Campus Recruitment

Introduction

For selecting a candidate in Campus recruitment, there are various factors which are considered. Students from a particular University also want to know the factors upon which he/she will get placed.

Objective

Follow the Data Science Methodology. Analyze the data set and identify the parameters that affects the campus recruitment. Answer the questions like

Which factor influenced a candidate in getting placed?

Do percentage matters for one to get placed?

Which degree specialization is much demanded by corporate?

Dataset

The dataset, **Placement\_Data\_Full\_Class.csv**, is taken from Kaggle. It has been uploaded to canvas.

This data set consists of Placement data of students in XYZ campus. It includes secondary and higher secondary school percentage and specialization. It also includes degree specialization, type and Work experience and salary offers to the placed students.

Tasks in this assignment

1) Write a Data Science Proposal for achieving the objective mentioned.

2) Perform exploratory analysis on the data and describe your understanding of the data.

* 1. 3) Perform data pre-processing. E.g., missing data, normalization, discretization, etc.,.
  2. 4) Apply any two feature selection engineering techniques
  3. 5) Compare the two selected feature engineering techniques.
  4. 6) Use Any regression techniques you have studied to predict.
  5. 7) Compare the performance of the regression technique before and after applying the pre-processing and feature engineering techniques.
  6. 8) Present the conclusions/results (Answers to the above-mentioned objectives and tasks) in the format shared.

# **Expected Submissions**

Two files are expected as the assignment submission.

1. The summary of the work including the answers to the objectives and the inference from the analysis in the template provided. Rubrics can be used as reference to the work to be done. (you may fill only the boxes relevant to this problem statement)
2. The executed ipynb file with clear subdivision of the codes and brief description of the purpose of respective code. All the executed tables or graphs and results should be present in the ipynb file.