Readme

This Project Consist of 8 files written in python.

A. Short description of source code files are as follow:

1. **data_preprocessing.py**
   This file contains function-*get_dataset()*- for reading dataset from csv file and pre-processing it. It also contains function *normalize()* to normalize data and return in np.array format.
2. **elbow_method.py**
   This file contains function-*elbow_method()*- it maps within cluster sum of squares values to number of clusters and plots graph for same.
3. **Kmeans_plus_plus.py**
   This file basically generates initial centroid for KMeans clustering. *initialize_random_centroid()* function is used to select 1st centroid and then using *final_centroids()* function other centroid are found
4. **kmeans.py**
   This file contains *k_means()* function which uses above file to generate clusters and return cluster number for data and list of Centroids of each clusters.
5. **plot.py**
   This file basically contains method *plot()* and *subplot()* to draw graph for given data.
6. **encoding.py**
   This file is used to convert continuous attributes to discrete attributes. It has *encoding()* function which does the job of encoding the attributes. It also has a function *write()* to write generated clusters to csv files.
7. **Associatiooin_rule_mining.py**
   This file has function *mine_rules()*, which generates frequent itemsets and association rules for given data. It also has a function minsup() which changes value of minimum support according to the size of dataset.
8. **main.py**
   This is a helper file which imports all other files and makes calls to function of other files in proper order to carry out the required task.

B. Installing required Dependencies [For Windows]:

Note: Python 3 and pip should be installed on the system.

Run following commands in order:

1. pip install numpy
2. pip install pandas
3. pip install sklearn
4. python -m pip install -U pip setuptools
   python -m pip install matplotlib
5. pip install mlxtend

C. How to Run:

1. Open Command Prompt at the location of source files.
2. Enter command- python main.py
3. Press Enter
4. Output Files are generated as notified in the output of the code.