

CODE HACKATHON PROBLEM STATEMENTS

Problem Statement 1

Problem Statement: Sentiment Analysis of Customer Reviews

Description: Sentiment analysis is a technique for determining the sentiment or emotion expressed in a piece of text, such as a customer review or social media post. This information can be useful for businesses to understand customer opinions and feedback about their products and services.

Please Find the dataset for this Project in below link :

https://drive.google.com/drive/folders/1Rdn0kuDpinjwfDknh9T5vJ95AlqjJ3SO?usp=share_link

Goal :

The goal of this project is to build a sentiment analysis model that can accurately classify customer reviews as positive, negative, or neutral. Participants will be provided with a dataset of customer reviews from a range of product categories, such as electronics, clothing, and food. The data includes both the text of the review and the sentiment label.

Participants should explore different algorithms and techniques for sentiment analysis, such as logistic regression, support vector machines, and deep learning models. They should also experiment with different preprocessing steps, such as removing stop words and stemming, as well as feature engineering, such as word embeddings and n-grams.

Evaluation: The success of the project will be measured by the accuracy of the sentiment analysis model on a held-out test set of customer reviews. The performance of the solution will be evaluated using metrics such as accuracy, precision, recall, and F1-score. Participants should also consider the interpretability of the model, such as the most important features and the reasoning behind its predictions.

Deliverables: Participants should deliver a working sentiment analysis model that can be tested on the customer review dataset. The source code can be submitted via **GitHub**, you should also save the Model file using **HDF5** or any other Format. Additionally, participants should provide a report on how they have managed to train a model, they should specify their approach (Algorithm).

Note:

You can use any IDE's like Jupyter Notebook or VS Code etc, Your model will also be cross verified during evaluation at our end.

Problem Statement 2

Problem Statement: Image Classification of Land Use/Land Cover Types

Description: Image classification is a computer vision task that involves categorizing an image into one of several predefined classes, such as land use and land cover types. This task can be used for a variety of applications, such as land use planning, environmental monitoring, and resource management.

Dataset:

https://drive.google.com/drive/folders/1C-kp-AxfBZ8NkAf2AvVKfyvYVc6MG_jK?usp=share_link

Please read the Data description carefully, before processing the data.

Goal:

The goal of this project is to build an image classification model that can accurately identify different land use and land cover types from satellite images. Participants will be provided with a dataset of satellite images, each labeled with the type of land use or cover it represents, such as residential, commercial, industrial, and agricultural areas.

Participants should explore different algorithms and techniques for image classification, such as convolutional neural networks (CNNs), transfer learning, and data augmentation. They should also experiment with different preprocessing steps, such as resizing, normalization, and feature scaling.

Evaluation: The success of the project will be measured by the accuracy of the image classification model on a held-out test set of satellite images. The performance of the solution will be evaluated using metrics such as accuracy, precision, recall, and F1-score. Participants should also consider the interpretability of the model, such as the most important features and the reasoning behind its predictions.

Deliverables: Participants should deliver a working Classification model that can be tested on the Intel Image Classification dataset. The source code can be submitted via **GitHub**, and the model file should be saved in a format such as **HDF5** or TensorFlow. Additionally, participants should provide a report on how they have managed to train a model, they should specify their approach (Algorithm).

Problem Statement 3

Problem Statement: Customer Segmentation Using Machine Learning

Description: Customer segmentation is the process of dividing a customer base into groups of individuals that are similar in specific ways relevant to marketing, such as age, gender, interests, and behaviors. By using machine learning algorithms, businesses can gain valuable insights into customer behavior, identify patterns and trends, and develop targeted marketing strategies.

Dataset:

https://drive.google.com/drive/folders/1zuVaiAJkpluVtfUSfB1nhybmL9PqckTw?usp=share_link

Goal:

The goal of this project is to use machine learning techniques to segment customers into meaningful groups. Participants will be provided with a dataset of customer demographics and purchase behaviors. They should use different clustering algorithms such as k-means, hierarchical clustering, and DBSCAN to identify distinct customer segments.

Evaluation: The success of the project will be measured by the quality of the customer segments identified, including the degree of similarity within each segment and the distinctiveness between segments. Participants should also consider the interpretability of the segments and the ability to identify meaningful customer characteristics and behaviors.

Deliverables: Participants should deliver a working customer segmentation model that can be applied to the customer behavior dataset. The source code should be submitted via GitHub, and the final model should be saved in a format such as HDF5. Additionally, participants should provide a report on their approach, including the algorithms and techniques used, the results and findings, and any limitations or future work.

Note:

You can use any IDE's like Jupyter Notebook or VS Code etc, Your model will also be cross verified during evaluation at our end.