

Laborator 8

Folosind un set de date de la <http://archive.ics.uci.edu/ml/datasets.html?format=&task=&att=&area=&numAtt=&numIns=&type=text&sort=taskUp&view=table> (<http://archive.ics.uci.edu/ml/datasets.html?format=&task=&att=&area=&numAtt=&numIns=&type=text&sort=taskUp&view=table>) sau resurse indicate in cursul 8, sa se rezolve o problema de clasificare a textelor.

Se poate opta pentru codificare BOW, n-grams, word2vec sau altele adecvate; clasificarea se poate face folosind knn cu distanta cosinus, naive Bayes, logistic regression sau alte modele de clasificare.

Prezentarea temei se va face in saptamana a zecea.