

BFC AI Summer Training – Machine Learning

Final Project Description

Group 4

You are required to implement your own final project as *individual*, you have some requirements to do in criteria of **100 points**:

- **Implementation – (50 Points)**
 - Perform Exploratory Data Analysis with Visualization (10 points)
 - Clean your data as much as possible including removing outliers if exist (5 points)
 - Encode any categorical attribute in your dataset if exist (5 points)
 - Standardize your attributes if possible (10 points)
 - Use **Linear Regression** to classify your records (10 points)
 - Tune your hyperparameters using GridSearchCV to give better performance (5 points)
 - Visualize your model performance (train against validation) including ROC-AUC curve and training performance and confusion matrix (5 points)
- **Evaluation – (10 Points)**
 - Your model Performance will be evaluated using R2 Coefficient of Determination metric against hidden test set.
 - You should achieve at least 90% of the specified metric.
- **Presentation – (20 Points)**
 - Submit your notebook as part of the presentation (5 points)
 - Create your own PowerPoint presentation at maximum 10 Slides showing your proposed work (5 points)
 - Discuss your project in maximum of 5 Minutes in front of your mates (10 points)
- **Communication – (20 Points)**
 - Record a video about 2 minutes showing your notebook and the proposed work (5 points)
 - Post your video with description on your **LINKED IN** profile and add a mention to the trainer @Yousef Elbaroudy and add hashtag of #BFC AI and #MachineLearning (5 points)
 - Upload your notebook on your Kaggle account and make it public (5 points)
 - Upload your proposed work on your account through Github platform (5 points)

Trainer: Yousef Elbaroudy