COSMIC CLASSIFIER

THEME

The Galactic Classification Challenge (GCC) invites participants to step into the shoes of humanity's saviors, tasked with decoding Dr. Klaus Reinhardt's final transmission. Using advanced machine learning techniques, participants must classify planets based on their survival potential and resource availability.

The fate of humanity is in your hands—will you rise to the challenge?

Himanshu Kumar
Head Events
Cognizance, IIT Roorkee

Prof. Tarun Sharma
Co-Faculty Coordinator
Cognizance, IIT Roorkee

Prof. Neetesh Kumar
Co-Faculty Coordinator
Cognizance, IIT Roorkee

Prof. Siddhartha Khare
Faculty Coordinator
Cognizance, IIT Roorkee

Prof. M.V. Sunil Krishna
Associate Dean of Students' Welfare, IIT Roorkee

Prof. Barjeev Tyagi
Dean of Students' Welfare
IIT Roorkee

INTRODUCTION

In 2547, Dr. Klaus Reinhardt, a renowned scientist and explorer, made a groundbreaking discovery while traversing a distant galaxy. After years of experimentation, he successfully categorized planets based on their potential for human survival and use. However, as fate would have it, his ship was caught in the gravitational pull of a black hole. In his final moments, he transmitted the data to Earth before being consumed by the singularity.

Unfortunately, the immense gravitational forces of the black hole caused interference during transmission, permanently removing some part of data while adding noise to some parts. The names of the categories, though not meaning anything, are standard use in the scientific community. Humanity is relying on you to classify planets accurately to secure our survival.

PROBLEM STATEMENT

Dataset Details

Input Features (10 attributes of planets):

1. Atmospheric Density: Measure of the thickness of the planet's atmosphere (in kg/m^3).

2. Surface Temperature: Average surface temperature of the planet (in Kelvin).
3. Gravity: Surface gravity of the planet (in m/s^2).
4. Water Content: Percentage of the planet's surface covered by water (0-100%).
5. Mineral Abundance: Index representing the availability of valuable minerals (scale 0-1).
6. Orbital Period: Time the planet takes to orbit its star (in Earth days).
7. Proximity to Star: Distance from the planet to its star (in AU).
8. Magnetic Field Strength: Measure of the planet's magnetic field (in Tesla).
9. Radiation Levels: Average radiation levels on the planet's surface (in Sieverts/year).
10. Atmospheric Composition Index: Index measuring the suitability of the atmosphere for human life (scale 0-1).

Output Classes (10 planet types in German):

The transmitted classifications are as follows, but their meanings remain unclear.

1. Bewohnbar
2. Terraformierbar
3. Rohstoffreich
4. Wissenschaftlich

5. Gasriese
6. Wüstenplanet
7. Eiswelt
8. Toxischetmosäre
9. Hohestrahlung
10. Toterahswelt

Your mission is to map planets to these categories using the given attributes, even though we do not fully understand each classification.

GOAL

Participants must develop a classification model that accurately predicts the type of a planet based on its attributes. Due to interference during data transmission, the dataset may contain some noise, adding an extra layer of challenge. The model's predictions will be evaluated using accuracy as the primary metric.

COMPETITION RULES

1. Dataset: The dataset is provided in two files: train.csv (60,000 samples with labels) and test.csv (10,000 samples without labels).
2. Submissions: Participants can submit predictions on the test set in a submission.csv file, containing two columns: Planet_ID and Predicted_Class.

3. Evaluation: Submissions will be evaluated based on accuracy, calculated as the percentage of correctly classified planets in the test set.

4. Fair Play: External datasets and pre-trained models are not allowed. All work must be done within the scope of the provided dataset.

ROUND-1 (CODE SUBMISSION)

The problem sets (training dataset) is given here : cosmic classifier data

Participants have to submit the code in a Python notebook by **5th March 2025**.

Format of the file: Python notebook with extension (.ipynb)

Submission should be done on our website.

This round serves as a qualifier for advancement to Round 2

ROUND-2 (OFFLINE TESTING)

- This round will be held in offline mode at IIT Roorkee from **21-23 March 2025**.
- The participants will be given the test datasets on **21 March 2025** and need to process the output by running the already submitted code at the time of the event.

● The participants have to build an illustrated blog of code, i.e., it should be a python notebook with code and an explanation of the method. Also, it should contain prediction pipelines.

ABOUT THE DATA

The data will be stored in the train directory containing the CSV file with several parameters. The test data will be placed in the test_data directory containing a similar CSV file to the train data.
Test dataset will be made available to participants during the main fest at **IIT Roorkee (21 - 23 March 2024).**
There are some missing values of labels in the train and data indicated by large negative numbers that the participants must filter out as a data processing step.

REGISTRATION PROCEDURE

- The registration shall be done through the Cognizance website.
- Each Member needs to register on the website. This will generate a unique Cognizance Id, after email verification.
- The Team Leader (which you will select yourself) needs to login into the website with his username and password.

- After logging in to his Cogni ID, the team leader needs to select Cosmic Classifier from the dashboard in order to participate in this particular event.
- The Team Details include the following details of all team members:

  Name: **City:**

  Branch: **E-mail:**

  Institute Name: **Contact No:**
- Enter the Cognizance ID of the team leader and other team members.

GUIDELINES

- Eligibility: Students pursuing an Undergraduate/Master's Degree in any discipline.
- Team size: Maximum size of up to 5 members.
- No double-trouble: Only one entry would be acceptable. In the case of multiple entries, the first one would be taken for granted by default.
- Plagiarism will not be allowed.

TIMELINE

Competition Start: To Be Decided

Submission Deadline: To Be Decided

QUERIES

For any queries, you can contact the coordinator:

KRISH JAIN | +91 89491 32182

AARUSH MANGLECHA | +91 74108 87910

Or mail your queries with the **Subject "Query | Event Name | Cogni ID | Team Name" to** submission.cognizance@iitr.ac.in