Project2: Machine Learning Mine Versus Rock
Name: Divyesh Rathod
ASU ID: 1225916954

Number of PCA components that achieved maximum accuracy: 7
Maximum accuracy: 87.3

Confusion matrix

[[23 9]
 [4 27]]

The confusion matrix shows the results of the classification model's predictions. In this case, there were 23 true negatives (TN), 9 false positives (FP), 4 false negatives (FN), and 27 true positives (TP). This means that the model correctly identified 23 out of 32 rocks and 27 out of 31 mines. However, it misclassified 9 rocks as mines and 4 mines as rocks.

Number of components

The PCA analysis found that using 7 principal components resulted in the highest test accuracy of 87.3%. The plot showed a general increase in accuracy with increasing components until reaching a peak at 7, after which it slightly declined, suggesting additional components may introduce noise or overfitting.

General shape of Plot

The general shape of the plot is a concave curve, suggesting that as the number of components increases, there is a rapid increase in test accuracy, followed by a leveling off. This is a typical pattern in PCA analysis and reflects the balance between capturing more data information and avoiding overfitting.

Parameters for MLP classifier

The parameters for the MLPClassifier were chosen through trial-and-error and common best practices. The logistic activation function was used for binary classification tasks, while the solver was set to 'adam', a popular optimization algorithm.

Conclusion

In this project, the MLPClassifier model was trained and tested on the Sonar dataset using PCA for feature reduction. The model achieved a maximum accuracy of 87.3% with 7 PCA components. The confusion matrix shows that the model has a higher accuracy in predicting rocks compared to mines. According to this, there is an 87.3% likelihood of correctly identifying a mine or a rock. We therefore have an 87.3% chance of surviving a mine.