Machine Learning and Data Mining Lab

Practice Lab-2

22 Jul 2020

Instructions

- Dataset contains some binary features. The objective is to compute pairwise similarity.

- Note that binary feature can be treated as a special case of nominal attribute

- $d(i,j) = (p-m)/p$, where p is total number of attribute, m is number of matches

- Load the data set: Here

- Separate Features and Label (label is first column)

- What is the number of features?

- What is the order of the data matrix?

- Write a function to compute pairwise similarity

- Compute pairwise similarity for all pairs

- Write output to a csv file

- Upload output CSV and your notebook (as zip file) to Teams assignment link

- Due date: 23 Jul 2020, 11:50 PM