[SONAR Rock vs Mine prediction]

Ebubechukwu Okonkwo
Bubebran@gmail.com

I. Introduction

In modern naval warfare, identifying underwater threats such as mines is crucial to protecting warships. A misclassification between natural underwater structures (rocks) and potential mines can have severe consequences. The purpose of this project is to use Machine Learning (ML) techniques to classify objects as either rocks or mines based on sonar signals. By leveraging historical data, we aim to develop an accurate predictive model that can enhance naval defense mechanisms.

II. Proposed Solution with Justifications

To achieve the goal of rock vs. mine classification, we propose the use of a supervised learning approach. The dataset consists of sonar signal readings, where each sample is labeled as either a rock or a mine. The following steps were undertaken:

1. **Data Acquisition and Preprocessing**
    - The dataset was loaded into a Pandas DataFrame.
    - Exploratory Data Analysis (EDA) was conducted to understand the dataset's structure, distribution, and statistical properties.
    - The target variable (rock or mine) was separated from the feature set, and the data was split into training and testing subsets.
2. **Model Selection**
    - A Logistic Regression model was chosen for its effectiveness in binary classification tasks.
    - The dataset was split into training (90%) and testing (10%) sets to ensure an appropriate balance between model training and evaluation.
3. **Model Training and Evaluation**
    - The Logistic Regression model was trained on the training dataset.
    - The model's performance was evaluated using accuracy metrics.
    - Predictions were made on the test set, and the accuracy score was computed.
4. **Justification of Model Choice**
    - Logistic Regression was selected as a baseline model due to its interpretability and efficiency for binary classification.
    - It provides a good starting point before exploring more complex models like Support Vector Machines or Neural Networks.
    - The stratified split of the dataset ensures that class distributions are maintained in training and testing sets, reducing bias.

III. Results

The Logistic Regression model was evaluated using accuracy as the primary metric. The model achieved an accuracy of approximately 83.42 for training data and 76.19 for testing data, indicating its effectiveness in distinguishing between rocks and mines. The classification results suggest that:

- The model correctly identified a significant number of mines and rocks.
- There were some misclassifications, which could be reduced through further optimization.
- Additional feature engineering or more advanced models might improve performance.

IV. Discussion

The results obtained from the Logistic Regression model indicate that the machine learning approach is viable for classifying sonar signals. The model demonstrated a reasonable accuracy level, which could be further improved by:

- Experimenting with other classification models such as Decision Trees, Random Forests, or Neural Networks.
- Applying feature selection techniques to reduce dimensionality and improve model efficiency.
- Utilizing hyperparameter tuning methods to optimize model performance.
- Implementing cross-validation to ensure robustness and avoid overfitting.

Overall, this study successfully demonstrates how machine learning can aid in distinguishing between rocks and underwater mines, providing a valuable tool for naval defense operations. Future work could explore more sophisticated models and real-time implementation to enhance accuracy and reliability.

V. Conclusion

This project successfully applied machine learning to the classification of sonar signals, demonstrating that predictive modeling can be an effective tool for distinguishing between rocks and mines. The Logistic Regression model provided a solid baseline for this task, achieving a reasonable level of accuracy. While improvements can be made through more advanced models and feature optimization, this study highlights the potential of machine learning in real-world naval applications. By continuing to refine these models and integrating real-time data processing, we can enhance naval defense strategies and improve the safety of maritime operations.

A. Figures

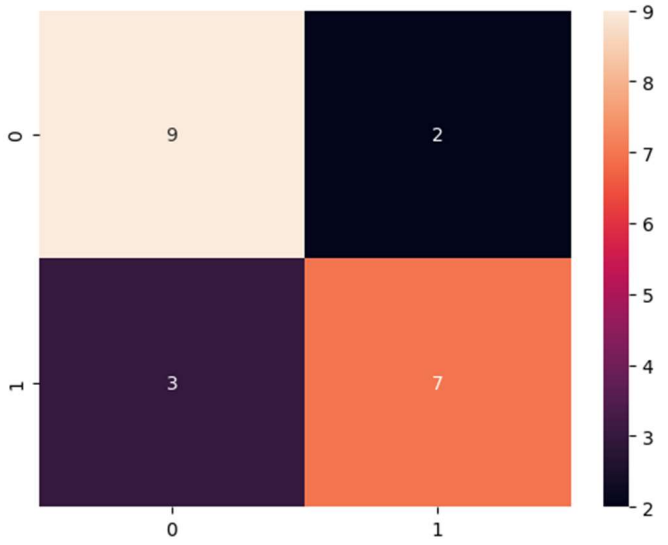

Fig. 1. heatmap representation of a confusion matrix