Assignment-4  
 Machine learning   
  
 **Total: 150 points**

# Part 1:

* Use file called Project.csv as your dataset.In the dataset “behavior\_type” is your classification column.
  + Clean the data (for example use if the behavior column value is “AdsBot-Google;Android browser;Chrome;Chrome Mobile”, change it to “AdsBot-Google”. But do it in a procedural way, not manually) **(5 points)**
* Provide the list of top 5 features for a random forest model to classify “behavior\_type”
  + Use a filter method **(5 points)**
  + Use a wrapper method **(5 points)**
  + Use an embedded method **(5 points)**
* Develop a random forest model to classify “behavior\_type”
  + Show the confusion matrix **(5 points)**
  + Provide the accuracy of the model **(5 points)**
  + Draw the ROC curve  **(5 points)**

# Part 2:

* Use file called Project.csv as your dataset.In the dataset “behavior\_type” is your classification column.
  + Develop a XGBoost model to classify “behavior\_type” **(10 points)**
  + Develop a deep neural network model to classify “behavior\_type” **(15 points)**

# Part 3:

* Provide a global and local explanation of an instance for your XGBoost model using associated explainable AI method using SHAP and LIME. **(20 + 20 points)**
* Provide a global and local explanation of an instance for your DNN model using associated explainable AI methods using SHAP and LIME **(25 + 25 points)**