# CS685: Data Mining
# Final Exam: Take-Home: Classification

## Total Marks: 50

## 8:00pm, 22nd November, 2021 – 12:00noon, 23rd November, 2021

This is part of the final examination. So, all the rules of an examination hold here. You are not supposed to use more than the allotted time for this examination.

Please note that there is a separate input file for each student. The files are available from `http://www.cse.iitk.ac.in/users/cs685/classification/`

If your student id is `xxx`, then you should **only** use the file `training-sxxx.csv` and nothing else. You can find your student id from the course website: `http://www.cse.iitk.ac.in/users/cs685/classification/students.csv`

There are 411 rows in the training file. Train a classifier using it. The first attribute is the class label.

You should submit an **executable file** that should run on the test file (which will be 100 rows) to produce an output of 100 lines with only the class label. The value should be marked as $C0$ or $C1$ or ..., etc.

Submit a file called `classifier-sxxx.sh` that corresponds to the answer text file. Replace `xxx` with your student id.

A sample truth file is given as `answer-sample.csv`. Your answer should follow the same format.

Also, submit separately all code, readme, etc. as a single zip file named `rollno-classification.zip`.

The file should be unzipped as a folder. The folder must contain a `Makefile` and a `Readme.txt` that will contain the instructions on how to run your classification detection method. It should **also** contain the file `classifier-sxxx.sh`.

The entire setup should run on a Linux operating system environment.

There is a *separate* truth file for each classification file. The results of your analysis will be run against the corresponding truth file only. Your score will be the **classification accuracy** of the test file rows times the maximum marks.

Please also note that your answer file will be run against all truth files. If it is found that your answer is giving *better* results with the truth file of another student, *both* of you will be *suspected of cheating*. And then, if it is established that you have cheated, you will get an **'F' grade** in the course, with possibly *stronger actions*.