

# STROJNO UČENJE

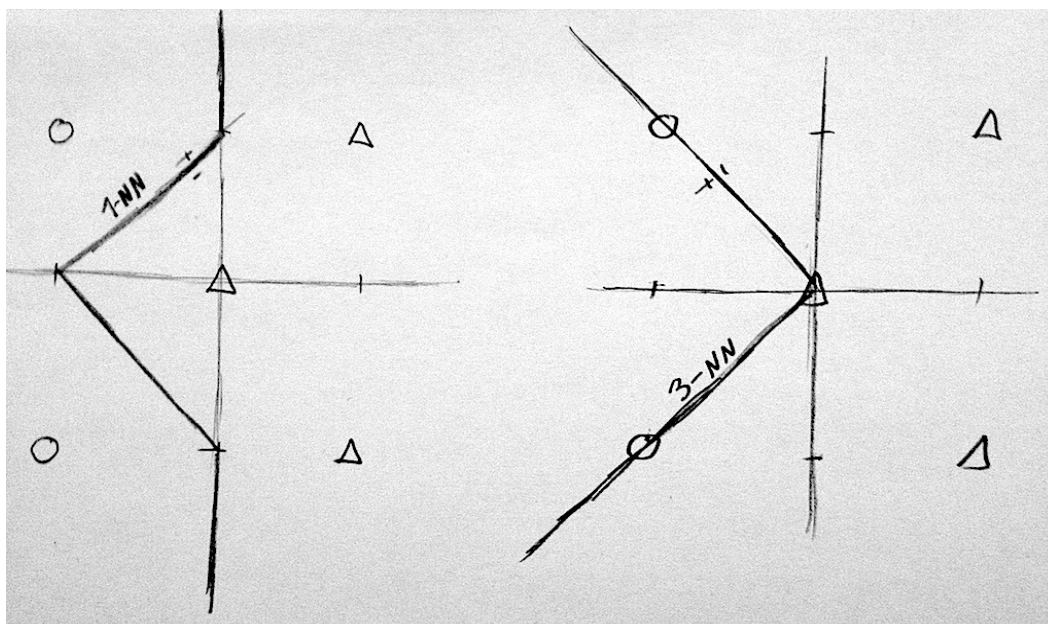
## 4. Domaća Zadaća

Krešimir Špes

0036419866

ak. god. 2011. / 2012.

a)



b) 3nn koristeći euklidsku udaljenost i jaccardijevu udaljenosti u oba slučaja klasificira 8. primjer kao pozitivan. Za potrebe ovog zadatka napravio sam skriptu u pythonu koja to izračunava. Datoteka se nalazi u zad2/zad2\_b.py

primjer izlaza skripte:

```
euklidska udaljenost: 3nn: ( 2 1 [[3, 2], [0, 4], [1, 6]] ) klasifikacija: 1
jaccardova udaljenost: 3nn: ( 2 1 [[3, 0.33], [0, 0.57142], [1, 0.75]] ) klasifikacija: 1
```

c) euklidska udaljenost će biti 1 a jaccardovu nije moguće odrediti bez konkretnih vektora jer udaljenost ovisi o broju jedinica u oba vektora.

Pa na primjer, razlike u prosjeku ne želimo tretirati jednako jer očito da razlika između niskog i visokog prosjeka nije ista kao i razlika između niskog i srednjeg prosjeka. Jedna od ideja ako bi koristili binarni vektor jest da napravimo za svaku značajku matricu udaljenosti između pojedinih značajki i to koristili kao mjeru udaljenosti, ili za množenje mjere udaljenosti.

Ako ne bi gledali te značajke kao binarne vektore mogli i pretvoriti značajke u brojeve i rasponom brojeva korigirati udaljenosti na koje utječe ta značajka. Npr niski prosjek bi bio -1, srednji 0, a visoki 1.