Predictioneer (Problem Statement)

Introduction:

In the ongoing fight against infectious diseases, machine learning has become a powerful tool for predicting, understanding, and mitigating the spread of diseases. By analyzing large datasets with the help of machine learning models, researchers can identify disease hotspots, predict future outbreaks, and uncover patterns that may have otherwise gone unnoticed.

Machine learning analysis can provide insights into the epidemiological factors influencing disease transmission, including geography, demographics, and environmental conditions. This allows for early detection and timely intervention, minimizing the impact of outbreaks. Additionally, predictive models can help optimize resource allocation, ensure that healthcare services are directed to the most affected areas, and aid in the development of more effective vaccines and treatments.

Challenge:

Ebola, like many infectious diseases, poses significant public health risks. Early prediction of expected cases and deaths in various regions is critical to effective resource allocation, timely intervention, and containment efforts. In this challenge, participants are tasked with predicting the number of deaths and expected cases and case fatality ratio for multiple regions, based on available data.

Given the case fatality ratio (CFR) for a few regions, along with the number of deaths reported in some areas(death data are missing for some locations), and the latitude and longitude for tracking the geographical location of the regions, participants need to build a machine learning model that can forecast the number of deaths, confirmed cases and case fatality ratio for other unreported regions.

Potential Applications :

Machine learning analysis of diseases like Ebola outbreaks has significant potential in early detection, resource optimization, and public health planning. By analyzing existing data, predictive models can forecast future outbreaks and identify high-risk regions, allowing health authorities to take proactive measures such as targeted resource distribution, vaccination efforts, and tailored health interventions. This foresight ensures that areas with the highest predicted caseloads receive necessary medical supplies and support, optimizing the response to the outbreak.

Additionally, machine learning can enhance global health security by identifying potential hotspots and enabling timely international cooperation. By predicting the spread of the disease, countries and organizations like the WHO can implement measures such as travel restrictions, quarantine policies, and real-time monitoring to prevent the disease from spreading across borders. This contributes to a more efficient and coordinated global response, ultimately saving lives and reducing the overall impact of the outbreak.

The submission comprises of **three** parts:

Part I:
Predicted confirmed cases and deaths for all the geographical locations

Part II: Model postulation behind the predicted cases with solid correlations in a **MATLAB / Python / R** or Excel file.

Part III:
- The logic behind the code/model summarized in the form of bullet points or a flow-chart
- Any references used for constructing the model or ideating on the logic Discussion on
- how the model can be applied to new locations that have not been seen before.

(submit files as a word or pdf file)

Submission Requirements:

To participate in the competition, you will need to submit the following:
- A **Python/R/MATLAB/Excel** script that implements your model.
- A written report that describes your model and its performance.

Evaluation Criteria:

The points weightage for the three parts of the problem statement are as follows:

Part I (40 %)
The accuracy of the results will be assessed using Root Mean Square(RMS) values between the actual water volume and predicted water volume in each water body

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Predicted_i - Actual_i)^2}{N}}$$

Part II (30 %)
The marks for this part will be allowed for the format and effiiency of the workof the code

Part III (30 %)
This part will be graded on the overall presentation and the problem-solving approach. In this section, creativity will be put to the test. It will be desirable to predict the occurrence of crucial activities and processes in environment and their effect on water availability.

Participation Procedure:

- A group of **2 to 4** members are allowed per team.
- No participant can register in more than **one team**. In such an instance, both teams will be subjected to **disqualification**.

Certification Prizes:

- The top **5 teams** will be awarded cash prizes worth **20K INR** and the **top 10 teams** will receive a **certificate of merit.**
- The rest of the teams making a valid final submission and qualifying all the eligibility criteria will receive a **certificate of appreciation**.

NOTE: The final decision-making authority lies with team AZeotropy. Participants also consent to the free use of their submissions by the team AZeotropy. Proper attribution shall be given to the authors in case their submissions are published on the symposium

In case of any queries related to the Problem statement, participants can contact:

Shivani (Competitions Head)
shivani.azeotropy@gmail.com
Mobile No. +91-7668572809