

ZAVRŠNI RAD br. 6985

# Identifikacija ponavljajućih uzoraka kod mjerenja u testnim platformama

---

Jan Čapek

MENTOR  
doc. dr. sc. Marko Đurasević

# Problem

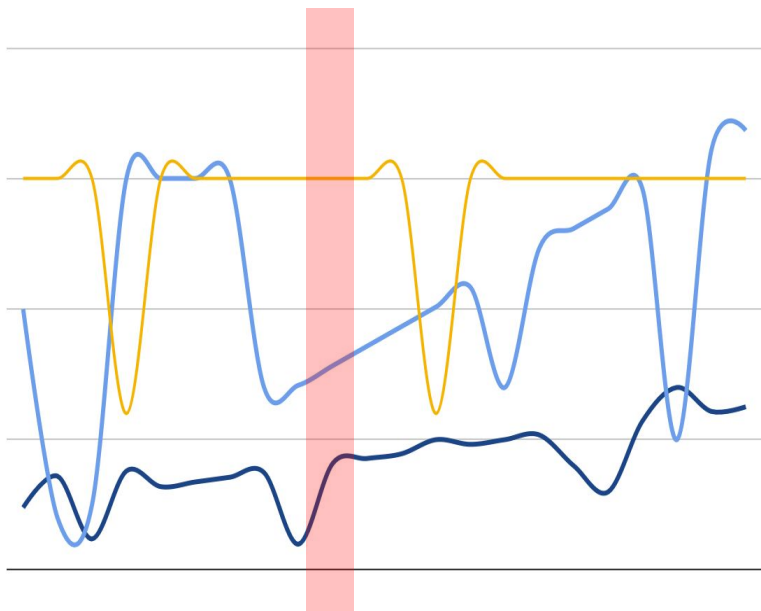
- Testiranje motora u dvije faze
  - Kalibracija motora s različitim početnim stanjima
  - Stabilnost
- Obe faze traju jako dugo

RJEŠENJE

=

utvrditi stabilnost iz prve faze

# Podaci

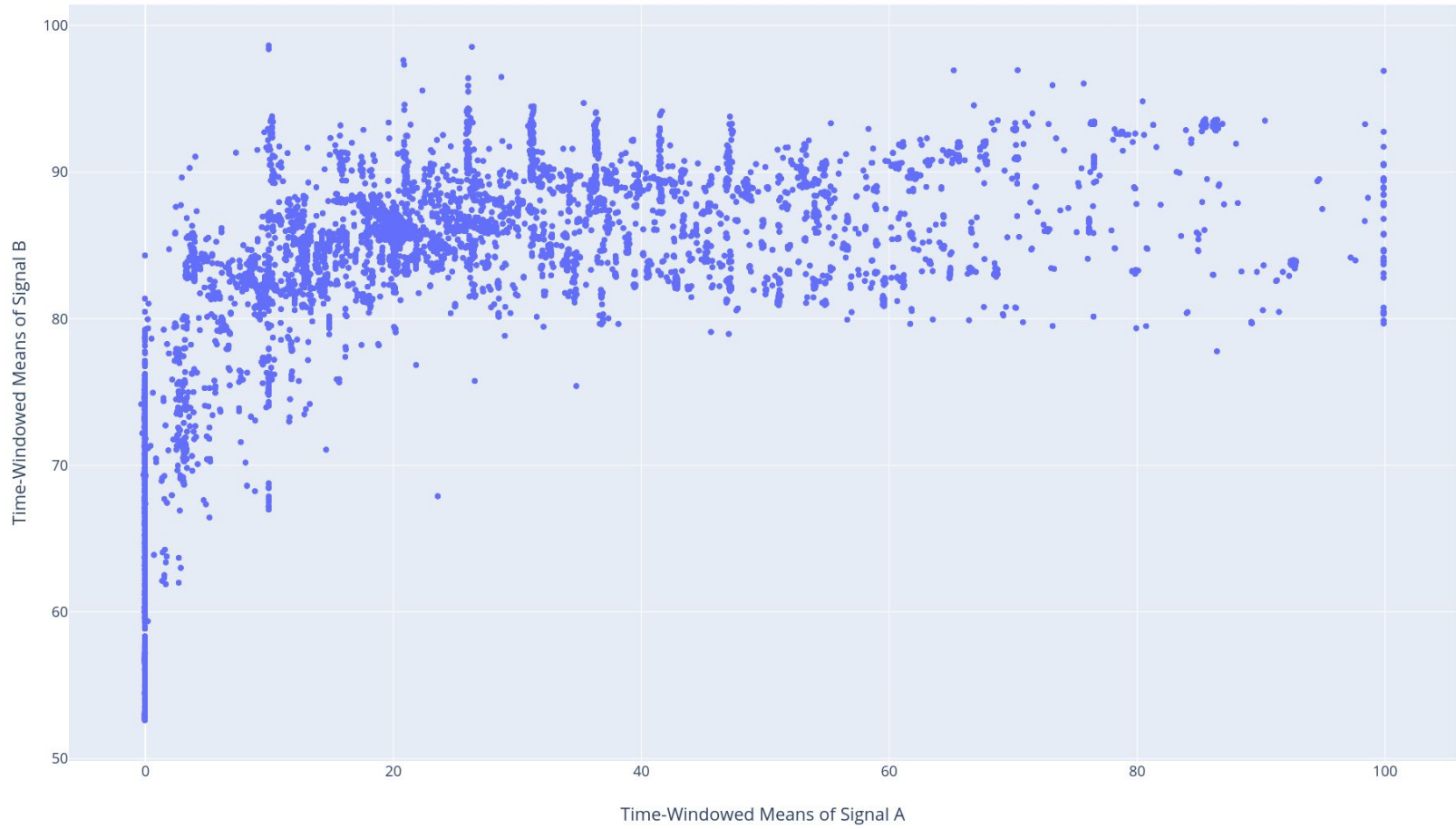


UZORAK

=

uređena N-torka srednjih  
vrijednosti dijelova signala

$$p_i = (m_1, m_2, \dots, m_N)$$



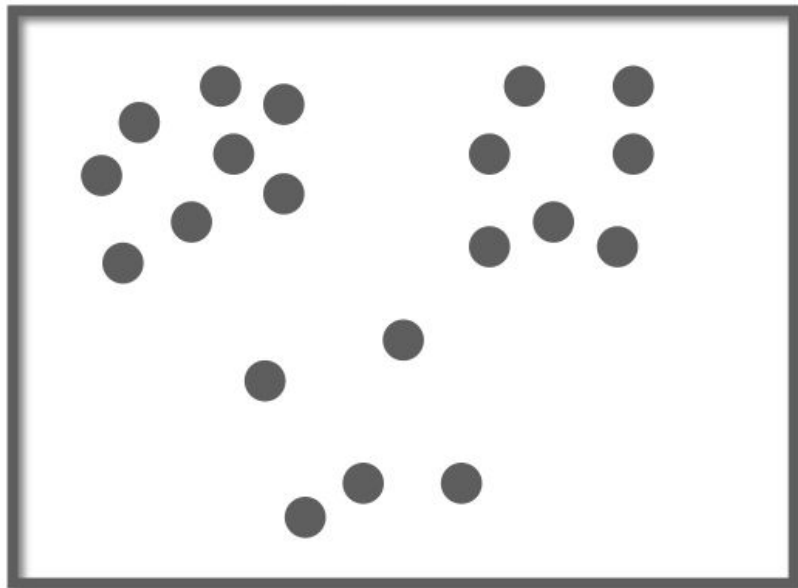
# Algoritmi grupiranja

# K-Means

# K-Means

- Baziran na udaljenosti
- Parametri
  - Broj željenih grupa
- Problemi
  - Potrebno unaprijed znati broj grupa
  - Nedeterminističan rezultat
  - Grupe ne mogu poprimiti kompleksne oblike

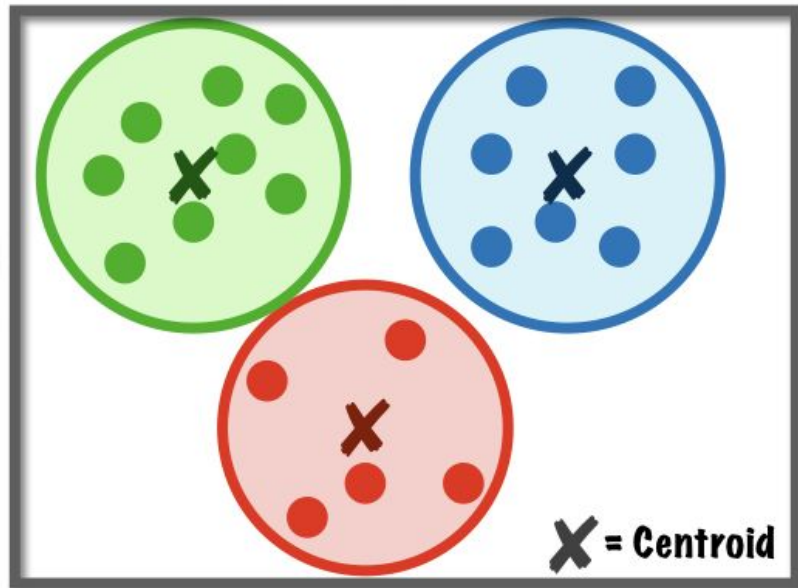
## Unlabelled Data



**K-means**



## Labelled Clusters



By Alan Jeffares -

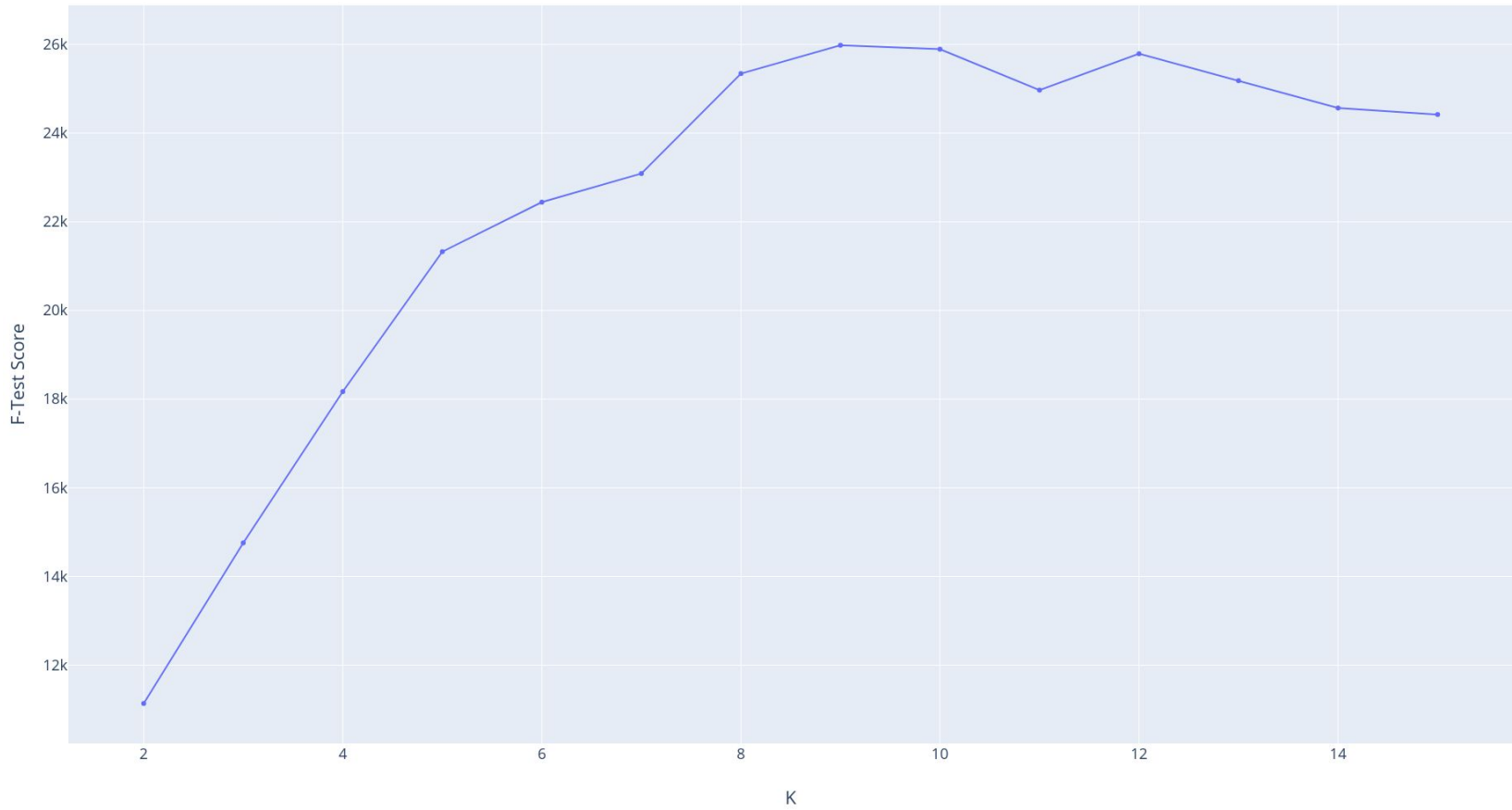
<https://towardsdatascience.com/k-means-a-complete-introduction-1702af9cd8c>



# Metoda lakta

- Heuristička metoda za pronalazak optimalnog broja grupa
- Iz grafa se pronalazi zglob te se odabire broj grupa koji odgovara koordinati zgloba
- Osi grafa
  - X - broj grupa
  - Y - odnos između varijance svih grupa i varijance unutar grupa (ANOVA F-Test)

$$F = \frac{\text{between group variance}}{\text{within group variance}} = \frac{\frac{\sum_{k=1}^K n_k * (\mu_k - \bar{X})^2}{K-1}}{\frac{\sum_{k=1}^K \sum_{i=1}^{n_k} (X_{ki} - \mu_k)^2}{N-K}}$$





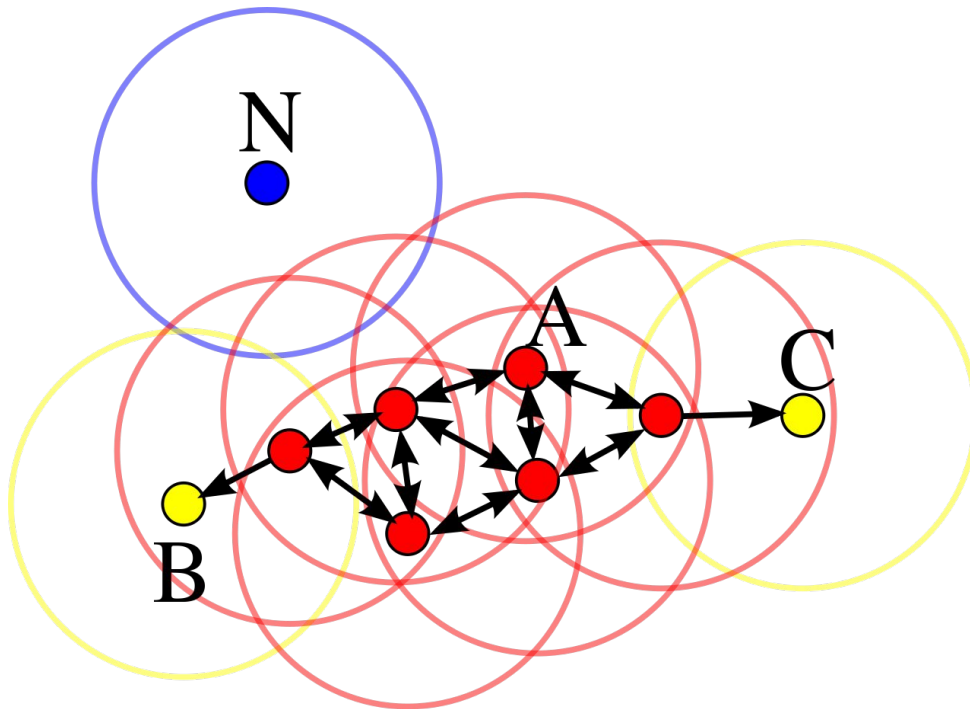
# DBSCAN

# DBSCAN

- Baziran na gustoći
- Prepoznaje šum
- Parametri
  - Radijus unutar kojeg se točke mogu nazvati susjedima
  - Minimalan broj susjeda koje točka mora imati da bi spadala u jezgru grupe
- Problemi
  - Skupovi podataka s velikim razlikama u gustoći grupa
  - Optimalan radijus

# Terminologija

- Točka jezgre
- Granična točka
- Šum



By Chire - Own work, CC BY-SA 3.0,  
<https://commons.wikimedia.org/w/index.php?curid=17045963>



# OPTICS

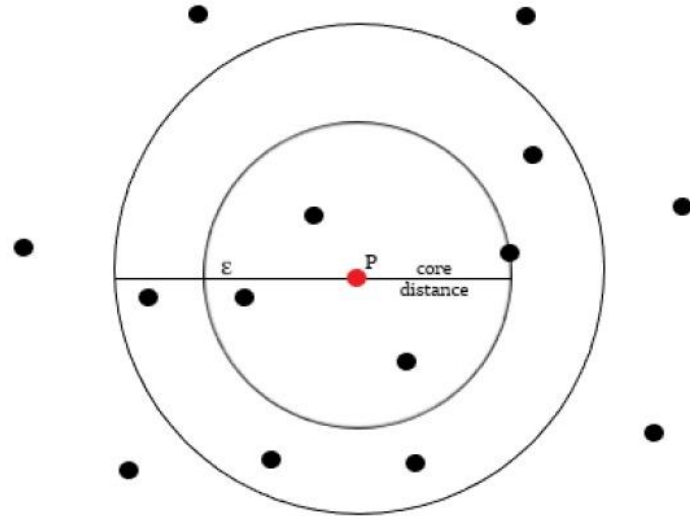


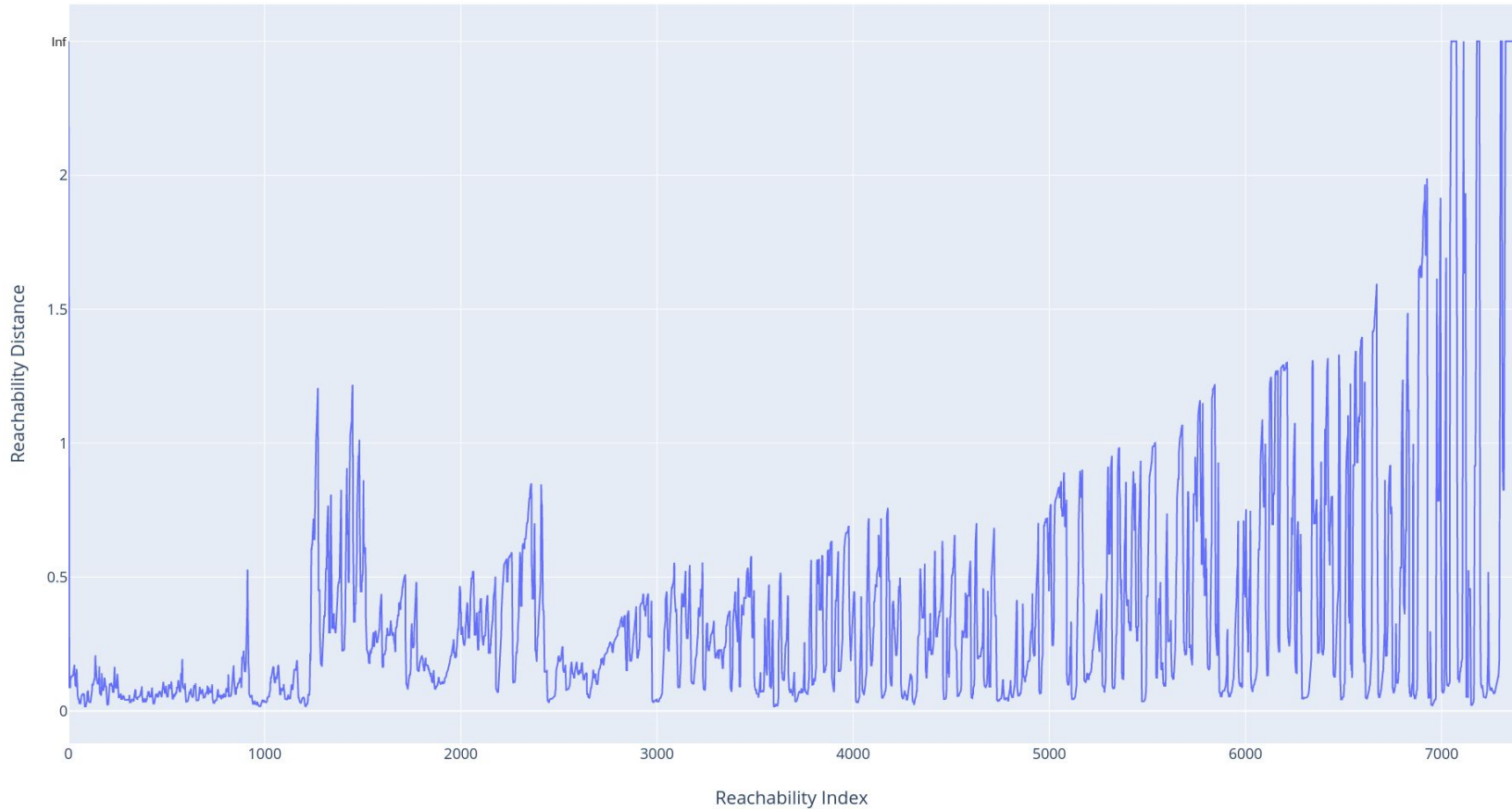
# OPTICS

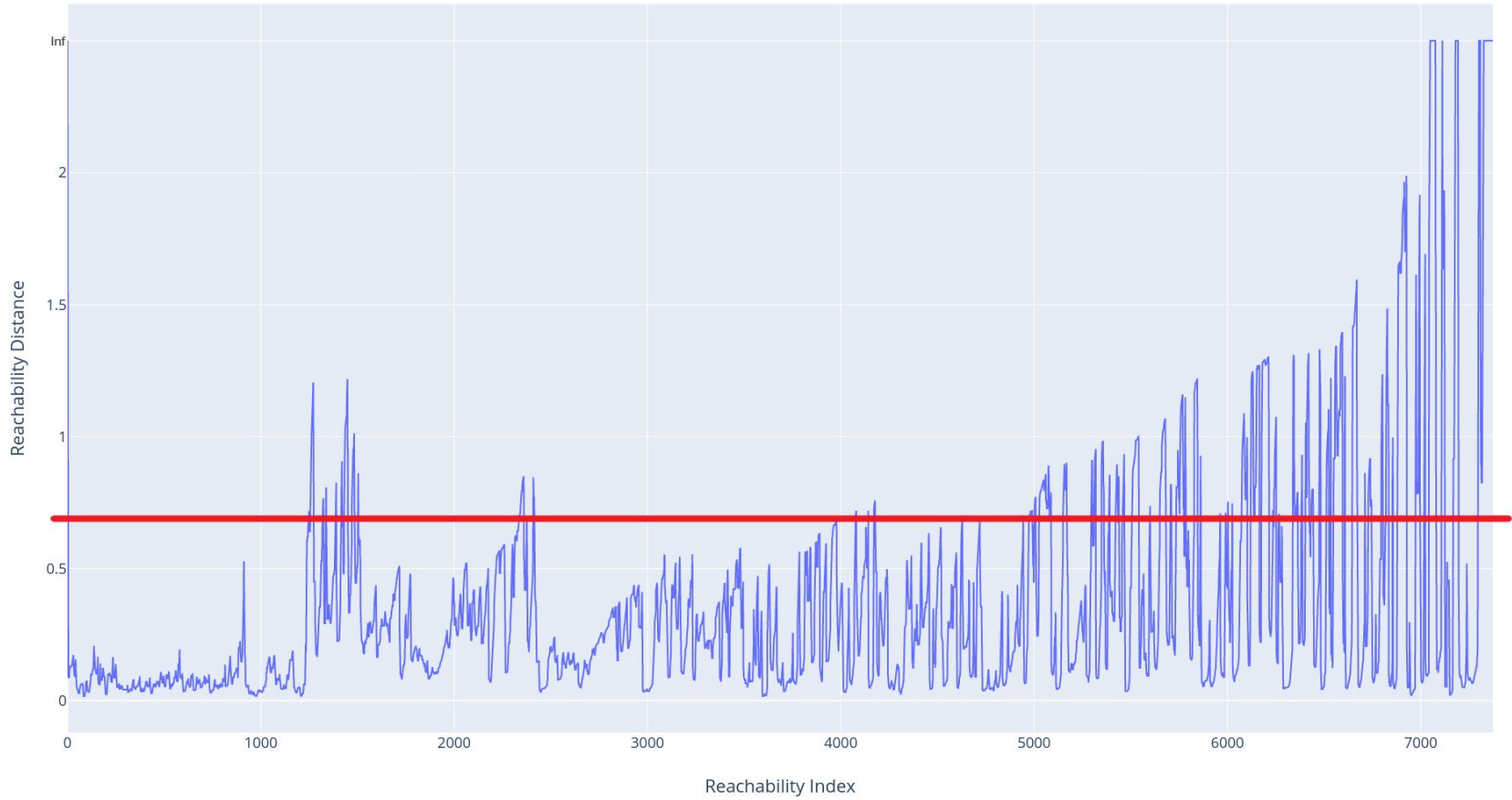
- Baziran na gustoći
- Ne grupira nego sortira
- Parametri
  - Maksimalan radijus unutar kojeg je točka “dohvatljiva”
  - Minimalan broj susjeda koje točka mora imati da bi spadala u jezgru grupe
- Problemi
  - Rezultat algoritma nisu grupe
  - Teško odrediti minimalan broj susjeda

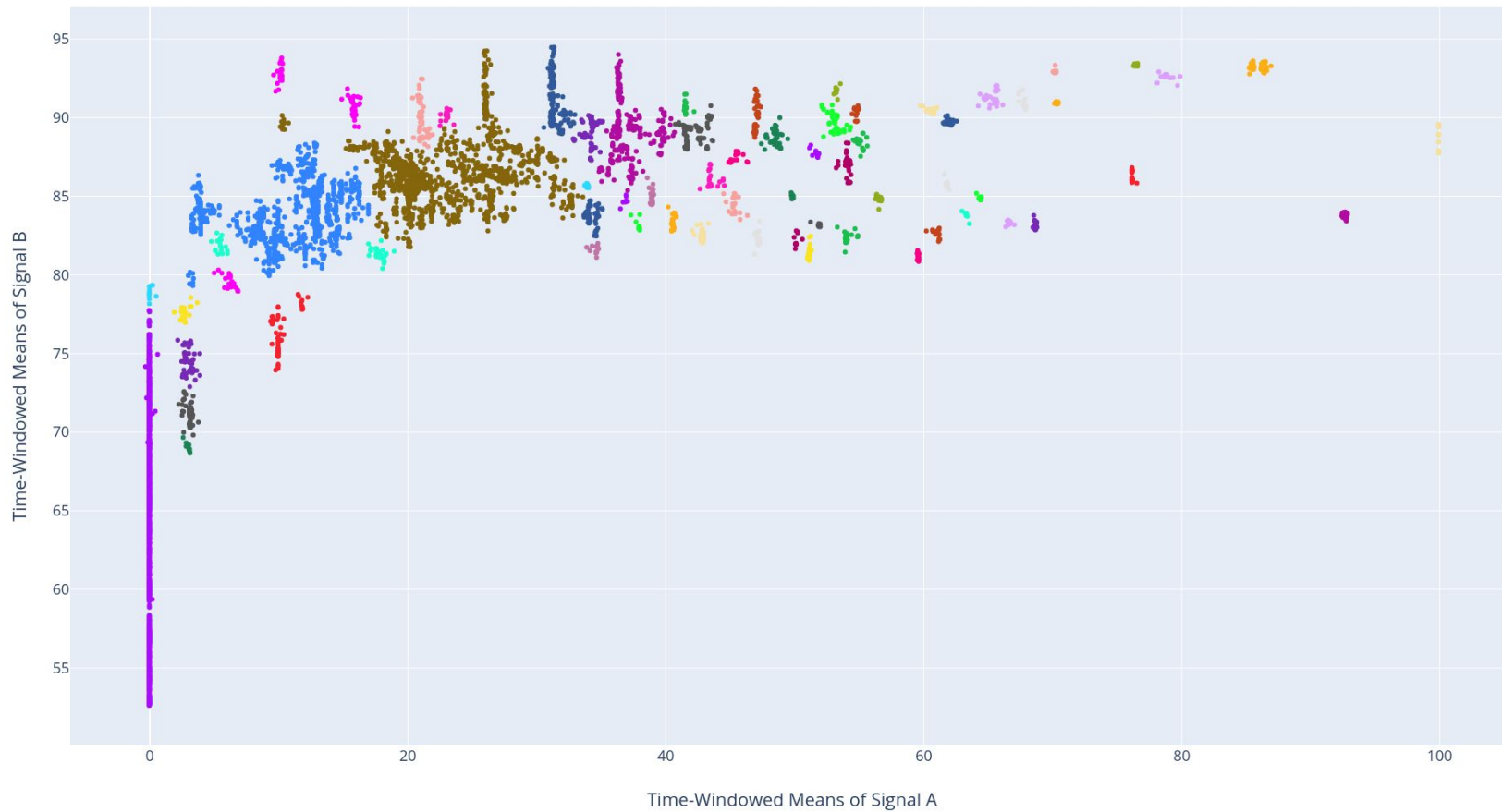
# Terminologija

- Udaljenost jezgre
- Udaljenost dohvatljivosti

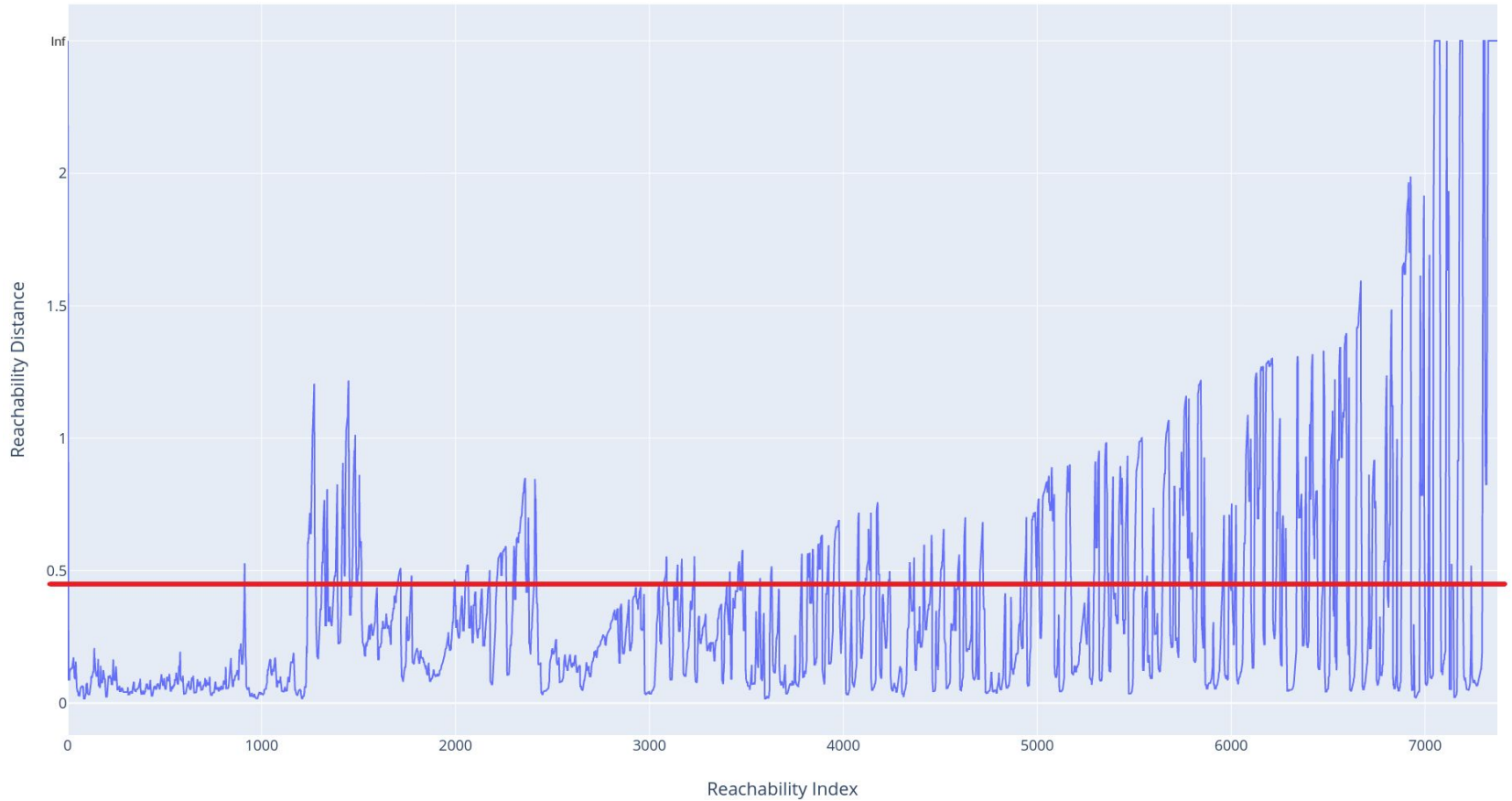








DBSCAN rezultat za radijus 0.65 i broj susjeda 10





DBSCAN rezultat za radijus 0.45 i broj susjeda 10

# Zaključak

- Spadaju li sve točke u uzorke koji se ponavljaju?
- Koliko je teško odrediti parametre algoritma unutar domene testiranja motora?
- Mogu li grupe biti različitih gustoća?