EECS 837 project:

KU ID: 2813662

Name: Ranjith Kumar Sompalli

Please follow the below instructions to execute the project:

1. There are two zipped files included in the mail. The one named as DataMiningProject_Discretization.rar is the eclipse project. If you are running the project through eclipse just extract this rar file and load it into eclipse. Running instructions specified below.
2. The other zipped file named DataMiningProject_Discretization_commandPrompt.rar is for command prompt on Linux/windows machine. If you are running the code on EECS Linux servers, extract this file, go to DataMiningProject_Discretization folder. All the source files are present in that folder. Running instructions specified below.
3. This folder also has two subfolders input and output. All the input files used to test the application need to be placed in DataMiningProject_Discretization /input folder.
4. To run the application:

a. If you are running through command prompt:

   i. Place the input file in DataMiningProject_Discretization /input folder.
   ii. In DataMiningProject_Discretization folder run the command:
   javac *.java
   iii. Then run the application as:
   java DiscretizationAlgorithms
   iv. This starts the application. When asked for input file name, just give the name of the file. The application looks for the file in input folder. (e.g.: filename.txt)
   v. Choose the algorithm you want to run among a, b and c.
   vi. Output files will be stored in DiscretizationAlgorithms/output folder. Two files will be generated one named as test.init which contains the intermediate cut point values and the other test.data which has the final discretized table.

b. If you are running through eclipse:

   i. Import the project extracted from DataMiningProject_Discretization.rar file into eclipse as specified in point 1.
   ii. Place the input file in DataMiningProject_Discretization /input folder.
   iii. All the source files are present in default package in src folder.
   iv. To run the application, run the DiscretizationAlgorithms.java file and provide the inputs as explained above.
   v. Output file will be saved as explained above on the output folder.

4. The application also runs good for large size files. It may take around 10 minutes to generate the rules for large files.