**Step 1: Setup & Dataset (30–45 min)**

**Tasks:**

* Install required libraries: pandas, numpy, scikit-learn, matplotlib, seaborn, joblib
* Download dataset (UCI German Credit Dataset or Kaggle Credit Dataset)
* Load dataset into a **Pandas DataFrame**

**Learning:**

* Practice **Python data handling** with Pandas
* Understand the **structure of tabular datasets**

**Step 2: Exploratory Data Analysis (EDA) (1.5–2 hours)**

**Tasks:**

* Check dataset shape, missing values, data types
* Plot **feature distributions** (histograms, boxplots)
* Visualize **correlations** with heatmaps
* Detect **outliers** and consider handling them

**Learning:**

* Understand **data distribution & relationships**
* Identify features that may need **transformation or encoding**

**Step 3: Data Preprocessing (1.5–2 hours)**

**Tasks:**

* Handle missing values (mean, median, or KNN imputation)
* Encode categorical variables (one-hot or ordinal encoding)
* Scale numeric features (StandardScaler or MinMaxScaler)
* Split data into **train/test sets** (80%/20%)

**Learning:**

* Learn **feature engineering & preprocessing techniques**
* Practice **train/test splitting & scaling** for ML models

**Step 4: Model Training (2–2.5 hours)**

**Tasks:**

* Train **Logistic Regression**
* Train **Decision Tree Classifier**
* Train **Random Forest Classifier**
* Evaluate each model using **accuracy, precision, recall, F1-score**

**Learning:**

* Understand **classification algorithms**
* Learn **basic model evaluation**

**Step 5: Model Evaluation & Visualization (1–1.5 hours)**

**Tasks:**

* Plot **confusion matrix** for each model
* Plot **ROC curve & calculate ROC-AUC**
* Compare all models to **select the best performer**

**Learning:**

* Learn to **interpret evaluation metrics**
* Visualize results to **make insights actionable**

**Step 6: Hyperparameter Tuning (1–1.5 hours)**

**Tasks:**

* Use **GridSearchCV or RandomizedSearchCV** on Decision Tree & Random Forest
* Find the **best parameters** and retrain models
* Compare performance before/after tuning

**Learning:**

* Learn **how hyperparameters impact model performance**
* Understand **cross-validation**

**Step 7: Save & Reuse Model (30 min)**

**Tasks:**

* Save the trained model using **joblib or pickle**
* Test loading and predicting with saved model

**Learning:**

* Learn **model persistence** for real-world deployment

**Step 8: Documentation & GitHub (1–1.5 hours)**

**Tasks:**

* Create **README.md** with project overview, instructions, and results
* Upload **code notebook or Python scripts** to GitHub
* Include **visualizations and metrics** in README
* Optional: add a small **requirements.txt** file

**Learning:**

* Learn **professional project presentation**
* Make your work **portfolio-ready**

**Step 9: LinkedIn Post (30 min)**

**Tasks:**

* Write a short post explaining your **project objective, approach, and results**
* Add **GitHub link**
* Optional: share **screenshots of metrics/plots**

**Learning:**

* Learn to **showcase your ML work professionally**
* Start **personal branding for AI & ML**