## KMeans algoritam

### Način implementacije algoritma

1. Korak – Nasumičan odabir instanci dataset-a koje predstavljaju početne centre klastera (Figure 1)



Figure 1

Zatim se formira while petlja u kojoj se primenjuju koraci dati u nastavku:

1. Korak –Za svaki uzorak se određuje koliko je udaljen od svakog centra (Figure 2)

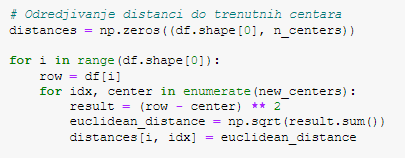


Figure 2

1. Korak – Određivanje koji centar je najbliži (Figure 3)

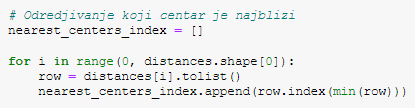


Figure 3

1. Korak – Određivanje novih centara (Figure 4)

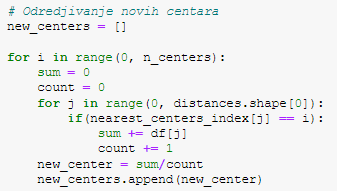


Figure 4

1. Korak – Određivanje da li je došlo do konvergiranja (Figure 5). Do konvergiranja je došlo ukoliko se centri nisu promenili u odnosu na prethodnu iteraciju.

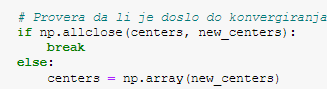


Figure 5

### Skup podataka

Skup podataka o cvetu irisa je multivarijantni skup podataka. Skup podataka se sastoji od 50 uzoraka iz svake od tri vrste perunika (Iris Setosa, Iris virginica i Iris versicolor). Iz svakog uzorka merene su četiri karakteristike: dužina i širina listova i latica, u centimetrima.

### Zaključak

Nakon primene algoritma KMeans sa zadatim parametrom n\_centers = 3, svi podaci se grupišu u tri klastera. Pošto je izbor početnih centara nasumičan onda će se sa svakim pozivom algoritma formirati drugačiji klasteri. U nastavku je data 2D i 3D vizuelizacija dobijenih klastera za jedan od poziva algoritma. Crvenim krstićima su označeni centri klastera

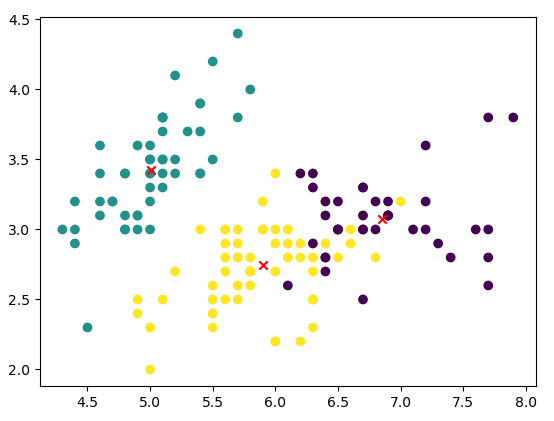


Figure 6

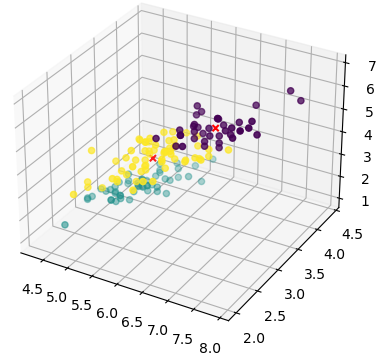


Figure 7