**Major Project for Machine learning using Python Batch I**

**Task 1 using the Dataset for Regression Task:**

1. **Import Data, head, tai, describe, shape & dtypes**
2. **Box Plot, Histogram & Density Curve for Target Variable – Purchases**
3. **Hypothesis Testing**
4. **Gender with Purchases**
5. **Age bins with Purchases**
6. **City Category with Purchases**
7. **Marital Status with Purchases**
8. **Occupation with City Category**
9. **Occupation with Gender**

**For Hypothesis testing first do the groupby, Choose the relevant hypothesis test & frame the Null & Alternate Hypothesis. Interpret & write the Analysis.**

1. **Clean the Missing values using the relevant strategy**
2. **Label Encode the relevant Variables**
3. **Run the following Algorithms:**
4. **Multiple Linear Regression**
5. **Decision Tree Regressor**
6. **Random Forest Regressor**
7. **Gradient Boosting Regressor**
8. **Neural Network Regressor**
9. **Support Vector Machine Regressor**
10. **KNN Regressor**
11. **Predict for Each Model, Calculate the Residuals & Root Mean Square Error.**
12. **Identify which is the best and worst model.**

**Task 2 using the Dataset for Classification task:**

1. **Import Data, head, tai, describe, shape & dtypes**
2. **Conduct Hypothesis tests on the following and analyze**
3. **Gender vs Attrition**
4. **Education Field vs Attrition**
5. **Gender vs Monthly Income**
6. **Marital Status vs Monthly Income**
7. **Distance from Home vs Attrition**
8. **Gender vs Years at Current role**
9. **Marital Status vs Years at Current role**

**For Hypothesis testing first do the groupby, Choose the relevant hypothesis test & frame the Null & Alternate Hypothesis. Interpret & write the Analysis**

1. **NO MISSING DATA**
2. **Label Encode the Relevant Variables**
3. **Run the following Algorithms:**
4. **Binary Logistic Regression**
5. **Decision Tree Classifier**
6. **Random Forest Classifier**
7. **Gradient Boosting Classifier**
8. **Neural Network Classifier**
9. **Support Vector Machine Classifier**
10. **KNN Classifier**
11. **Predict for each model, create the confusion matrix and calculate the Accuracy**
12. **Identify which is the best and worst model.**

**Complete the Assessment in the next 10 Days and send it to the same email Id to which the Minor Projects were sent with your Name & payment id, etc. as specified earlier during the Minor Projects.**