

# Assignment: Predictive Analytic

- Select datasets.
- Perform the experiments.
  - Select 2 classification methods for creating models.
  - Create models
    - non-spark (i.e., sklearn) and
    - spark platforms (pyspark.mlib).
  - Experiment with 5-fold cross validation.
  - Predictive evaluation:  
Confusion matrix, Accuracy, Precision, Recall, F1.
  - Processing time comparison (training phase and testing phase)
- Write the report.
- Presentation