**Mine vs. Rock Summary**

**Observations**  
After running the analysis, the highest test accuracy achieved was **92%** using **37 principal components**. This suggests that the model, trained with the first 37 components, has a 92% chance of accurately detecting mines in a real-world scenario. The confusion matrix, shown in Table 1, provides further insights into the model's classification performance.

|  | **Actual Class: Rock** | **Actual Class: Mine** |
| --- | --- | --- |
| Predicted Rock | 24 | 0 |
| Predicted Mine | 5 | 34 |

**Confusion Matrix**

The confusion matrix reveals that the model successfully classified all 34 mines (no false negatives), but incorrectly predicted **5 rocks as mines** (false positives). This is a favorable result since false negatives, where a mine is misclassified as a rock, are far more dangerous in this context. By minimizing false negatives, the model ensures a high safety level for real-world use in mine detection.

**Conclusions**  
From the confusion matrix, the total number of test samples is 63, with **29 labeled as rocks** and **34 labeled as mines**. The model's ability to identify all mines, without any misclassification, is critical, as mislabeling a mine as a rock could result in dangerous consequences. On the other hand, mislabeling rocks as mines (false positives) is less problematic, as it would result in unnecessary caution rather than danger.

The model achieved its peak accuracy at **37 components**, indicating that these components contain the most relevant features for differentiating between mines and rocks. As the number of components increased beyond 37, the accuracy began to decrease. This is likely because additional components did not provide useful information and may have introduced noise into the model.

**PCA Analysis and MLPClassifier Tuning**  
The graph of accuracy versus the number of components shows a steady increase in accuracy up to 37 components, followed by a decline. This suggests that only the first 37 components are highly correlated with the classification task, while the remaining components (38–60) contribute little to model performance.

For the **MLPClassifier**, the chosen parameters were:

* hidden\_layer\_sizes: (200, 100)
* activation: 'relu'
* max\_iter: 200
* alpha: 0.0001
* solver: 'adam'
* tol: 0.001
* learning\_rate: 'constant'
* random\_state: 1

These parameters were chosen based on a grid search to optimize performance. Setting random\_state to a fixed integer ensured consistent results across multiple iterations, making the comparison between different component counts fair.

**Alternative Results**  
An alternate configuration with **95% accuracy** was achieved by adjusting random\_state to 5. However, this configuration led to **two false negatives**, which poses a risk in the context of mine detection. Therefore, the configuration with **92% accuracy** was favored, as it resulted in **zero false negatives**, ensuring the safety of the submarine in a real minefield.

In conclusion, the selected model offers a high degree of accuracy with no false negative predictions, making it a reliable choice for detecting mines. However, further improvements could focus on reducing false positives to enhance overall performance.

*Figure 1: MLP Accuracy vs. Number of PCA components with hidden\_layer\_sizes=(200,100)*

