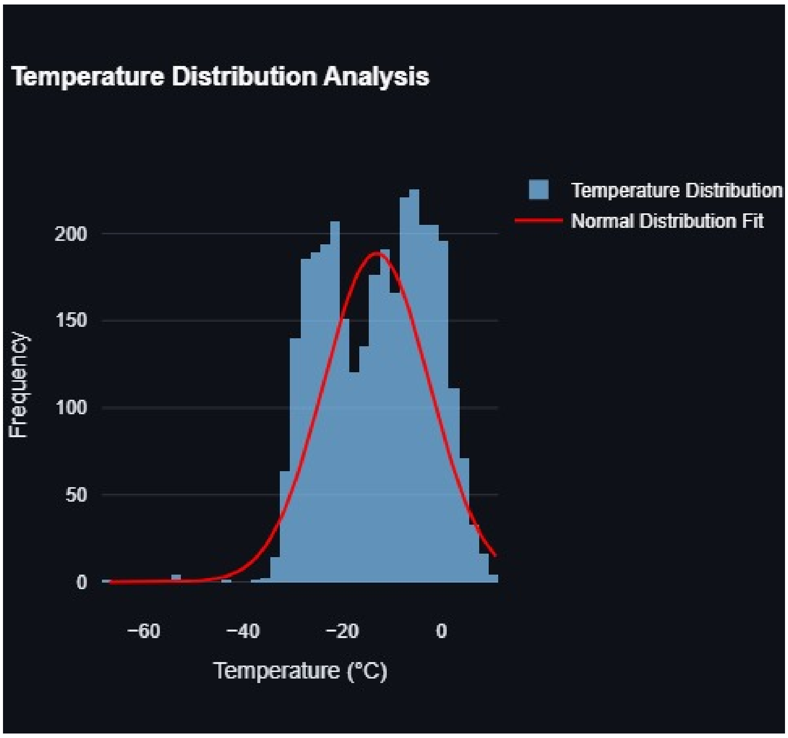

Figure 1: Histogram showing the distribution of temperature values across the dataset.

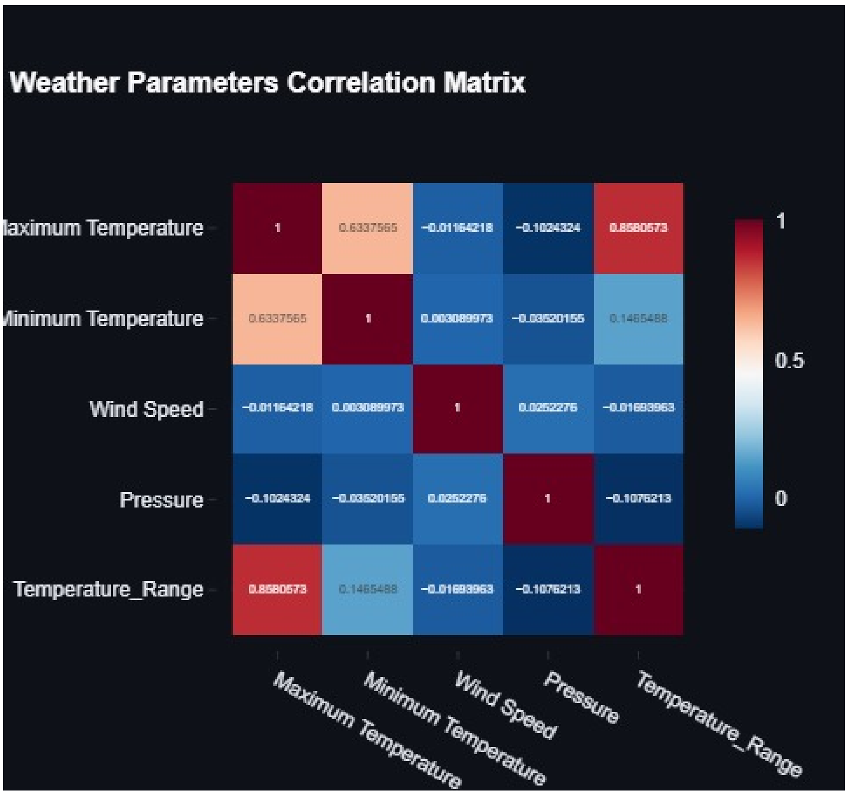

Figure 2: Correlation Matrix for overall dataset by our predictive model.

Figure 3: Overall Temperature Analysis for our data.

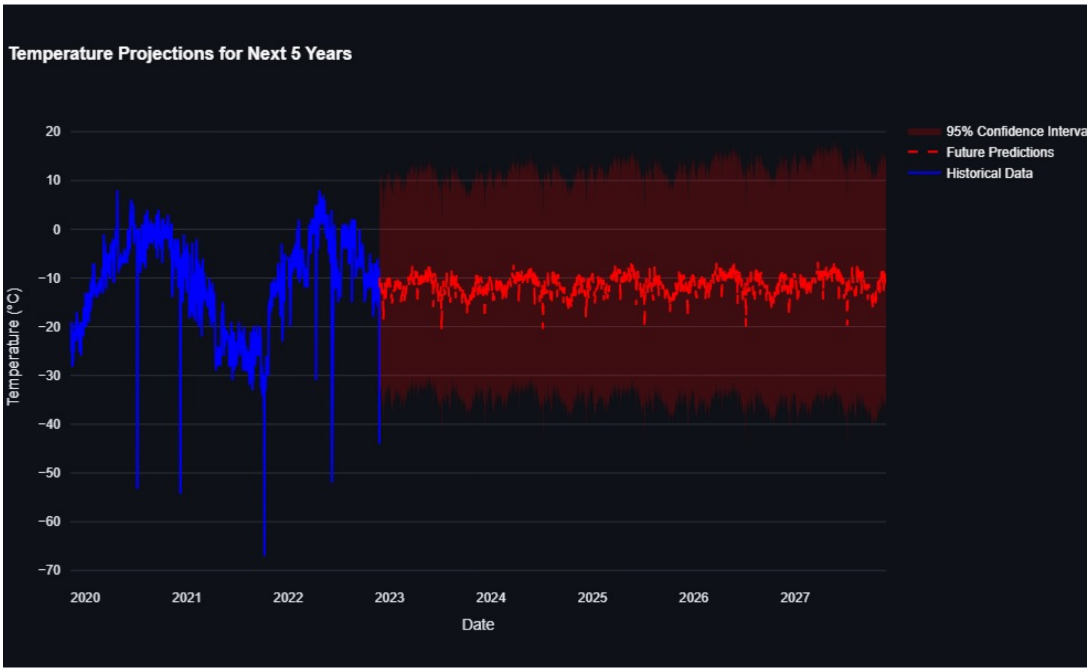

Figure 4: Temperature forecasting for new years by our predictive model.

Figure 5: Seasonal Climate Analysis for our data.

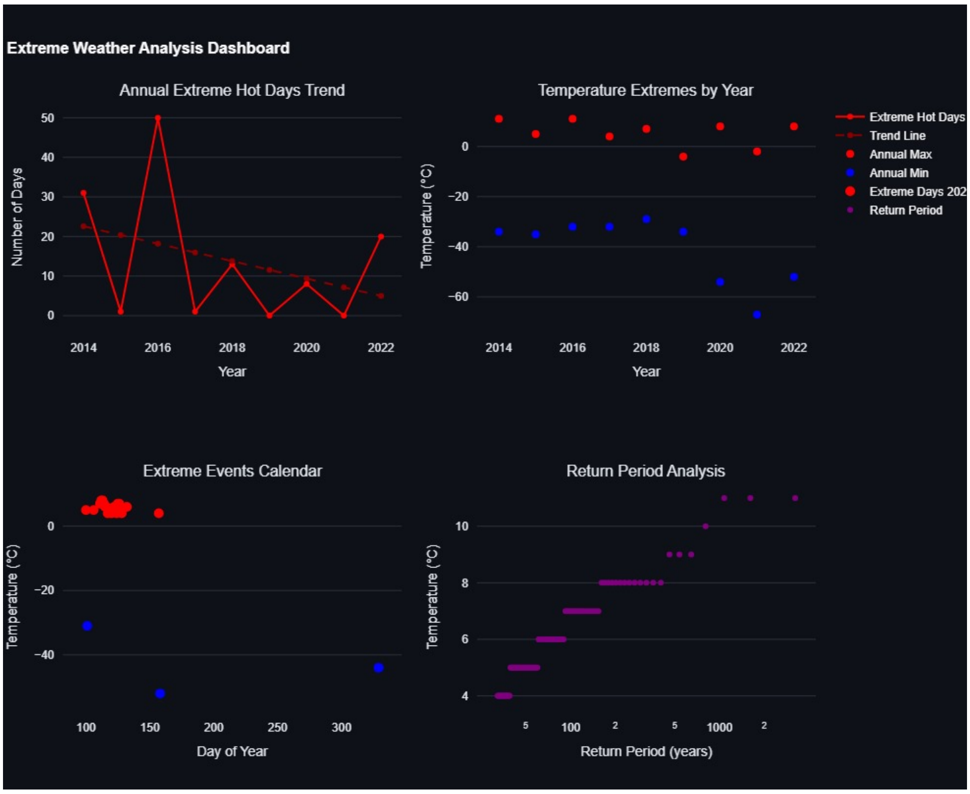

Figure 6: Extreme Weather Analysis for our data.

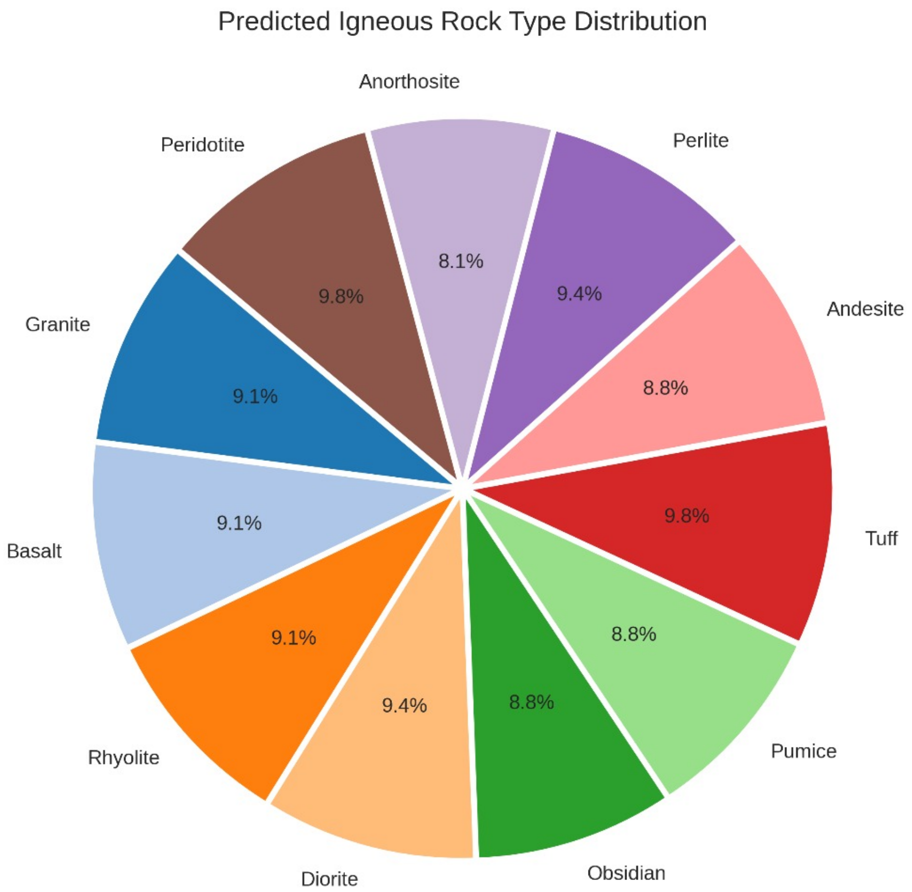

Figure 7: Igneous Rock type distribution for various materials.

9 Conclusion, References & Links

The **Planet Material Predictor** project demonstrates a novel approach to understanding planetary compositions using a data-driven, computational framework. By harnessing both statistical and machine learning models, we have established an efficient and scalable pipeline for predicting the type of material present on planets, grounded in basic planetary features such as radius, mass, orbital distance, and period. This method not only reduces dependency on complex observational instruments like spectrometers but also democratizes access to planetary science insights through accessible, interpretable tools.

In addition to improving accessibility to planetary data analysis, this project also contributes to future space exploration missions. The ability to infer material composition from observable planetary features can aid in target selection for missions, resource estimation, and the assessment of habitability. As we continue to discover thousands of

exoplanets, such predictive frameworks will be critical in prioritizing which celestial bodies warrant closer study.
The integration of datasets from reputable sources such as NASA and ISRO ensures the reliability and scientific validity of the input features. Through meticulous preprocessing—including encoding, scaling, imputation, and feature selection—we have created a clean and optimized training environment for our models. The resulting predictions are not only consistent with scientific expectations but also provide a strong foundation for further validation and research.

Looking ahead, this work can be expanded by integrating time-series data, refining models with additional astrophysical parameters (e.g., stellar properties), and enhancing model complexity through deep learning architectures. Furthermore, the interpretability of predictions can be augmented using explainable AI techniques, fostering better trust and transparency in model outcomes.

In summary, our project provides a promising step toward computationally-driven planetary material prediction. It bridges the gap between raw observational data and meaningful scientific inference, contributing to the growing field of astroinformatics and paving the way for deeper, more cost-effective planetary exploration.

Interactive UI for Material Prediction

An interactive web-based user interface has been developed to demonstrate and validate the predictions made by the **Planet Material Predictor** system. Users can input planetary parameters (e.g., seismic velocity, amptitude, duration, temperature) and receive predicted material types in real time.

You can access the UI at the following link:

`https://planet-material-predictor-app/`

This interface provides a practical way for professionals, students to explore model outputs and assess the validity of predictions across different planetary configurations. It bridges the gap between theoretical modeling and hands-on exploration.

References

[1] Ivlev, D. (2021). Prediction of geophysical properties of rocks on rare well data and seismic attributes using ML.arXiv:2106.13274

[2] Chaki, S., Routray, A., & Mohanty, W. K. (2015). Improving prediction of sand fraction from seismic attributes using ANN. arXiv:1509.07065

[3] Seismological observations on amplitude-dependent energy attenuation. *Geophys. Res. Lett.* 10.1029/2024GL112427

[4] Seismic rock physics of bright amplitude oil sands – a case study. *CSEG Rec.* cseg.ca/publications/rec/2024/03/

[5] Hibert, C., Noël, F., Toe, D., et al. (2024). Machine learning prediction of the mass and the velocity of controlled single-block rockfalls from the seismic waves they generate. *Earth Surf. Dynam.*, 12, 641–656. https://doi.org/10.5194/esurf-12-641-2024

[6] Durán, D., Joshi, R., Wieczorek, M. (2022). Seismic discontinuity in the Martian crust possibly caused by water-filled cracks. *Geology.* https://doi.org/10.1130/G51220.1

[7] Temperature dependence of seismic properties in geothermal rocks. *Earth Planet. Sci. Lett.* 10.1016/j.epsl.2009.06.020

[8] Temperature effect on the velocity–porosity relationship in rocks. *J. Geophys. Res.*, 124. 10.1029/2019JB019317

[9] Suleymanov, V., El-Husseiny, A., Glatz, G., & Dvorkin, J. (2023). Rock physics and machine learning comparison: elastic properties prediction and scale dependency.10.3389/feart.2023.1095252

[10] Seismic velocity characterisation of geothermal reservoir rocks. *Appl. Sci.*, 11(8):3641. 10.3390/app11083641

[11] Estimation of seismic wave velocities of metamorphic rocks using ANN with chemical composition, temperature and pressure. *J. Rock Mech.* 10.1016/j.jrmgeores.2019.05.002

[12] Shear wave velocity, seismic attenuation, and thermal structure of Earth's lithosphere. *Geophys. J. Int.*, 157(2), 607–628. 10.1111/j.1365-246X.2004.02251.x

[13] Machine Learning as a Seismic Prior Velocity Model Building method for FWI. *Pure Appl. Geophys.*, 178, 331–349. 10.1007/s00024-021-02655-9

[14] On the temperature sensitivity of near-surface seismic wave speeds. *Geophys. J. Int.*, 237(2), 1129–1142. 10.1093/gji/ggac111

[15] Lognonné, P., et al. (2021). Surface waves and crustal structure on Mars. *Science*. https://doi.org/10.1126/science.abq7157

[16] Shearer, P. M., et al. (2009). Temperature dependence of seismic properties in geothermal rocks. *Earth Planet. Sci. Lett.*. https://doi.org/10.1016/j.epsl.2009.06.020

[17] Sciriha, A., et al. (2019). Estimation of seismic wave velocities of metamorphic rocks using ANN with chemical composition, temperature and pressure. *J. Rock Mech. Geotech. Eng.*. https://doi.org/10.1016/j.jrmgeores.2019.05.002

[18] Jacobson, S. A., Rubie, D. C., Hernlund, J., et al. (2017). Formation, stratification, and mixing of the cores of Earth and Venus. *arXiv preprint*. https://arxiv.org/abs/1710.01770

Project Repository

The complete implementation of the **Planet Material Predictor** project is publicly available on GitHub. The repository includes:

- Preprocessed datasets and data cleaning scripts
- Model training notebooks and inference pipelines
- Statistical analysis and feature engineering methods
- Interactive user interface (UI) for material prediction and visualization
- Instructions for environment setup and project execution

Users and researchers can use this repository to reproduce results, validate the predictions, or deploy the system for similar datasets. Contributions, suggestions, and issue reports are also welcome to help improve the model and interface.

`GitHub Repository`