

# FUZZY CLUSTERING

## *UVOD:*

Cilj ovog seminarskog je bio da klasterujemo dokumente primenom fuzzy klasterovanja. Za potreba klasterovanja koristio sam skup podataka koji se nalazi u git repozitorijumu, a takodje inspiracija za implementaciju je bio naucni rad koji se nalazi isto u repozitorijumu.

## *OPIS SEMINARSKOG:*

Implementiran je fuzzy c-means algoritam koji koristi fuzzy logiku I c-means algoritam kao sto I sam naziv kaze. Ideja je bila da se izvrši klasifikacija reci iz skupa dataset pomocu implementiranog algoritma. Algoritam moze da se pokrece za proizvoljan broj klastera. Kao mera bliskosti je korisćena kosinusna mera, zato sto je ona pogodnija mera ako radimo sa tekstualnim podacima kao sto je slućaj u ovom radu. Takdoje imamo I klasu koja se bavi preprocesiranjem reci u svim dokumentima.

## *REZULTATI SEMINARSKOG:*

Za razliciti broj klastera dobio sam razlicite rezultate posto je skup statican moze se probati razlicit broj I videti koja je verovatnoća pripadnosti dokumenta kom klasteru. Algoritam je implementiran na Ubuntu 20.04 operativnom sistemu, koriscen je Python 3.8.5, takodje je nepohodna numpy biblioteka za pokretanje.

## *ZAKLJUČAK:*

Dalje unapredjivanje bi moglo bit u pravcu koriscnje nekih drugih mera za inicijalizaciju matrice, I racunanje rastojanja.

## *LITERATURA:*

1. Rad.pdf (koji se nalazi u ovo repo-u)
2. Skup podataka sa sajta asistentkinje Mirjane Maljkovic