Project Report :
Rock vs. Mine Prediction

1. Introduction:

The goal of this project is to develop a predictive model that can classify underwater objects as either "Rock" or "Mine" based on sonar data. This is a supervised learning problem where we utilize machine learning algorithms, specifically logistic regression, to train a model on labeled data and then make predictions on unseen data.

2. Data Collection and Preprocessing:

The dataset used in this project is the Sonar Dataset, which contains sonar signals bounced off different objects under water.

The data consists of 208 observations and 60 features, with each feature representing a specific aspect of the sonar signal.
Initially, we load the data into a pandas DataFrame and perform basic exploratory data analysis (EDA) to understand its structure and characteristics.

The dataset is then divided into features (X) and labels (y), where X represents the sonar signal features and y represents the target variable indicating whether the object is a rock (R) or a mine (M).

3. Model Training:

We split the dataset into training and testing sets using the train_test_split function from the scikit-learn library. The split is stratified to ensure equal distribution of classes in both training and testing sets.

Logistic Regression is chosen as the classification algorithm for its simplicity and effectiveness in binary classification tasks.

The model is trained on the training data using the fit method, where it learns to map the input features to the corresponding labels.

4. Model Evaluation:

The trained model is evaluated on both the training and testing datasets to assess its performance.

Accuracy score is used as the evaluation metric to measure the proportion of correctly classified instances.

The accuracy on both the training and testing data is calculated using the accuracy_score function from scikit-learn.

5. Prediction:

Finally, the trained model is used to make predictions on new, unseen data.

We provide an example input data point representing sonar readings of an underwater object.

The input data is passed through the trained model to predict whether the object is a rock or a mine.

The prediction is then displayed along with an interpretation of the result.

6. Conclusion:

In conclusion, the logistic regression model trained on the Sonar Dataset demonstrates reasonable accuracy in classifying underwater objects as rocks or mines based on sonar signals. Further improvements could be explored by experimenting with different algorithms, tuning hyperparameters, or incorporating additional feature engineering techniques. However, the current model serves as a solid foundation for automated classification of underwater objects in real-world scenarios.

7. Future Work:

Potential future work could involve:

Experimenting with other classification algorithms such as Random Forest, Support Vector Machines, or Neural Networks.
Performing more in-depth feature engineering to extract more relevant information from the sonar signals.
Collecting additional data or exploring alternative datasets to enhance the model's robustness and generalization capabilities.