**Code Work Distribution (R1-R5)**

**Key Areas**:

* **R2: Data Analysis and Preprocessing**
* **R3: Clustering Algorithms**
* **R4: Baseline Model Implementation**
* **R5: Neural Network Implementation**

**Data Preprocessing and EDA (R2):**

* **Assigned Members**: Anu,Fiyi,Ajay
* **Tasks**:
  + Data cleaning (handling missing values, normalization, etc.).
  + Exploratory Data Analysis (EDA): Visualizing trends, relationships between variables (correlation heatmaps, histograms, etc.).
  + Feature engineering: Creating new features, removing redundant ones.
  + **Output**: Clean dataset, visualizations for report and presentations.
  + **Code Deliverable**: Functions to clean and preprocess data, EDA scripts.

**Clustering Algorithms (R3):**

* **Assigned Members**: Mariam,Hari
* **Tasks**:
  + Implement two clustering algorithms (e.g., **K-Means** and **Hierarchical Clustering**).
  + Fine-tune hyperparameters (e.g., K-value in K-Means).
  + Analyze and interpret the clusters.
  + **Output**: Cluster analysis visualizations, cluster assignments for the dataset.
  + **Code Deliverable**: Clustering algorithm scripts with visualizations (e.g., scatter plots, dendrograms).

**Baseline Models (R4):**

* **Assigned Members**: Anu,Fiyi,Mariam
* **Tasks**:
  + Implement three machine learning algorithms: **Decision Trees**, and two others (e.g., **Naive Bayes**, **KNN**).
  + Evaluate models using metrics such as **accuracy**, **precision**, **recall**, **F1 score**, and **RMSE** (for regression if applicable).
  + Compare model performances.
  + **Output**: Model comparison table and performance metrics.
  + **Code Deliverable**: Baseline model scripts and evaluation functions.

**Neural Networks (R5):**

* **Assigned Members**: Ajay,Hari,Mariam,(Anu,Fiyi will back up in case)
* **Tasks**:
  + Implement neural network models (e.g., **MLP** and **CNN**).
  + Train, fine-tune, and evaluate the models.
  + Compare neural network performance against baseline models.
  + **Output**: Performance comparison with baselines, loss/accuracy curves.
  + **Code Deliverable**: Neural network implementation scripts, comparison plots.

**Report Writing Distribution**

The report has several key sections that correspond to the R1-R5 tasks. Each group member can be assigned both coding and the writing of specific sections related to their code tasks.

**Report Sections and Assignments:**

**1. Introduction & Problem Definition (R1)**

* **Assigned Members**: Anu,Mariam
* **Tasks**:
  + Write the **Introduction** outlining the problem and research questions.
  + Explain the hypothesis, goals, and importance of the project.
  + **Deliverable**: 1–1.5 pages of the report.

**2. Dataset Description and Data Preprocessing (R2)**

* **Assigned Members**: Ajay,Fiyi
* **Tasks**:
  + Describe the dataset (source, structure, key features).
  + Explain preprocessing steps (handling missing data, feature scaling, etc.).
  + Include **EDA** insights (patterns, trends, etc.).
  + **Deliverable**: 1–2 pages of the report with EDA visualizations.

**3. Clustering Analysis (R3)**

* **Assigned Members**: Hari,Mariam
* **Tasks**:
  + Describe clustering techniques used (K-Means, Hierarchical, etc.).
  + Provide insights from clustering analysis (e.g., cluster characteristics).
  + Present and explain clustering visualizations.
  + **Deliverable**: 1 page of the report.

**4. Baseline Model Implementation and Evaluation (R4)**

* **Assigned Members**: Anu,Ajay
* **Tasks**:
  + Explain the machine learning algorithms used (Decision Trees, Naive Bayes, KNN).
  + Provide details on model evaluation (metrics such as accuracy, F1 score, etc.).
  + Discuss results and compare model performances.
  + **Deliverable**: 1 page of the report, including tables/graphs of model performance.

**5. Neural Network Implementation and Comparison (R5)**

* **Assigned Members**: Fiyi,Hari,Mariam
* **Tasks**:
  + Explain neural network models (MLP, CNN) used.
  + Discuss training processes, hyperparameters, and model evaluation.
  + Compare neural networks with baseline models.
  + **Deliverable**: 1 page of the report with performance comparison tables/graphs.

**6. Results Discussion and Conclusion**

* **Assigned Members**: Hari, Ajay, Anu,Fiyi ,Mariam
* **Tasks**:
  + Summarize key findings from the clustering, baseline, and neural network models.
  + Provide insights, implications of results, and how they align with the research questions.
  + Discuss any limitations or improvements.
  + **Deliverable**: 1 page of the report.