# Istanbul Okan University

## **Faculty of Engineering**

**Data Mining Project**

Doctor Pinar

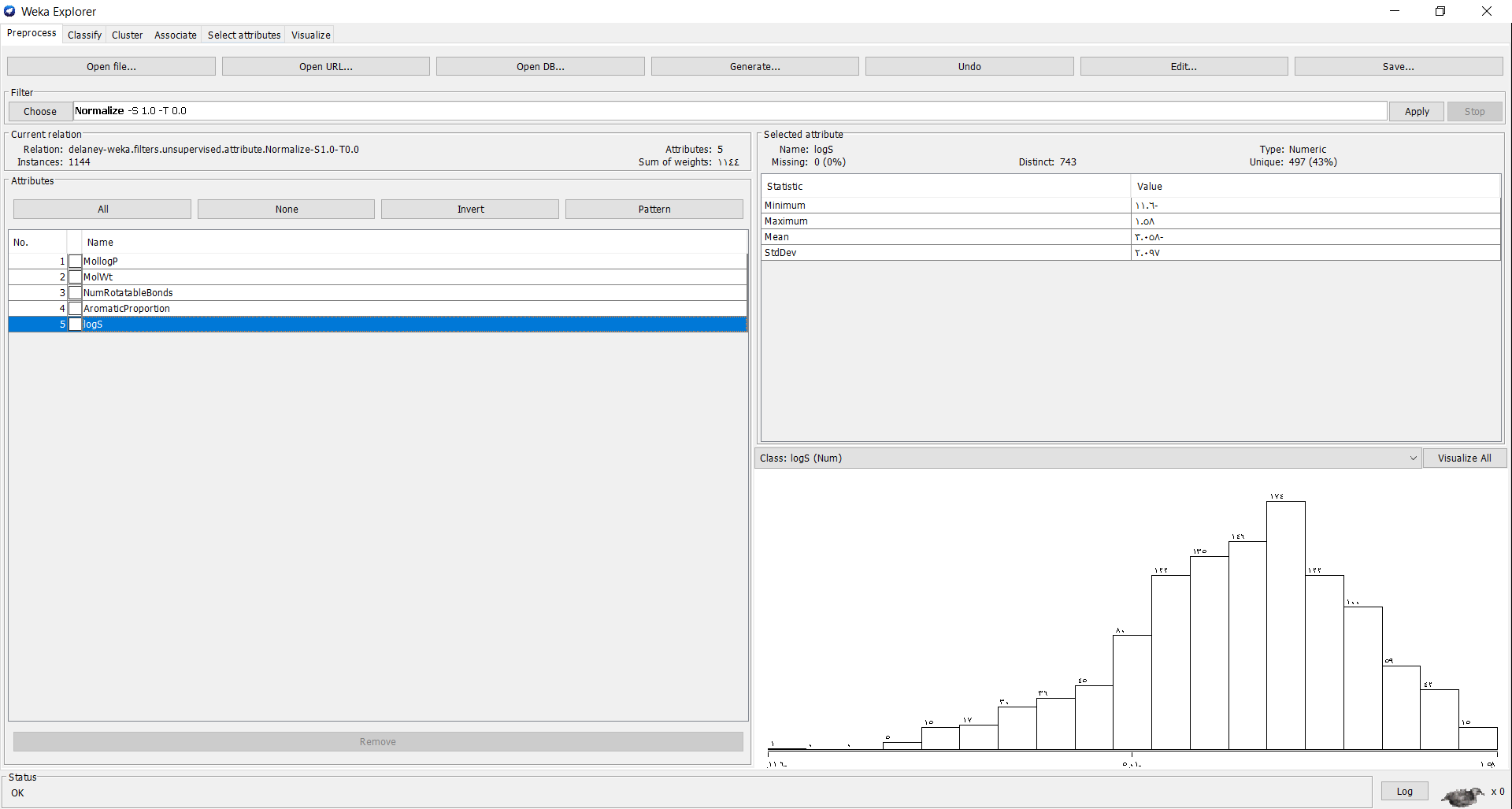
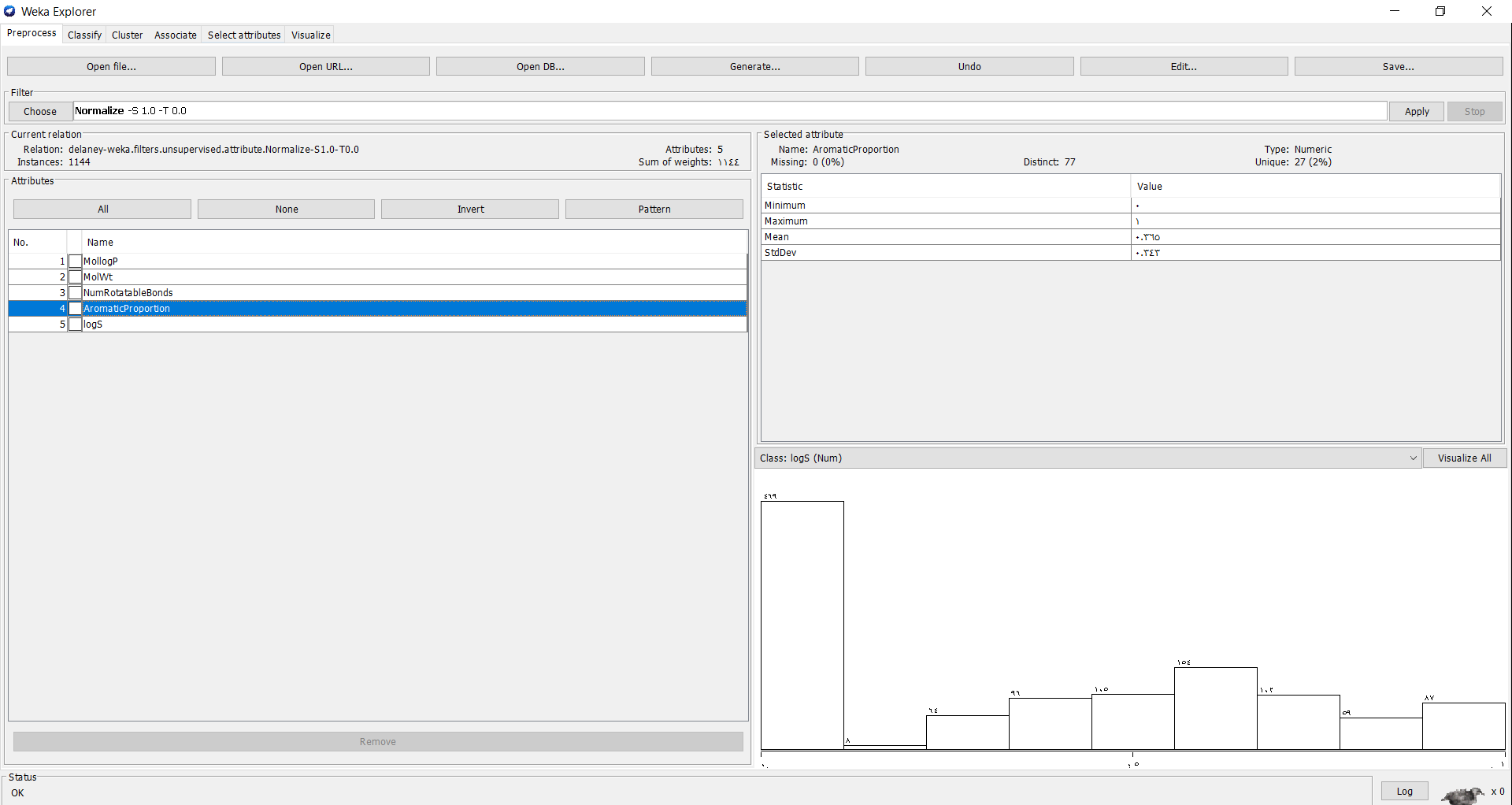
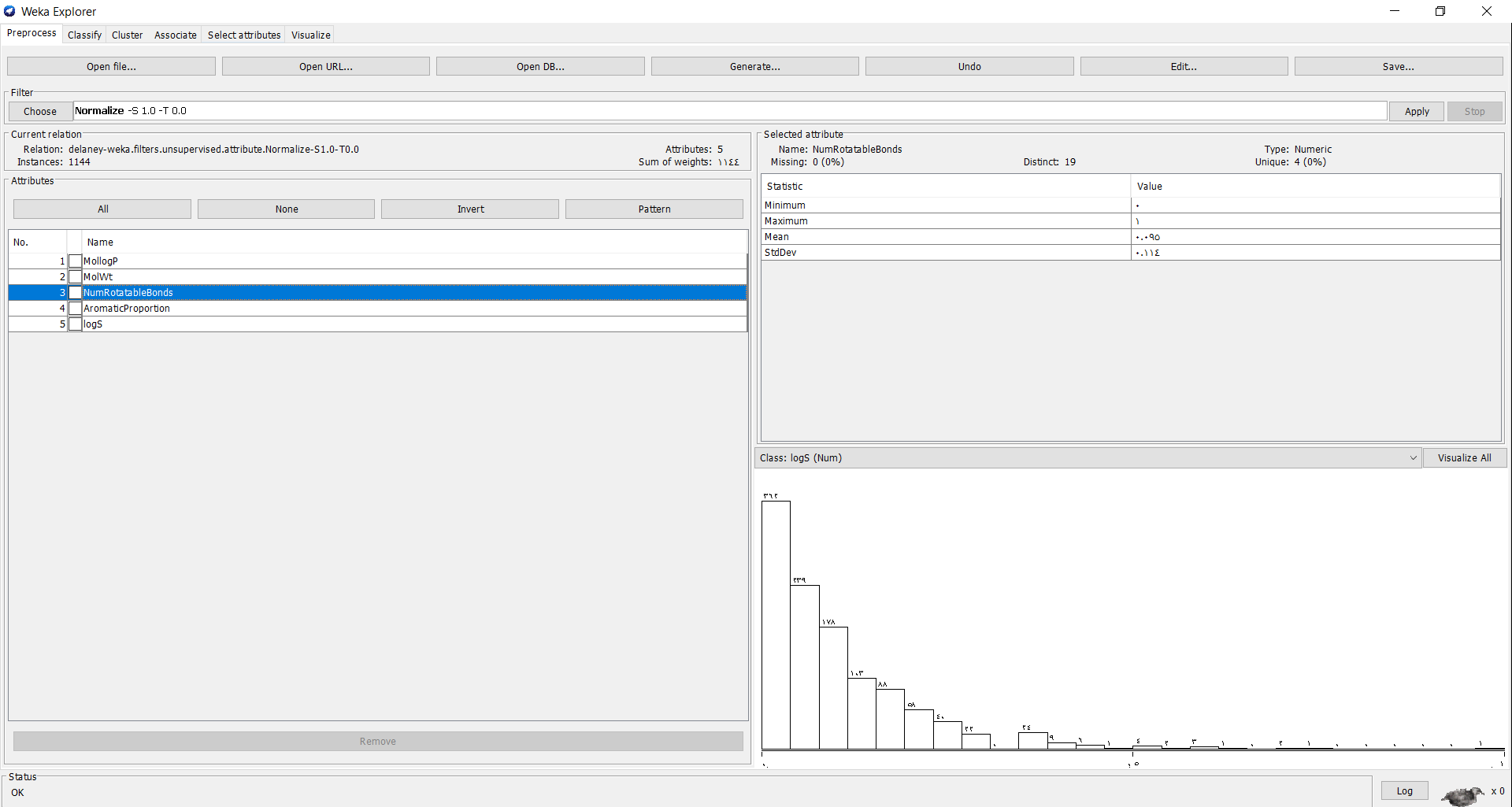
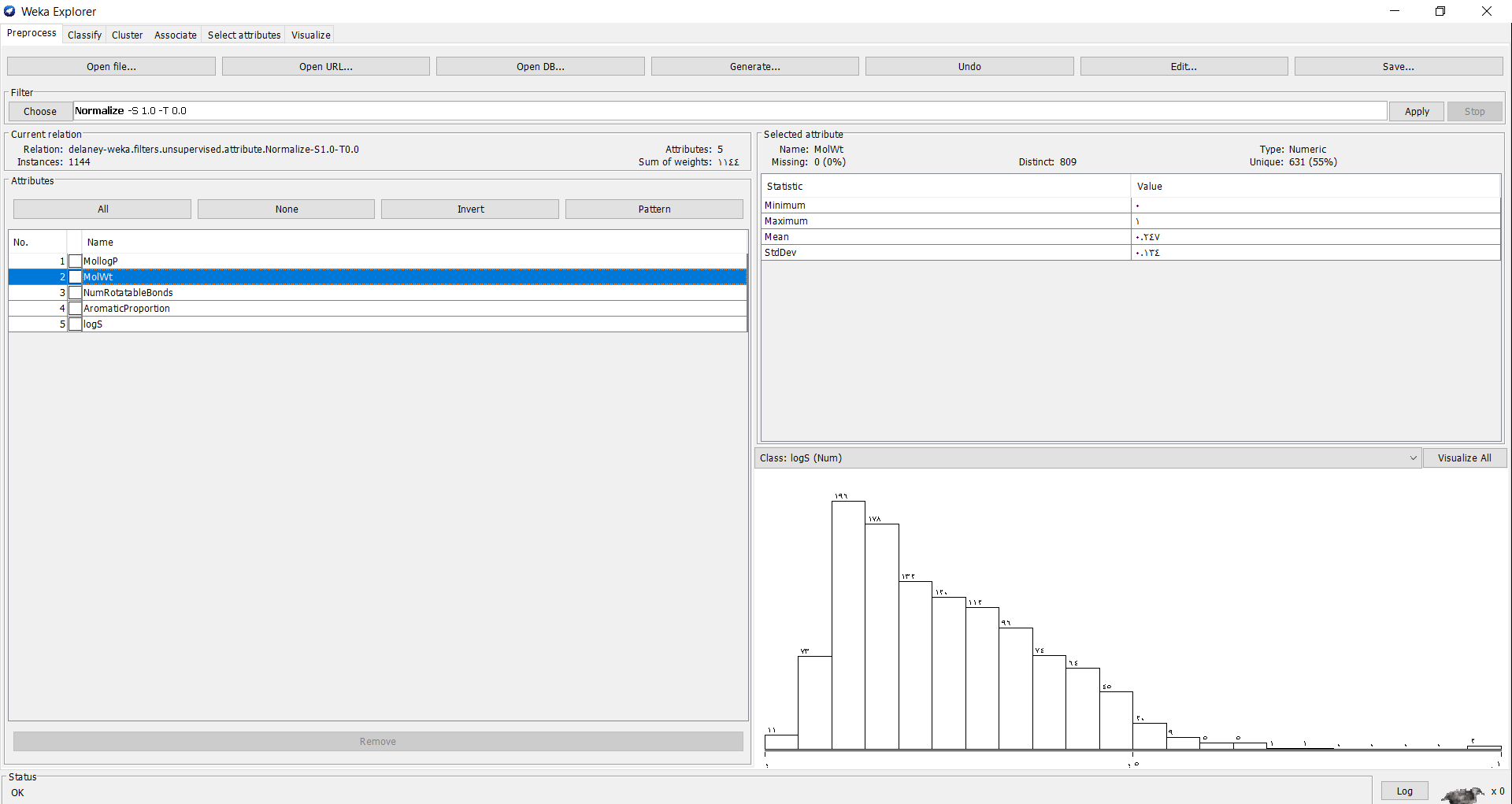
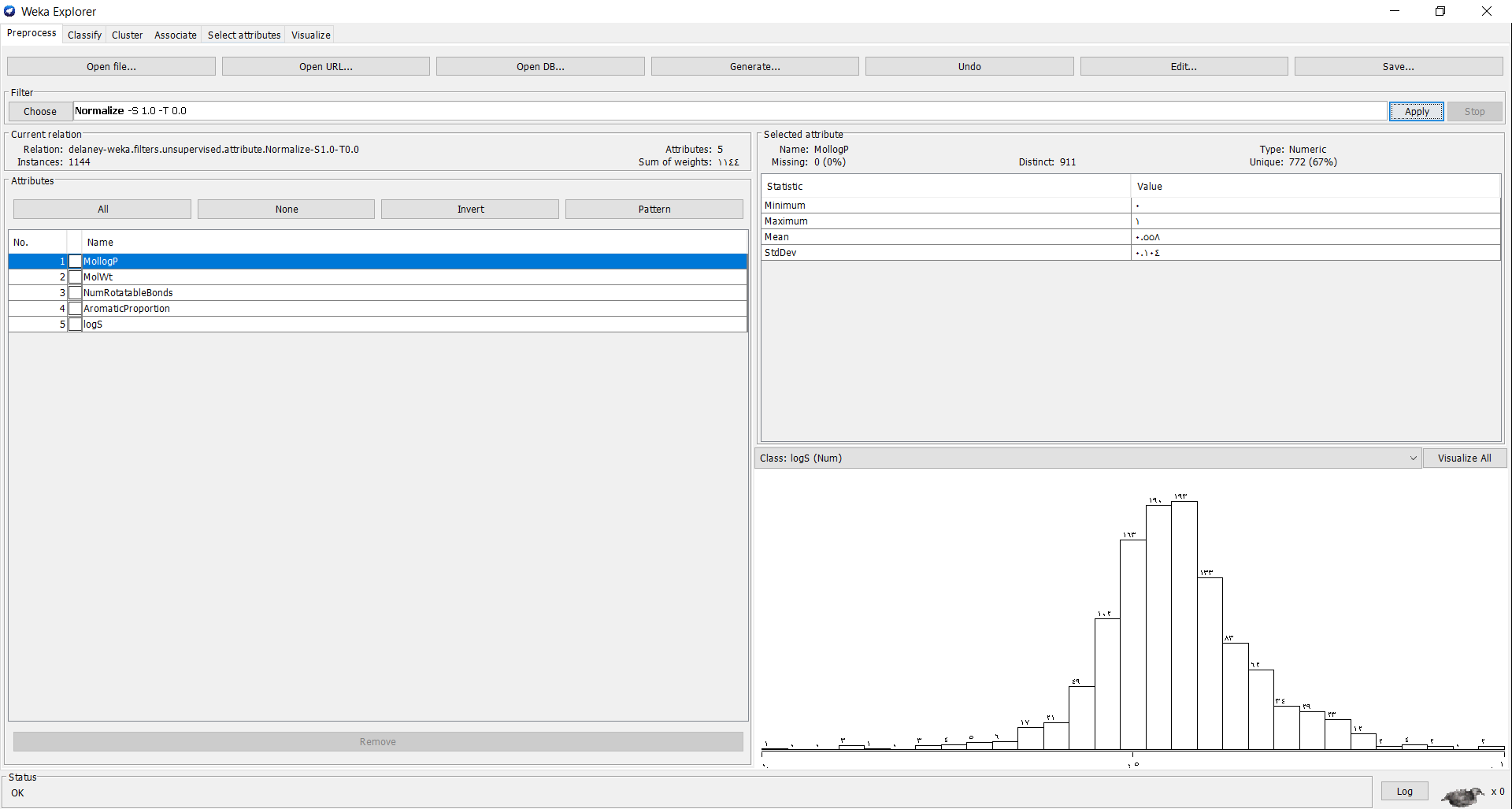
Student : Marwan Bakr Mohamed Saeid

University no : 220212347

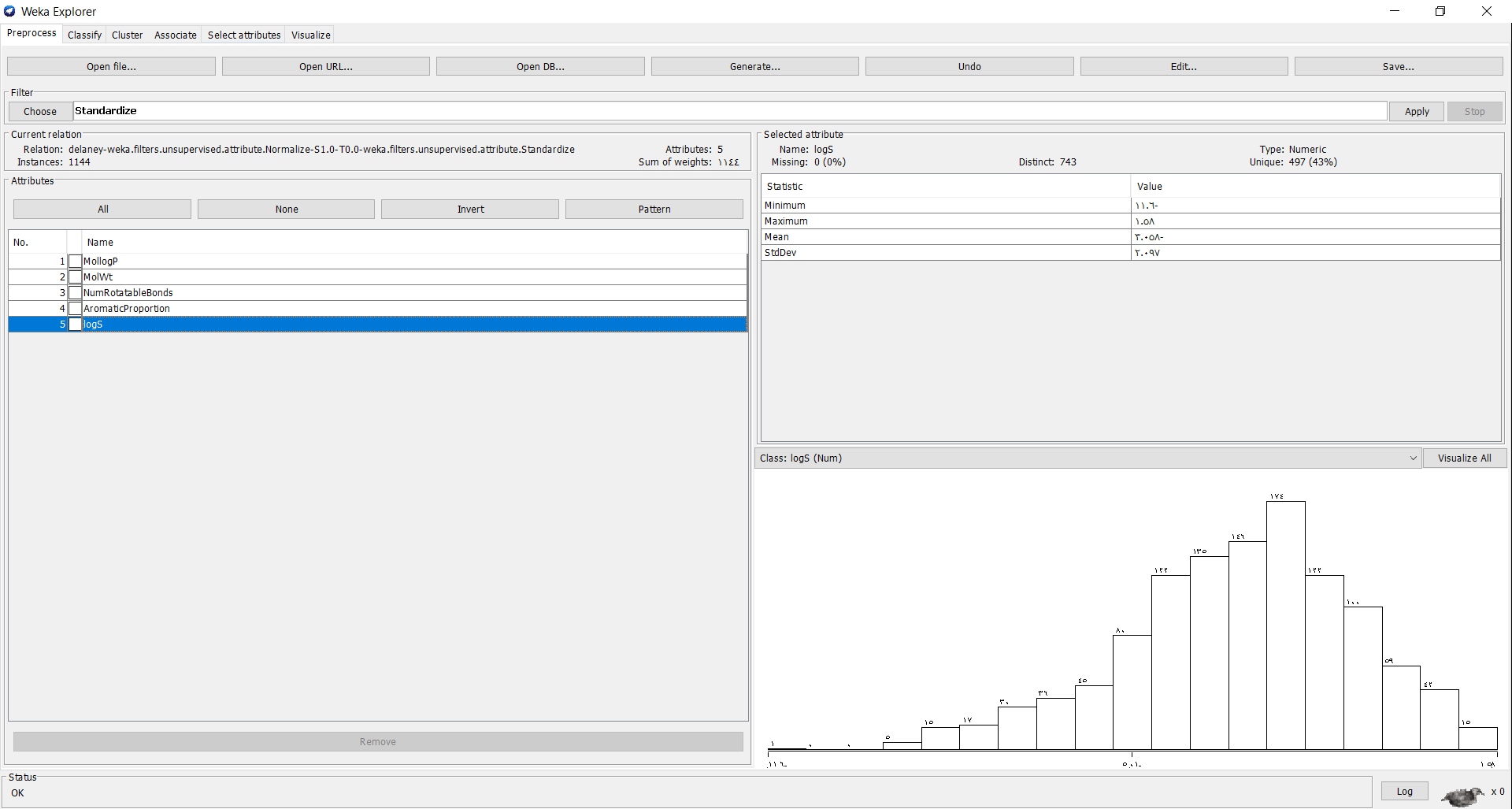
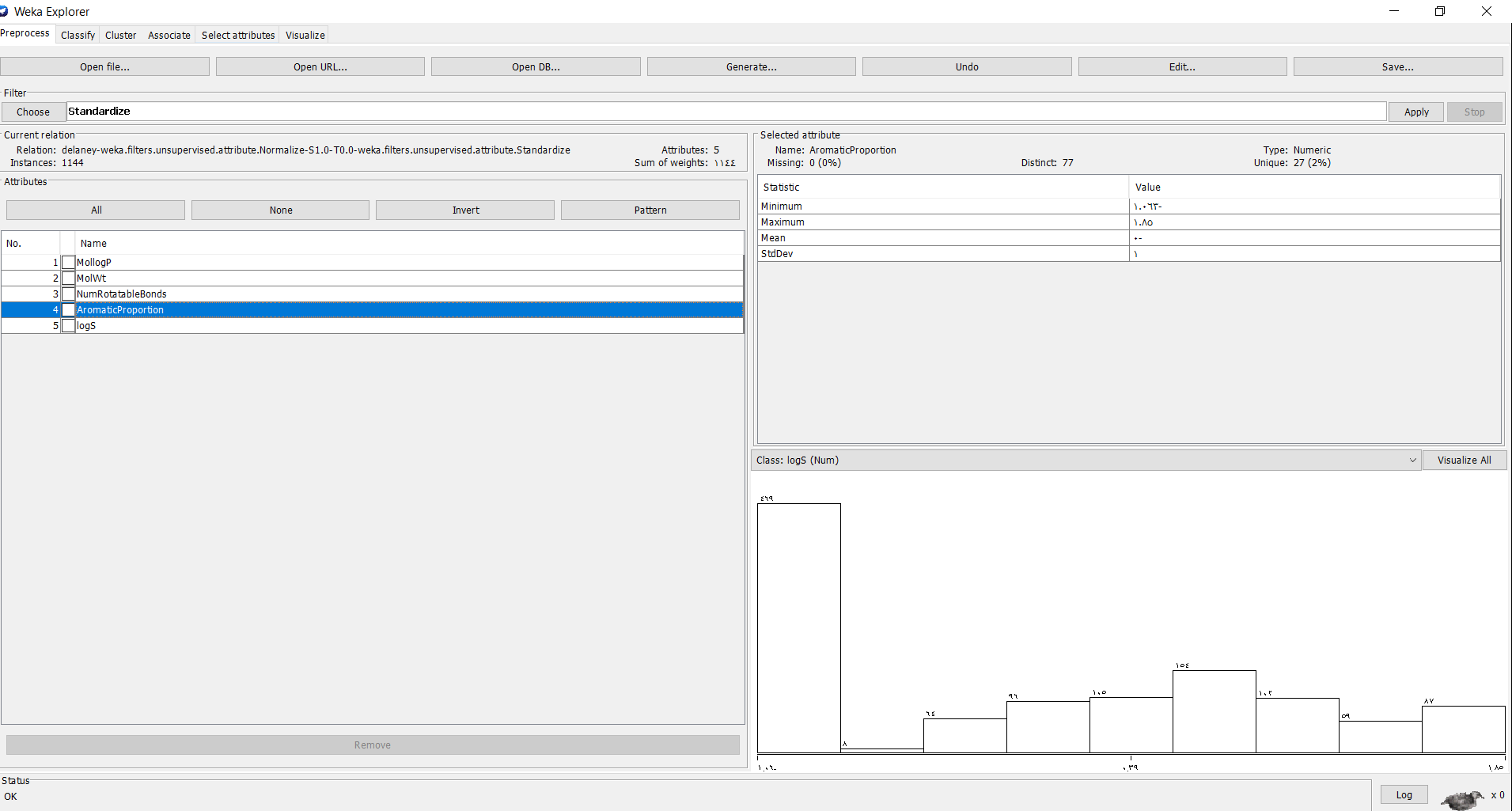
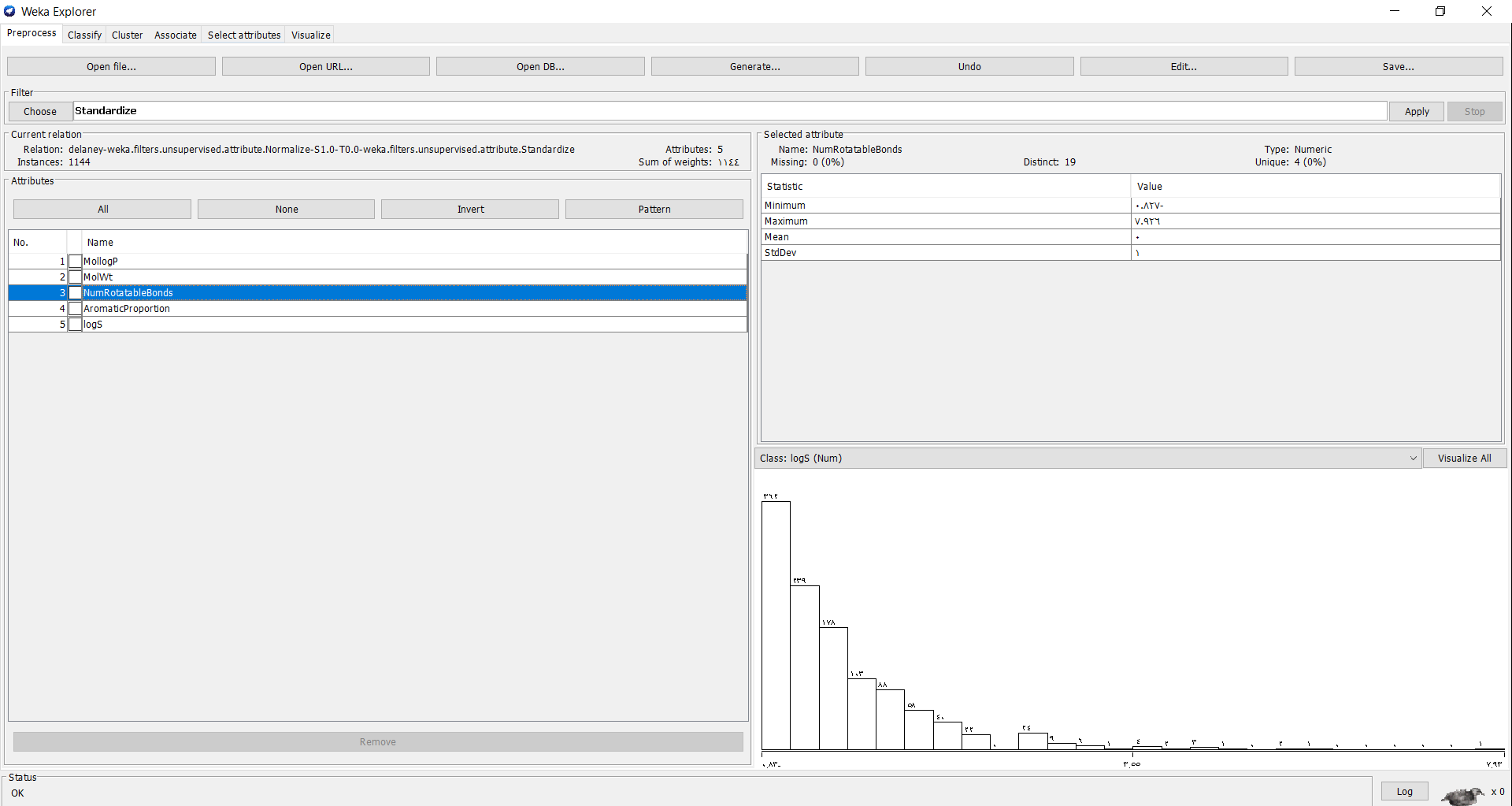
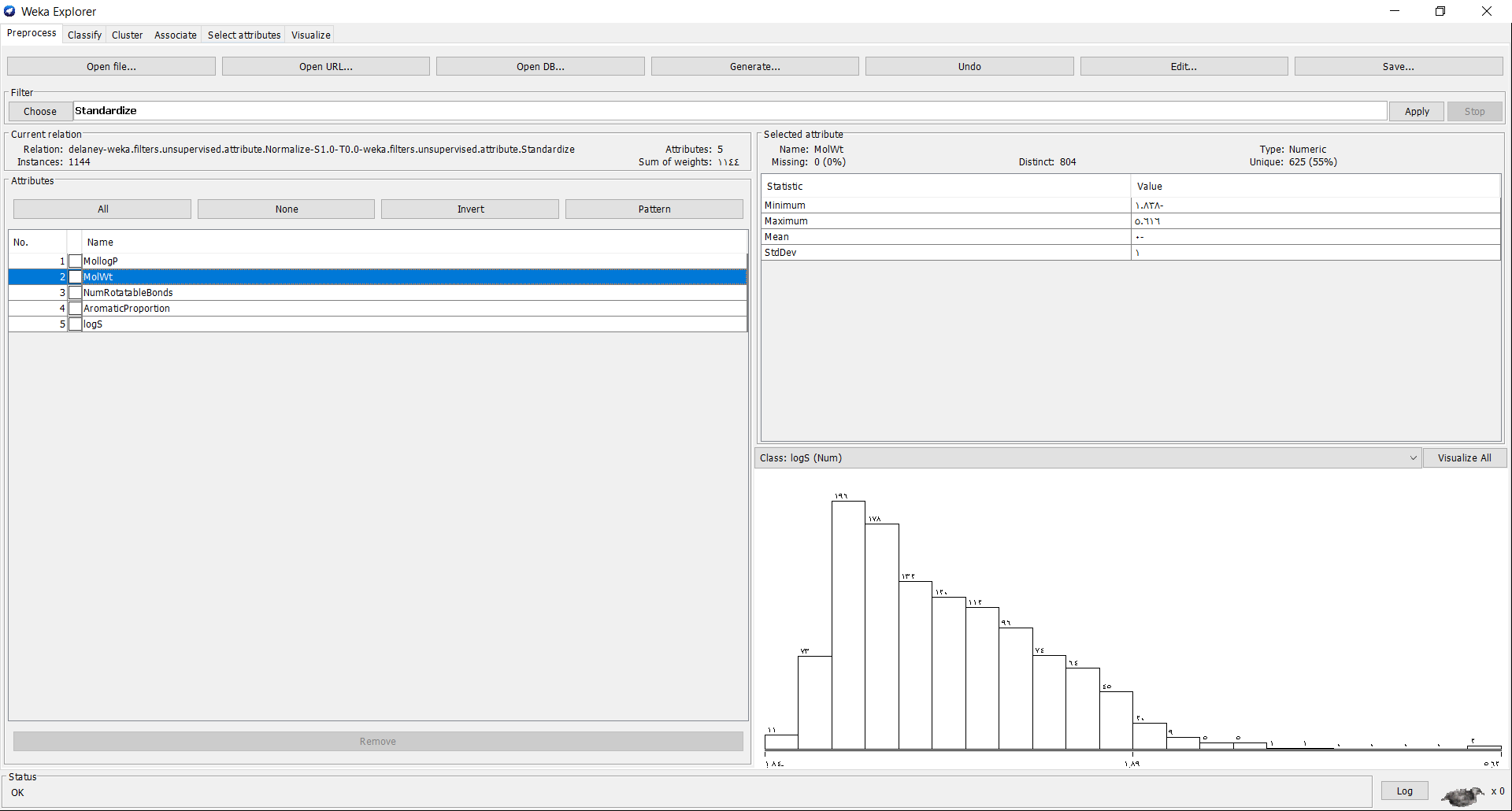
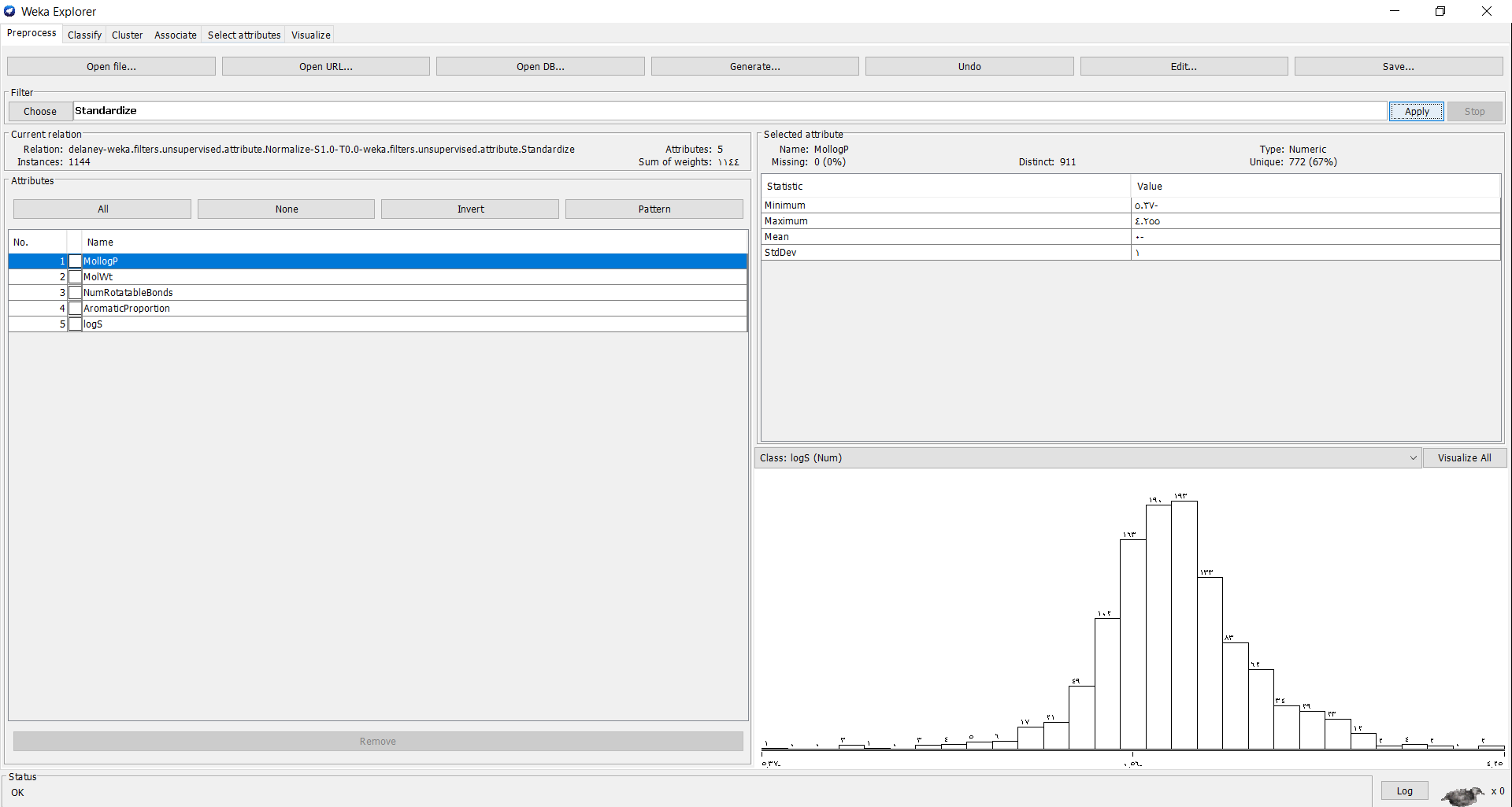
#### [Project](https://olearn.okan.edu.tr/webapps/assignment/uploadAssignment?content_id=_378262_1&course_id=_35503_1&group_id=&mode=view) summery:

The dataset is on delaney sloubility prediction . I used in the project of data mining the technique of classification and an algorithm on this subject is   
used and its performance is tested on the dataset.   
  
The dataset is obtained from github.  
  
  
  
Data preprocessing methods are applied to the dataset such as normalize and standardize.   
  
The chosen algorithm is implemented using Weka software .

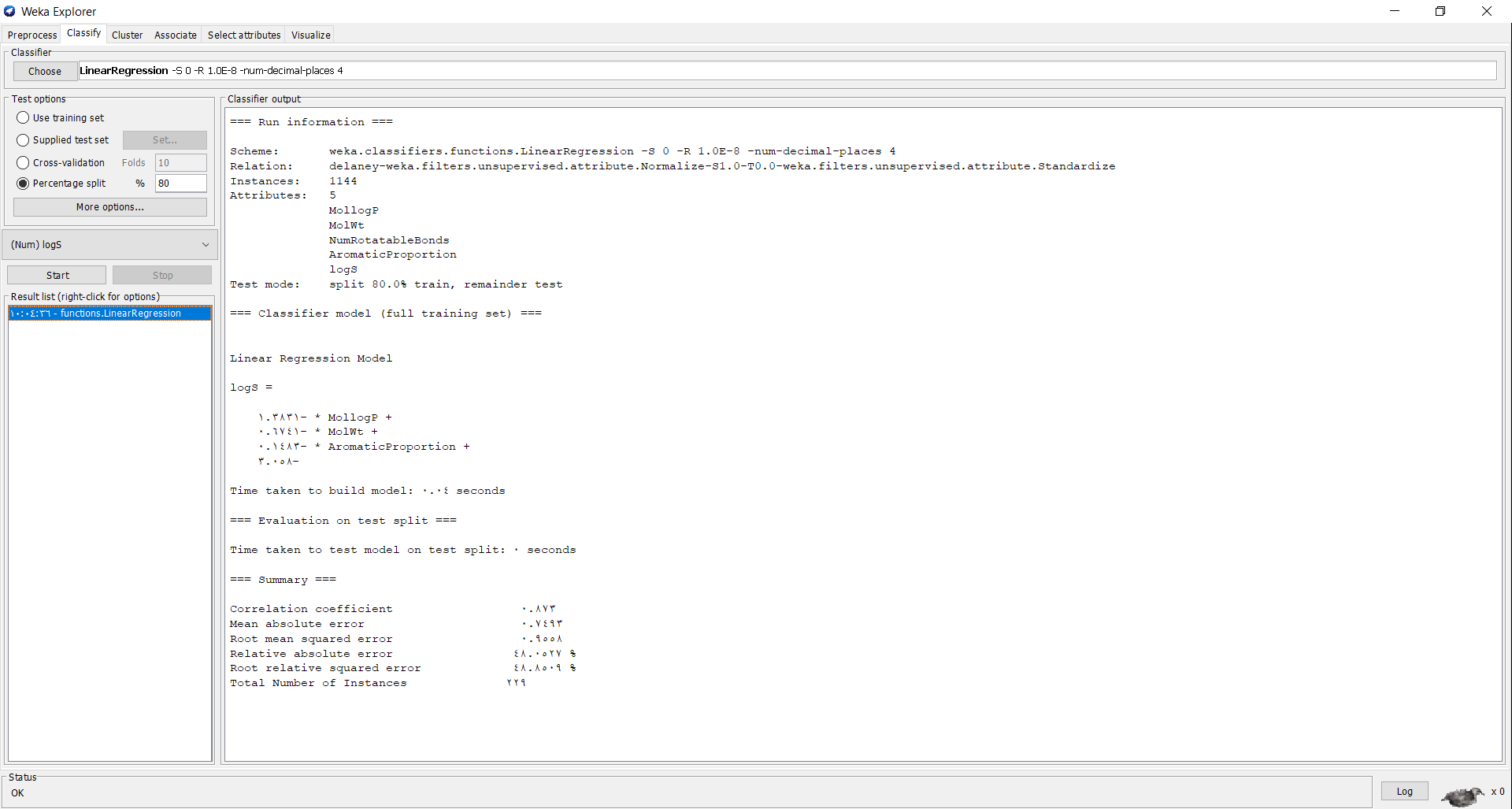
HERE ARE screenshots of the testing :

Preprocessing , normalizing : changes maximum value to 1 and minimum value to 0 for all attributes except the dependent variable that we’re predicting, it will be untouched ****

Preprocessing , standardize : changes mean value to 0 (mean certainty) and stadard deviation value to 1 (unit variance) for all attributes except the dependent variable that we’re predicting, it will be untouched

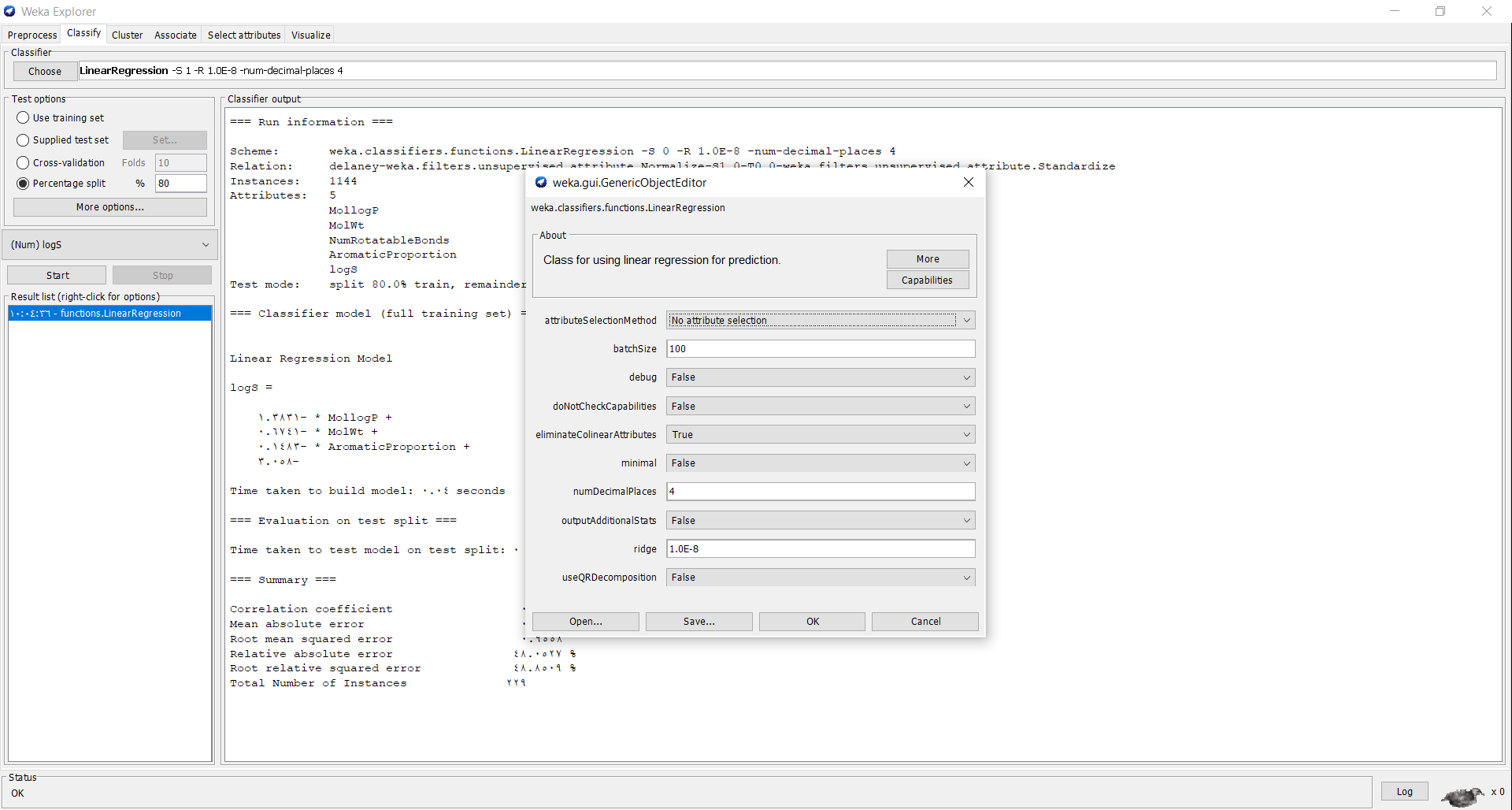
****

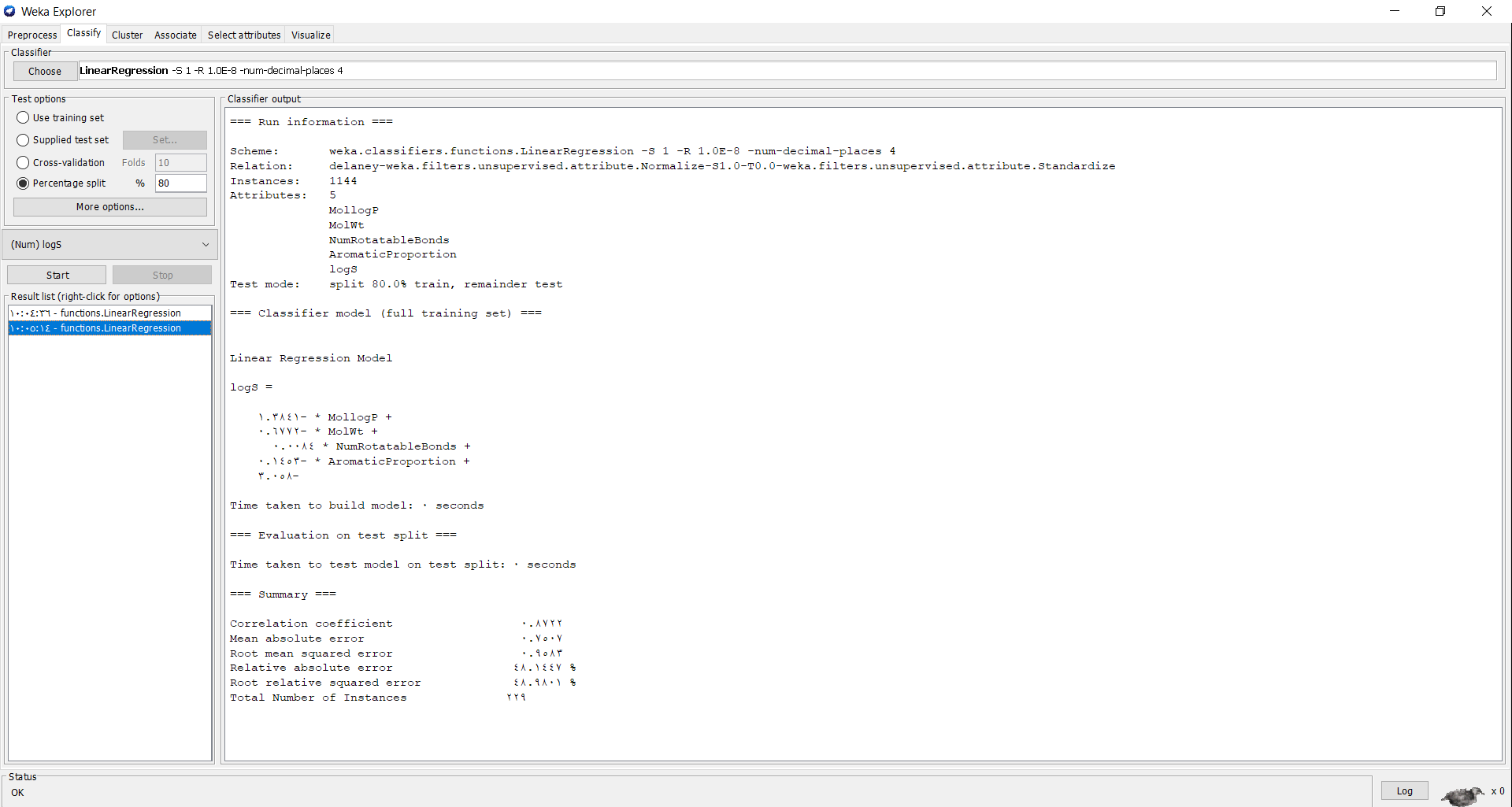
Data Mining Technique , Classification : Linear Regression , percentage split 80%

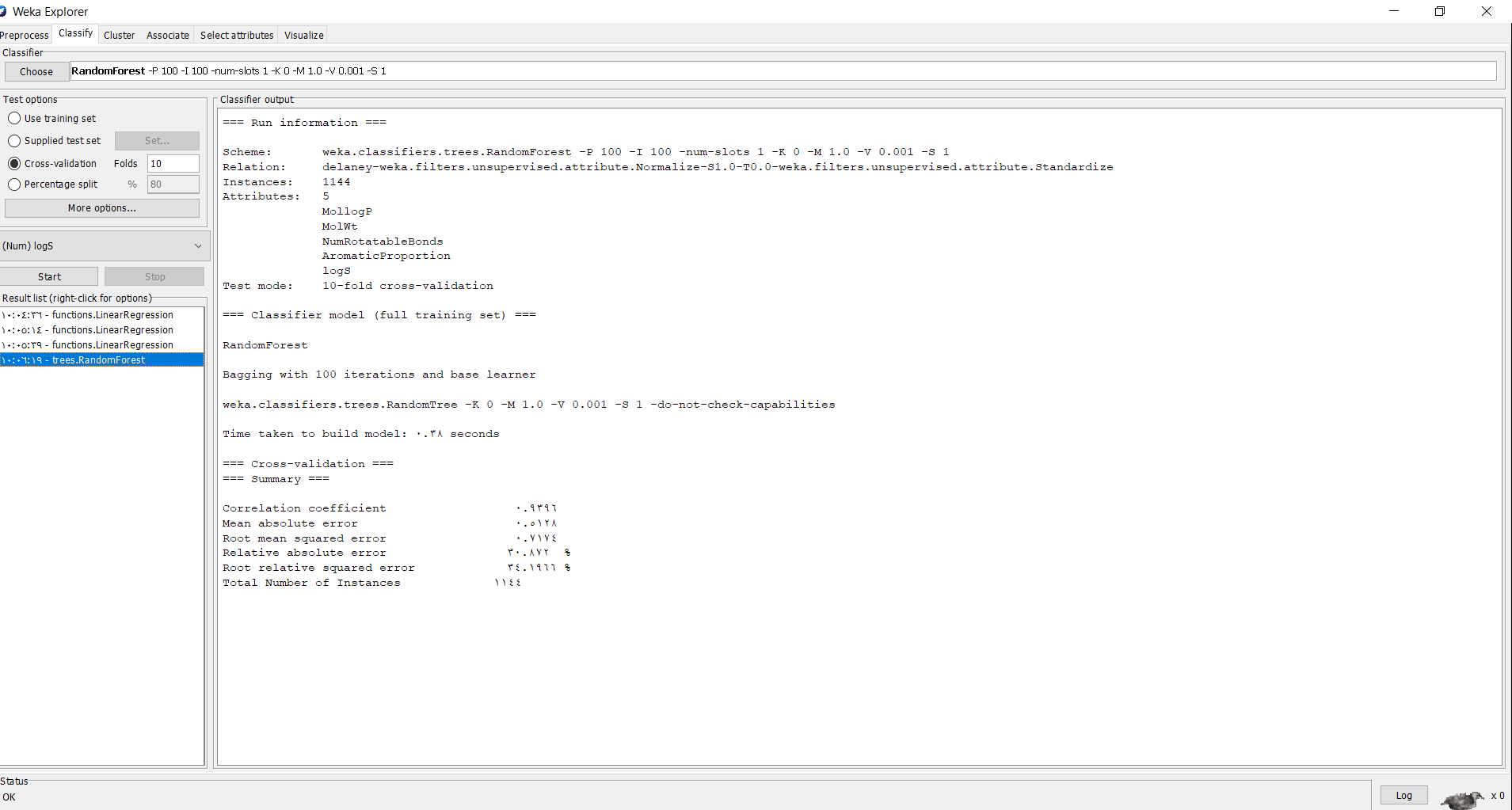
****

Data Mining Technique , Classification : Linear Regression , percentage split 80%

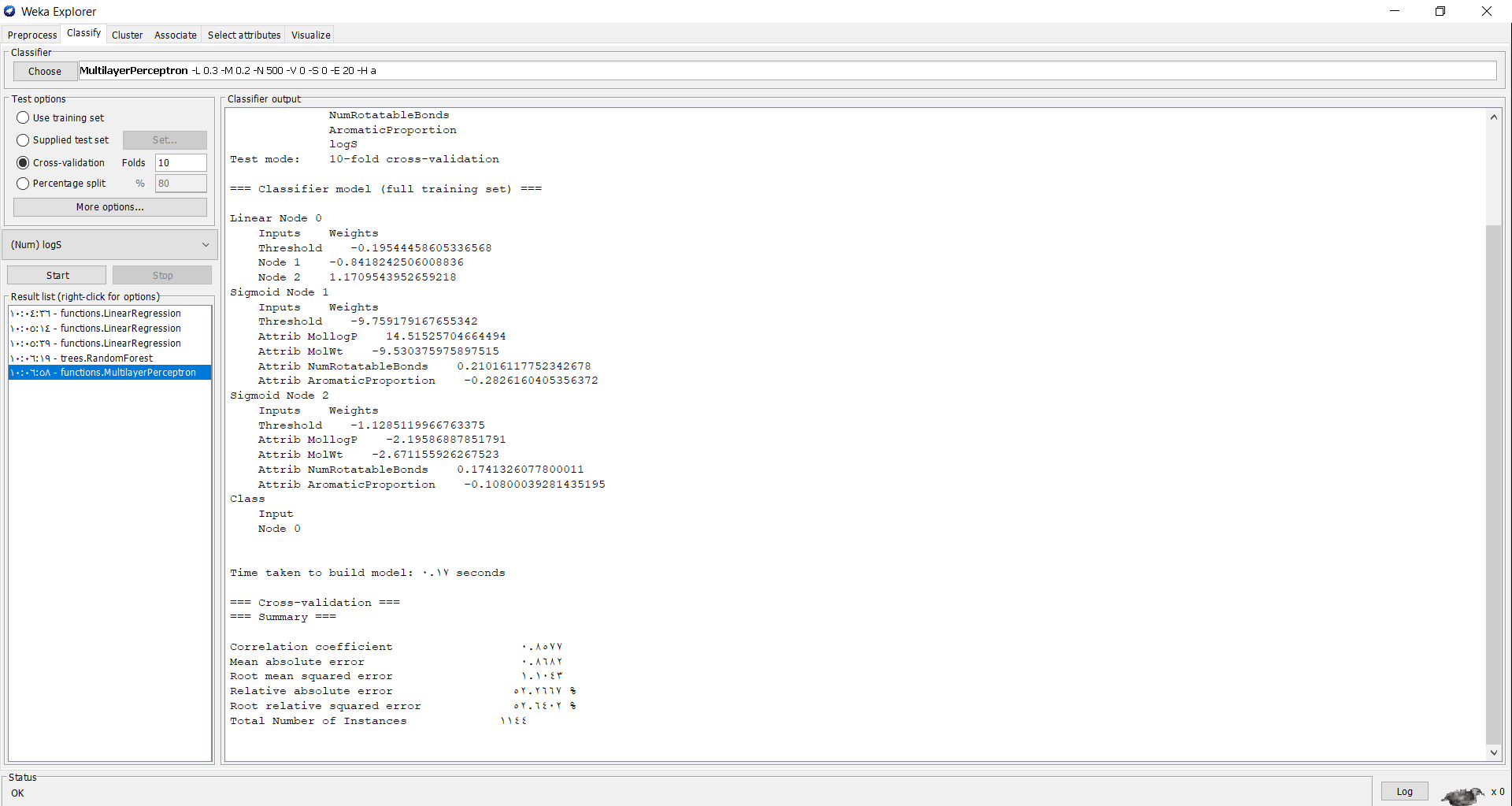
With no selection of attributes :

****

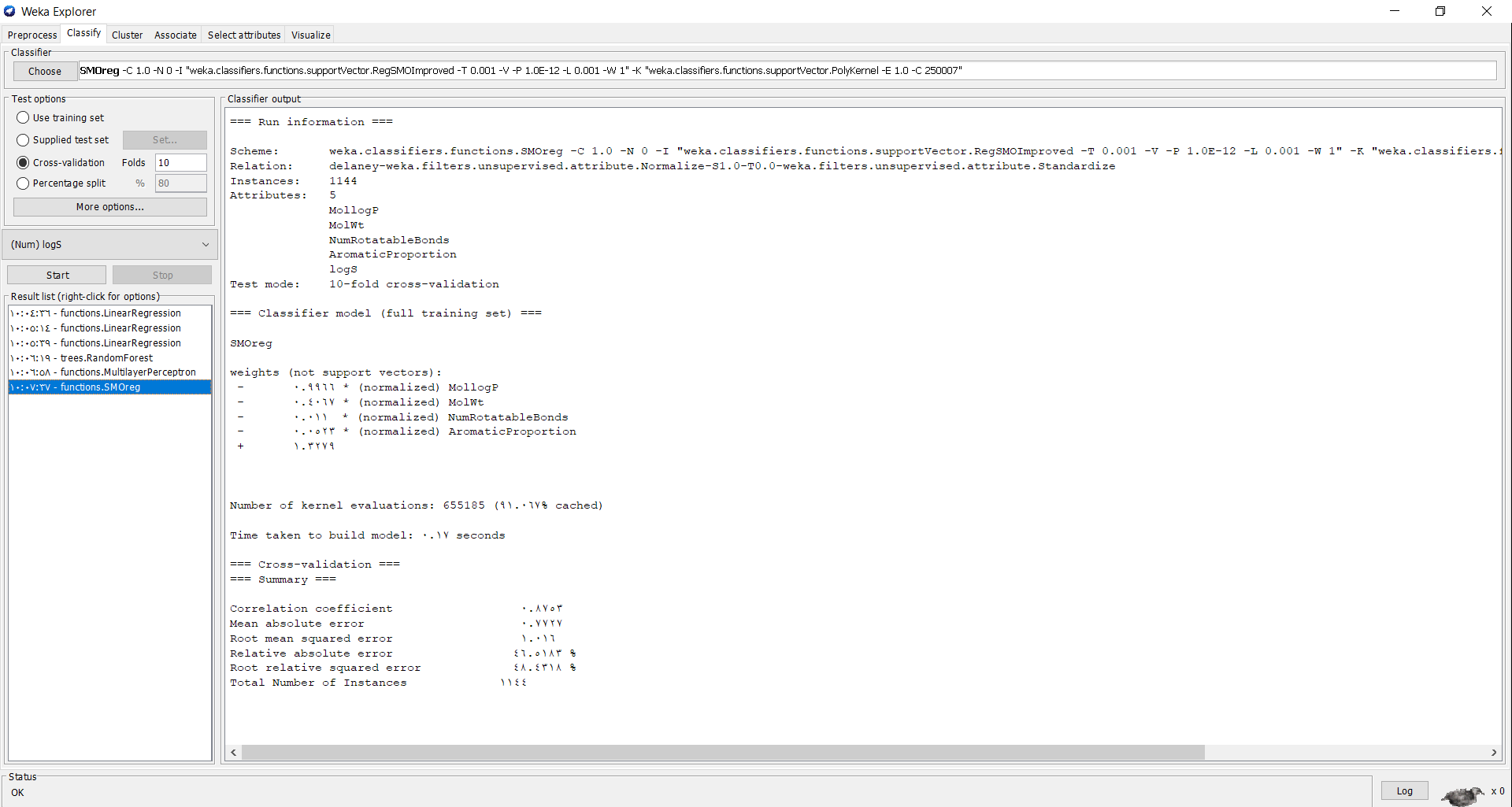
****

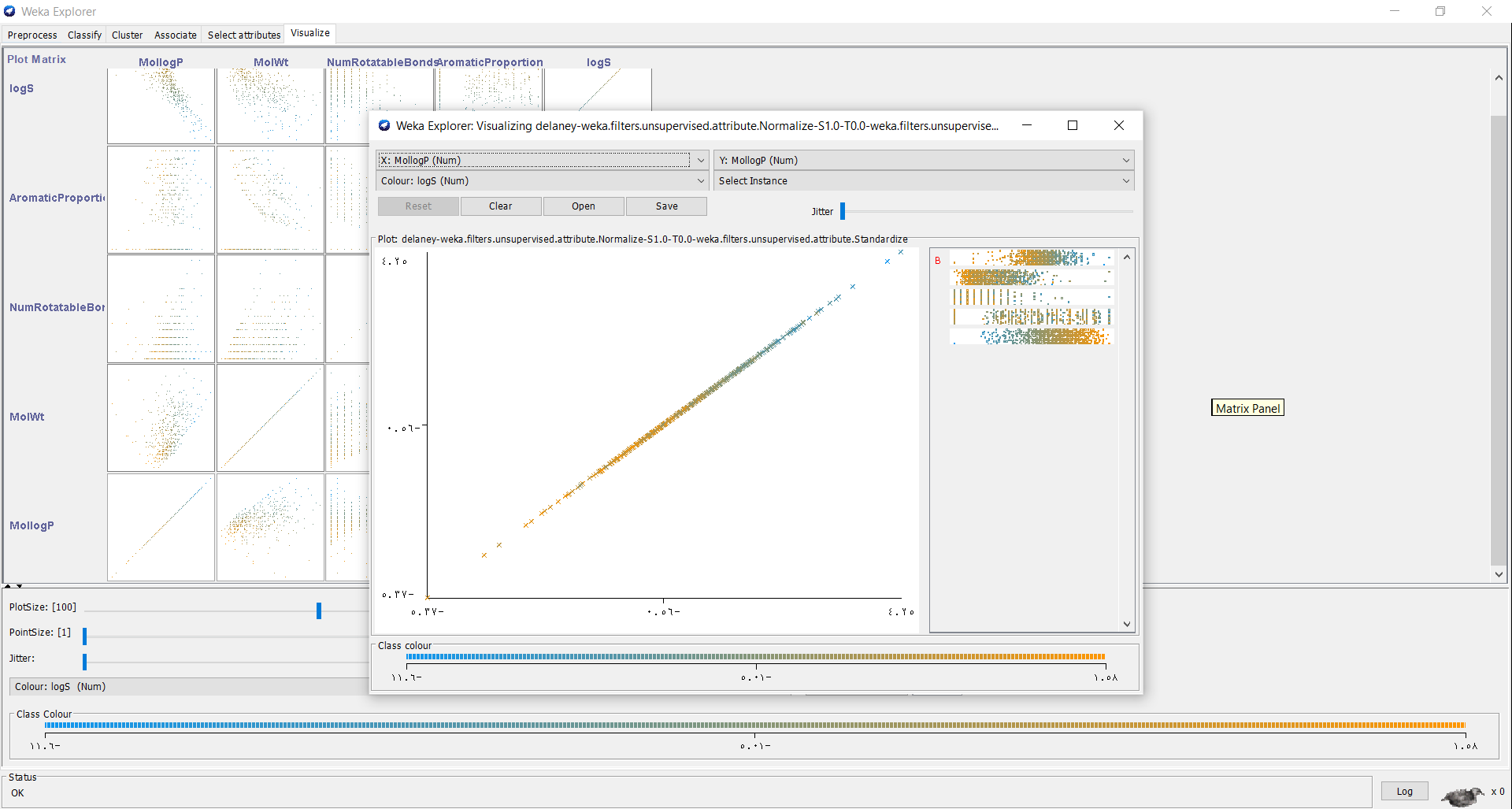
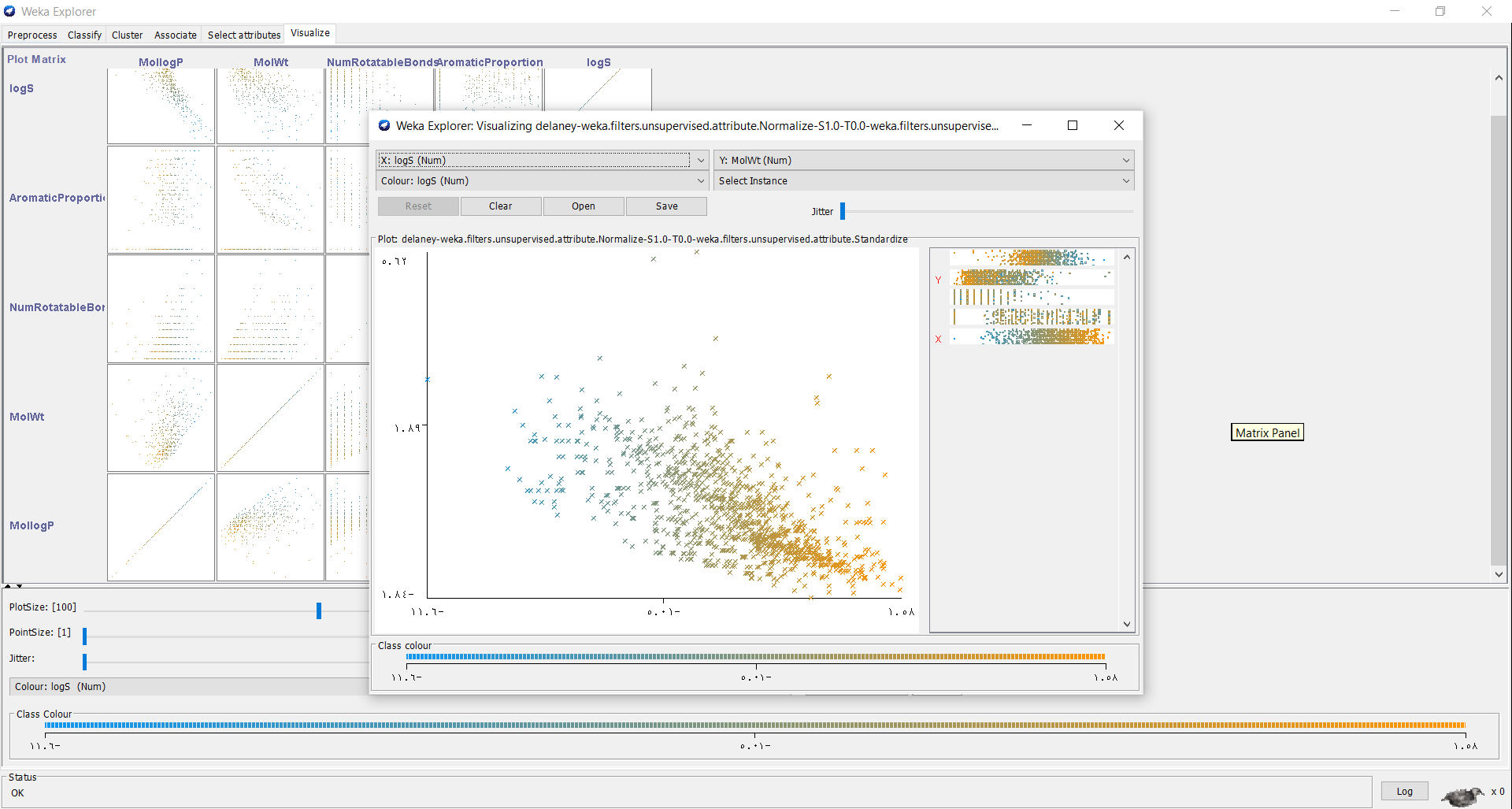
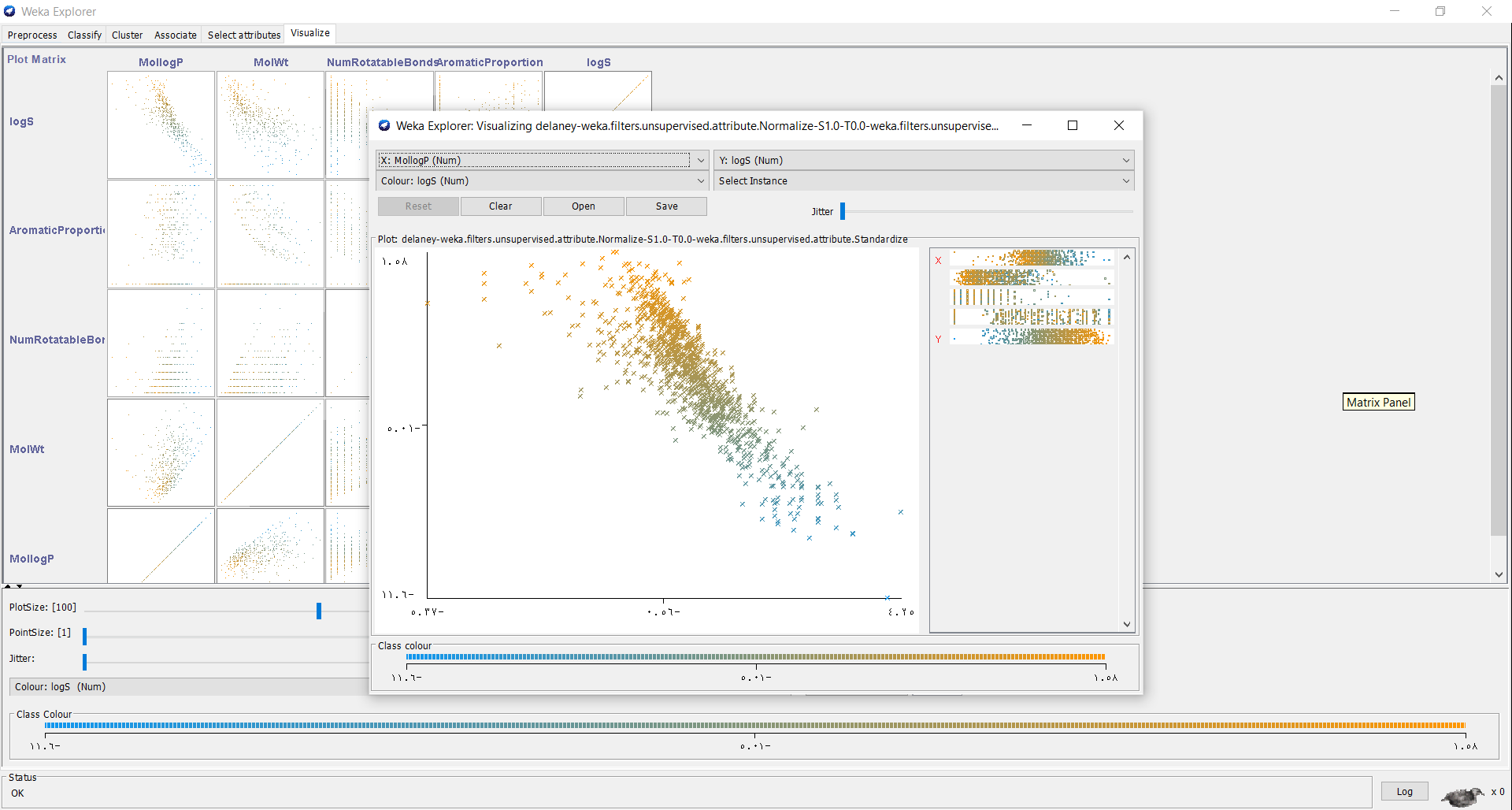
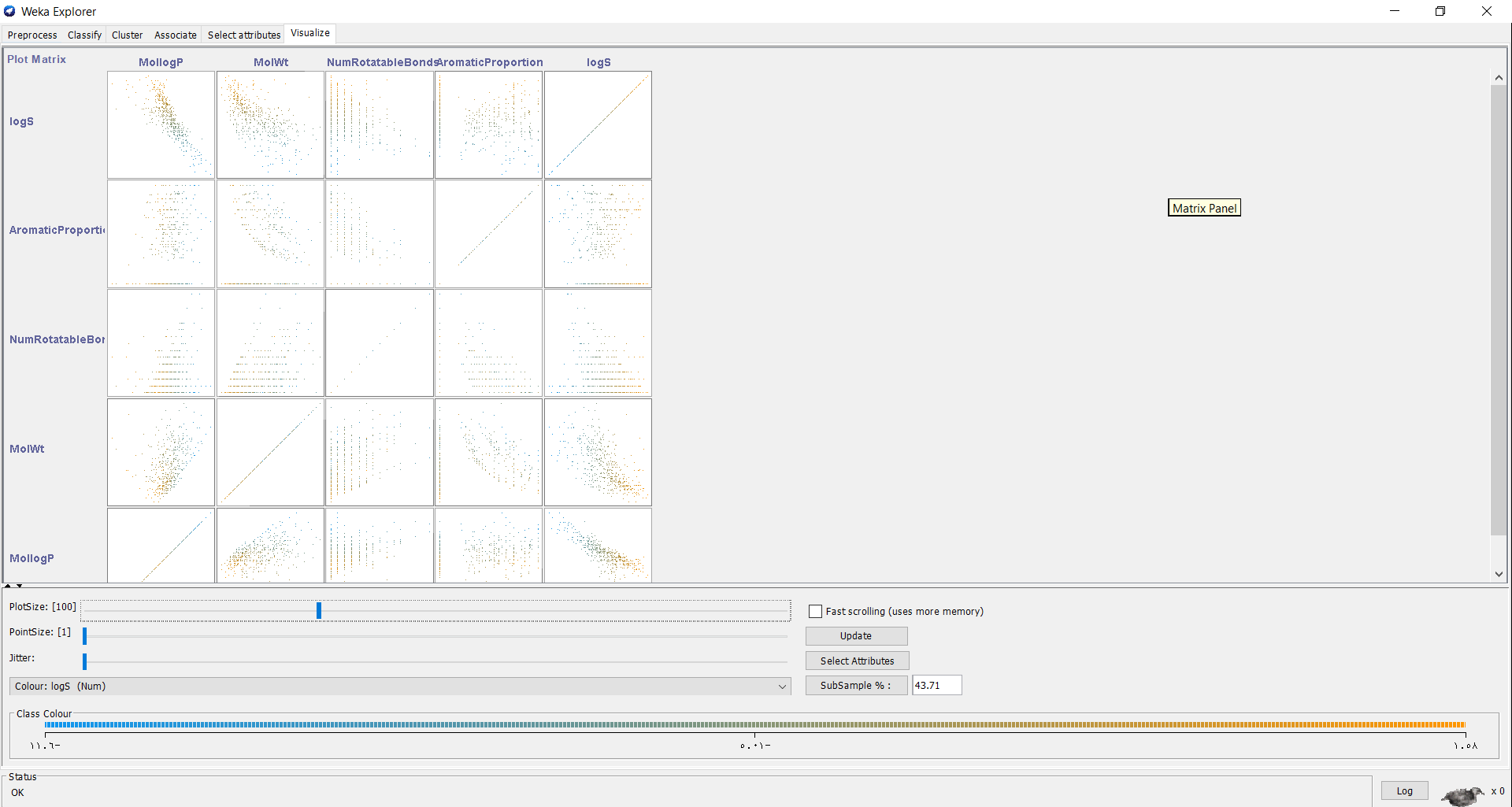
Data Mining Technique , Classification :RandomForest****

Data Mining Technique , Classification :MultilayerPerceptor

****

Data Mining Technique , Classification :SMOreg

****

Data Mining Technique , Visulaization samples ****

**This was my Data Mining Project ,thank you for reading and thank you.**