Kaggle Competition II
CSL2010: Introduction To Machine Learning
AY 2021-22

Competition Link : https://www.kaggle.com/t/04a239e37f69de0040810bc615b96063

Timeline:
Start Date: November 03, 2021, 18:30 IST
End Date: November 10, 2021, 23:59 IST

Problem Statement:
The goal of this competition is to work on the problem of label ranking. The label ranking problem can be posed as an extension of many supervised tasks such as multi-label classification or multi-label ranking. Surely, post-processing the output of a conventional classification model can help obtain the predictions for the tasks mentioned above.

You are hereby required to train a classifier that best fits the training samples and report the top-5 predictions on the test data.

Dataset Description :
1. The training data contains 10K samples, and each sample is represented by a 200-D feature vector. These are provided in the file "TrainFeatures.txt" (one feature vector per line).
2. There are 100 classes (or labels/categories) in the data set. The file "TrainClasses.txt" contains the class of all the training samples (one per line).
3. The file "TestFeatures.txt" contains the test data, with each line denoting one test sample represented by a feature vector.

Model Description:
Design an appropriate classification model using training samples and report the top-5 predictions on the test data provided to you (refer the sample output).

Submission Guidelines:

- You can make submissions in **.csv format** up to **3 times per day.**
- The final predictions from the classifier must be saved in a .csv file with the following format: (available in the Competition's data section , **SampleSubmission.csv**)

index	labels
0	1 2 3 4 5
1	6 7 8 9 10

Model Evaluation:

We will be using the **mAP@5** metric to evaluate the model.

Report Submission:

- You also need to submit a report by November 10, 2021, 23:59 IST. Name your report with your Roll No. For example, B20XX123.pdf
- Your report must include the following sections:
    - Model building and validation
    - The models you explored, your choice of selection, and how you arrived at your final model for the submission
- The page limit for the report is set to 5, including all the tables and graphs. Any content exceeding this limit will be ignored.
- You should use a normal A4 layout of Microsoft Word with the font as Calibri and font size as 11. For headings and titles, you may use bold and font size as 12. Also, write your name and roll no. in the header section.
- No need to upload your final code in google-classroom.

Important:

- Upon signing up for Kaggle and joining the competition, you need to change the display name in your Kaggle profile to your roll no., e.g., B20XX123, and profile picture to a picture of yours.
- Do not sign-up using more than one Kaggle account.
- This is a competition among individual participants. You should work on this problem on your own.
- Read the competition description and other instructions carefully.
- In case of doubt, drop a public comment in google classroom, and don't email the instructors/TAs individually.