**Assignment**

Objective is to create a model to predict remuneration of the vacancy/position in a data analytics firm.

In the train dataset, we have the input variables:

1. Experience
2. Job description
3. Job designation
4. Job Type
5. Key skills
6. Location
7. Salary
8. Company code

In the test dataset, we have all the above variables except **Salary** (which we need to predict)

Apart from the output, also answer briefly, the following points:

1. Final variable list used in the model
2. Pre-processing variables, e.g. cleaning variables, re-classification of categories, creation of dummy variables, outlier treatment, imputed variables, changed variable datatype/ structure, etc.
3. Model technique used
4. Variable importance/ strength/ contribution (if possible)
5. Any other information

**Note:** Please put visualizations wherever possible. Jupyter notebook to be used for assignment.