Date of Submission: **Max Marks:20**

NOTE: This is a Group Assignment and one submission per group is accepted. Submissions to this Assignment will only be accepted as Kernels in Kaggle. Please mention complete Team Member Details in your Kernel!

TOPIC: Classification And Clustering

Use the given dataset which describes records of absenteeism at work from July 2007 to July 2010 at a courier company in Brazil. Split the dataset into Training and Testing data. Use any 2 classification or clustering techniques to predict 'Absenteeism time in hours'. Justify any parameter (such as k (for k-NN or k-means), etc.) if pre-determined. Compare the techniques used and state which of the two is better for this data. Clearly interpret the results and support your interpretations with computations of suitable performance measures.

https://www.kaggle.com/chetnasureka/absenteeism-at-work