INNOV8: The Space Saga

ARIES IITD and Eighfold.ai

Introduction

As the captain of an advanced space exploration fleet, you have ventured across the far reaches of the galaxy, encountering numerous alien species. On this thrilling expedition, you will face multiple challenges that require your leadership and problem-solving skills to ensure the success of your mission.

Note : *There are 2 parts to this problem statement. Even if you are only able to do 1, submit that part as both are independent and will be evaluated seperately.*

Part 1: Decoding and Classifying Alien Communications

Upon landing on the planet Xernia, your crew intercepts a flood of mysterious transmissions from various alien species. To achieve your mission objectives, you must analyze and decode these messages to identify which species sent each one.

You are provided with a dataset (`data.csv`) containing intercepted alien messages, where each message is associated with specific attributes such as the alien species, number of fingers, and whether they possess a tail. Your task is to classify these transmissions based on these attributes.

Dataset (Alien Messages Corpus)

The intercepted transmissions are structured as follows:

- **Message**: The alien communication in text form.

- **Species**: The species that sent the message.

- **Number of Fingers**: The typical number of fingers for the species.

- **Tail**: A yes/no field indicating whether the species has a tail.

Example Dataset

Message	Species	Number of Fingers	Tail
"Zograk meerta xalool zetrak."	Zeta Empire	4	Yes
"Blorp vee xunee ruffle kerf."	Blorbians	6	No
"Xylkor pelvi masten ro'xx."	Xylkan Tribe	3	Yes
"Wektre zon puugree zor rika."	Zeta Empire	5	Yes
"Kerflon blorp nexaru xevkerp."	Blorbians	11	Yes

Task

You talked to a lot of aliens this time and want to test how accurate your prediction skill has become. You thus log all your conversations, observe the number of fingers and the existence of tails for these new aliens in `test.csv`. Your goal is to identify the species of aliens in `test.csv` based on the intercepted messages, number of fingers, and presence of a tail. Submit a `result.csv` file only containing a column titled 'Species' that contains species for each row of `test.csv` based on your predictions. This part will be graded on the accuracy of the predictions in the `result.csv` file and the approach and cleanliness of your solution (`solution.ipynb`). Ensure that your approach is clearly written with comments and markdown cells in the python notebook that you submit.

Part 2: Predicting Troop Betrayal in the War Against the Phrygians

As you prepare for the inevitable war against the Phrygians, you face a dangerous internal challenge: **betrayal within your own ranks**. Historically, many Xernian soldiers have been lured by the Phrygians' promises of wealth and power, abandoning their clan and joining the enemy. As the troop leader, your mission is to predict which soldiers in your current army are most likely to betray your clan and join the Phrygians.

To prevent betrayal, you must design a decision-making system that analyzes various factors about your troops and flags those who are most likely to defect. This system will help you manage your forces strategically and prevent treachery.

The Challenge

Your task is to architect a system that evaluates the likelihood of betrayal within your army. Consider various factors that could influence a soldier's decision to betray the clan, including greed, temptation, poor respect, and more such factors. Focus on quantifying these factors into data points that can be used as features for your system.

Objective

Develop a decision-making system that evaluates each soldier based on the data you define from your hypotheses and predicts the likelihood of betrayal. Focus on:

- Identifying key factors (features) from your hypotheses and converting them into data that can indicate a higher likelihood of betrayal.

- Designing a workflow for how your system would analyze the data and rank each soldier based on their risk level.

- Presenting a well-thought-out approach to decision-making, outlining how the system evolves and adapts as new data is gathered, making it more scalable.

Evaluation Criteria

This task is open-ended, and you will be evaluated on:

- The clarity of your thought process and hypotheses for betrayal prediction. The quality of factors you choose and how you quantify them plays a huge role.

- The effectiveness and scalability of your decision-making algorithm in identifying potential traitors.

- Creativity in identifying relevant features and designing a system that can adapt to changing situations.

- Explaining the full stack carries points as well(like what all languages, libraries and resources you will use and how you will integrate these).

Submission Instructions

- For Part 1, submit `result.csv` corresponding to `test.csv` as input data. Also, submit `solution.ipynb` with your code and explanations.

- No code submission is required for Part 2 (although working code with README instructions and dataset will earn bonus points). Focus on presenting a clear thought process and execution flow.

- Prepare a well-structured PDF report explaining the steps mentioned in the objectives for Part 2.

- Submit all files before the deadline. Be as creative as possible in designing your decision-making system.