

## Data Science with Python Assignment 02



### Instructions

- Refer to the group list published in SCVLE in order to find your assigned group. Please note that mutual changes are not allowed.
- Find an appropriate dataset and come up with a suitable problem (Either prediction or classification of data) that can be solved using data science techniques that you have learned so far.
- After the submission of the working model, you will need to face a viva session in order to pass the assignment.
- Adhere to the following guidelines to gain maximum marks for your assignment.

### Guidelines for the report:

#### **Proposal submission (10%)**

- *You need to clearly indicate the dataset that you are supposed to use and provide an explanation of the problem that you are going to solve.*
- *Maximum 1 page proposal is required.*
- *This step will be conducted prior to the final presentation.*

#### **Pre-processing of the dataset (30%)**

- *The use of appropriate data cleaning and manipulation tasks.*
- *Any data visualization conducted to improve the insight of the cleaned dataset.*

#### **Model creation (30%)**

- *Creation of appropriate model to overcome or predict the expected outcome (Either Regression/classification or other)*
- *Clear explanation to use such a model to have the output.*

**Post-processing of the results (20%)**

- *Any usage of post processing procedures (pruning routines, rule quality processing, rule filtering, or even knowledge integration etc.) to provide a kind of symbolic filter for noisy, imprecise.*

**Evaluation of the data (10%)**

- *Clearly mentioning the accuracy and visualizations of the final output.*