**Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Each question carries equal marks (=20 marks). Please describe each question in details using examples, and Sci-kit learn (Python) functions where applicable. You can type-write the answers or write-with a pen/pencil and then scan it. In both the cases, you are going to upload one PDF/word/jpg file. Make necessary assumptions when necessary. Since this is an open book exam, plagiarism policy is strictly enforced and you are supposed to be working by yourself. If any trace of collaboration/cooperation is observed all the involved parties will receive an automatic ‘0’ (zero).

You are hired by a company X (company of your choice), answer these questions in relationship to **this company**.

1. Describe descriptive analysis, diagnostic analysis, predictive analysis and prescriptive analysis with one example for each type of analysis for this company.
2. Define and describe feature extraction and feature selection for a predictive analysis problem that you chose in **Question 1**. What are the differences between feature extraction and feature selection? Describe with example.
3. For the predictive analysis in Question 1 (e.g. next year’s revenue or will there be profit or loss for next year or similar example), describe your approach for choosing a final ML model. Let’s imagine you developed 6 different ML models for a problem for this company. Describe what is a ROC curve and AUC value. Describe with examples using Sci-kit learn functions.
4. For the same company, imagine a where the dataset has no labels (you can still use the same problem that you posed in Question 1), what type of ML algorithm(s)would you use? Give an example of the dataset of this type and discuss K-means clustering with example.
5. While developing ML models you use your training data but the ultimate goal is to be able to predict well on an unseen data, how do you make sure this happens? Describe this in relationship to over-fitting and under-fitting with example for this company?