***ISA Python Practical Lab session (6.30 pm to 9pm):***

***TEST 2***

***Total time: 2.30 hours. Full marks: 50 marks.***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

***Instructions:***

Read the instructions carefully as listed below for the practical Lab session:

Consider the following given dataset on a ***Loan classification Problem***.

The data has two parts: ***1)*** Train.csv ***2)*** Test.csv

Business objective: To frame a predictive model using Python to detect the Loan status (Yes/No) for customers. Build the model on the train.csv and finally validate the model in test.csv

***Part 1: (Lab session)*** ***Marks: 50***

***A) Data Exploration (15 marks).*  1 hour 25 mins.**

1) Import the data and explore it. ***Exploratory Data Analysis*** and ***Data Visualization.*** Exploration of variables include analysis of Categorical Vs Categorical, Continuous Vs Continuous variables as well. Statistical analysis of visualization as well as statistical measurement of significance/ relationship of predictor variables with the target variables should be done.

2) Data pre-processing : Checking/replacing missing values. Checking/treating presence of outliers in the data.

***B) Predictive modelling / Regression analysis: (25 marks).*  1 hour 15 mins.**

1) Data splitting.

2) Execution of Logistic regression and forming final model.

3) Decision tree analysis.

4) Random Forest Analysis.

PART 2

1) AdaBoost

2) Plotting decision trees from AdaBoost

3) XG Boost

4)