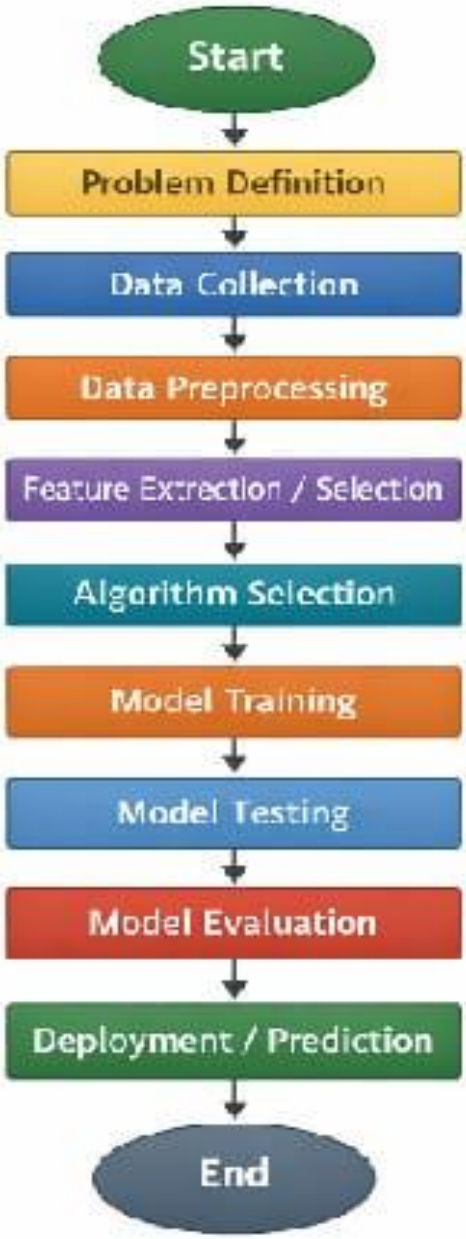
# Group Task 3: Module 3

## Machine learning: Concepts, Algorithms and Application

**Build a Simple ML Process Flow: Groups create a complete flowchart for a machine learning project, covering data collection, feature extraction, algorithm selection, training, testing, and evaluation.**



Here is the transcription of your notes with spelling and grammatical corrections to make them more meaningful, while keeping the original structure intact.

**1. Problem Definition**

- Clearly define what you want to predict or classify.
- Define what problem you are solving.
- Identify whether it is Classification, Regression, or Clustering.
- Decide success criteria:
- Identify the business or real-world problem.
- Example: Predict house prices; detect spam email.

**2. Data Collection**

- Gather raw data from different sources.
- Ensure data is relevant to the problem.
- Collect sufficient quality of data.
- Identify structured and unstructured data.
- Check data permissions and privacy rules.
- Example: House size, location, no. of rooms.

**3. Data Processing**

- Clean the data.
- Handle missing values, remove duplicates.
- Convert text/categorical data to numbers.
- Normalize or scale numerical values.
- Remove outliers.
- Example: Replace missing price with average value.

**4. Feature Extraction/Selection**

- Choose important input for the model.
- Remove unnecessary data to improve accuracy.
- Creating new useful features from raw data (feature engineering).
- Choosing only important features.

- Example: Select area and location as key features.

**5. Algorithm Selection**

- Choose an ML algorithm based on the problem type.
- Classification: Decision Tree, SVM.
- Regression: Linear Regression.
- Factors to consider: size of data, speed required, accuracy needed.
- Example: Linear Regression for price prediction.

**6. Model Training**

- Feed training data to the algorithm.
- The model learns patterns from the data.
- Data is divided into training and testing sets.
- The algorithm learns patterns from training data.
- Learn relationship between input and output.
- Model adjusts parameters.
- Example: Learning correlation between area and price.

**7. Model Testing**

- Test the trained model using unseen data.
- Check how well the model performs.
- Testing checks model performance on unseen data (use test dataset).
- Compare predicted output with actual output.
- Example: Predict price for new houses.

**8. Model Evaluation**

- Measure performance using metrics.
- Accuracy, Precision, Recall, RMSE.
- For Regression: Mean Absolute Error (MAE), Mean Squared Error (MSE).
- Example: 90% accuracy in predictions.

**9. Deployment / Prediction**

- Use the trained model in real-world applications.
- Model gives output for new inputs.
- Integrate into website or app.
- Monitor performance continuously.
- Update model when new data arrives.
- Use AI in cloud service.
- Example: Website showing predicted house price.