Stage 1: Data Collection

Steps:

- Collect Data: Gather all relevant data points, ensuring they are accurate and complete.
- Initial Storage: Store the raw data securely for future reference and verification.

Folder Structure:

```
project/
  data/
    raw/
      data.csv
  notebooks/
    data_collection.ipynb
  src/
    data_collection.py
  reports/
    data_collection_report.md
```

Stage 2: Data Preprocessing

Steps:

- Data Cleaning: Handle missing values, correct errors, and ensure consistent formatting.
- Data Transformation: Convert all measurements to a common unit if necessary (e.g., kilograms for grip strength). Normalize or standardize the data as needed.

Folder Structure:

```
project/
  data/
    raw/
      data.csv
    processed/
      cleaned_data.csv
  notebooks/
    data_preprocessing.ipynb
  src/
    data_cleaning.py
  reports/
    data_preprocessing_report.md
```

Stage 3: Data Analysis

Steps:

- Exploratory Data Analysis (EDA): Visualize the data to understand distributions, relationships, and potential outliers.
- Feature Selection: Identify which features are most predictive of frailty.
- Model Development: Develop and train machine learning models to predict frailty based on the provided features. Use techniques like logistic regression, decision trees, or other appropriate methods.
- Model Evaluation: Validate the models using appropriate metrics such as accuracy, precision, recall, and F1-score.

Folder Structure:

```
project/
  data/
    processed/
      cleaned_data.csv
  notebooks/
    data_analysis.ipynb
    model_development.ipynb
  src/
    feature_selection.py
    model_training.py
  models/
    logistic_regression.pkl
    decision_tree.pkl
  reports/
    eda_report.md
    model_evaluation_report.md
    visualizations/
      eda_plots.png
      model_performance_plots.png
  deployment/
    model_inference.py
    requirements.txt
  docs/
    README.md
    usage_instructions.md
```