**README.md**

Predicting Exercise Execution Manner Using Accelerometer Data

This project aims to predict the manner in which individuals perform weight lifting exercises using data collected from accelerometers placed on various parts of the body. The dataset utilized is the "Weight Lifting Exercise Dataset."

Dataset Description

The dataset comprises sensor data collected from six participants performing barbell lifts in both correct and incorrect manners. Sensors were placed on the belt, forearm, arm, and dumbbell of each participant. The primary files include:

- pml-training.csv: Contains 19,622 observations with 160 variables.

- pml-testing.csv: Contains 20 observations with 160 variables.

Each observation includes sensor measurements and a target variable `classe` that indicates the manner of exercise execution:

- `A`: Correct execution.

- `B` to `E`: Various types of incorrect executions.

Analysis Steps

1. Data Preprocessing:

- Remove columns with a majority of missing or irrelevant values.

- Split the data into training and testing sets.

2. Model Selection:

- Implement the Random Forest algorithm for classification due to its robustness in handling datasets with numerous variables and its ability to manage overfitting.

3. Model Validation:

- Apply cross-validation techniques to evaluate model performance and ensure generalizability to new data.

4. Performance Evaluation:

- Use accuracy metrics to assess the model's performance on the testing set.

Results

The Random Forest model achieved high accuracy in predicting the manner of exercise execution, demonstrating the effectiveness of using accelerometer sensor data to assess exercise quality.

References

- Velloso, E., Bulling, A., Gellersen, H., Ugulino, W., & Fuks, H. (2013). Qualitative Activity Recognition of Weight Lifting Exercises. \*Proceedings of the 4th International Conference in Cooperation with SIGCHI (Augmented Human '13)\*. Stuttgart, Germany: ACM SIGCHI.

**Guidelines for HTML Report**

An HTML report provides a comprehensive and interactive medium to present your analysis. You can create this report using tools like R Markdown or Jupyter Notebook, which allow for the integration of code, visualizations, and narrative text. The report should include:

1. **Introduction**: Explain the project's objective and provide a brief description of the dataset.
2. **Data Preprocessing**: Detail the steps taken to clean and prepare the data for analysis, including handling missing values and feature selection.
3. **Model Selection and Training**: Discuss the rationale behind choosing the Random Forest algorithm, the training process, and any parameter tuning performed.
4. **Validation and Evaluation**: Describe the cross-validation approach used and present the model's performance metrics, such as accuracy, precision, recall, and F1-score.
5. **Conclusion**: Summarize the findings, discuss the model's strengths and potential limitations, and suggest possible directions for future work.

Ensure that the report includes relevant visualizations, such as accuracy plots, confusion matrices, and feature importance charts, to provide deeper insights into the model's performance and the data's characteristics.