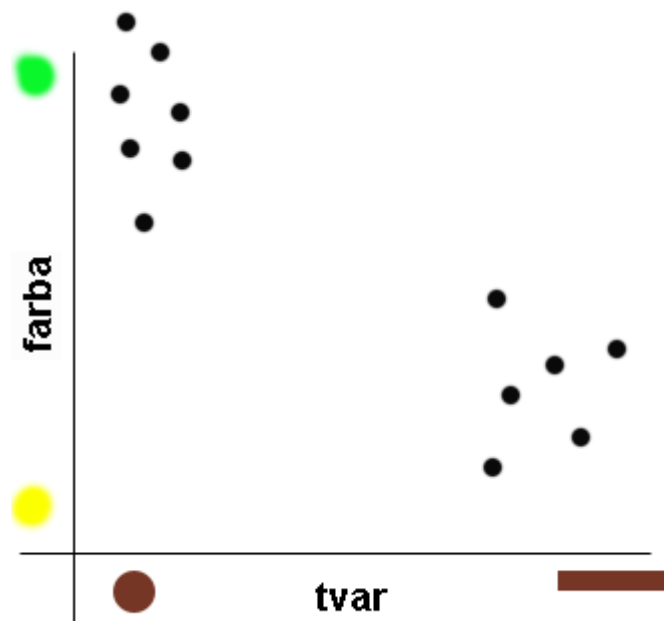


Samorganizácia

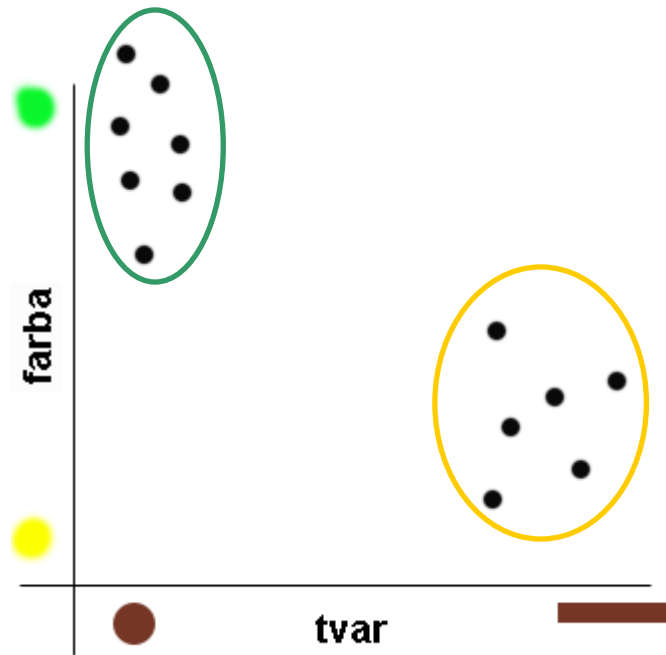
Učenie bez učiteľa

- Jablká a banány



Učenie bez učiteľa

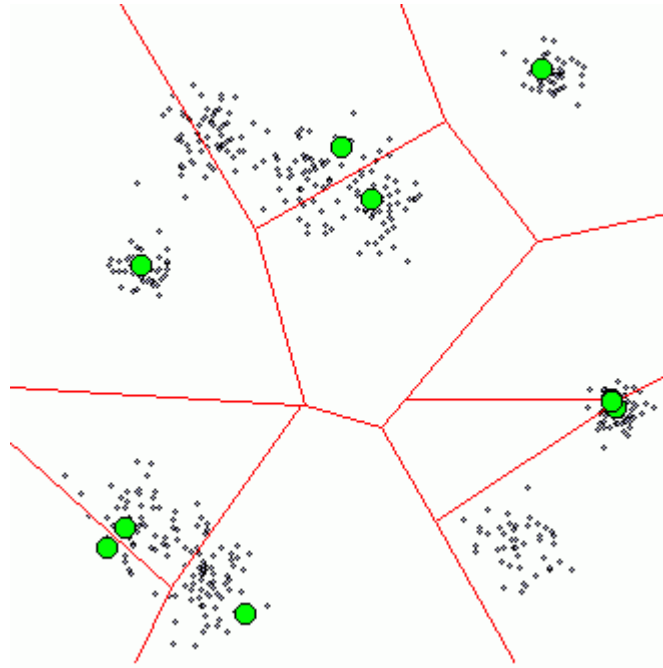
- Jablká a banány



Klastrová analýza

- K-means algoritmus
 - Zvoľ K centier – náhodne / spomedzi vstupných vektorov
 - Pre každý vektor zisti, ku ktorému z K centier je najbližšie (väčšinou Euklidovsky)
 - Vypočítaj nové centrá ako priemer z vektorov, ktoré prislúchajú danému centru
 - Opakuj určitý počet krát, alebo ak už len malé zmeny

K-means algoritmus

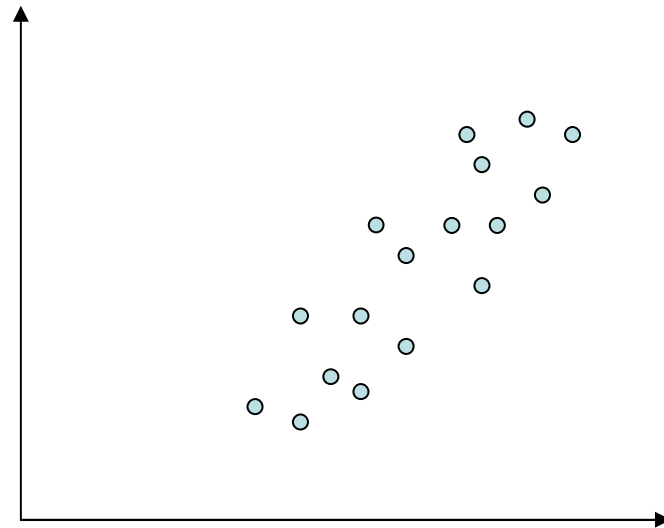
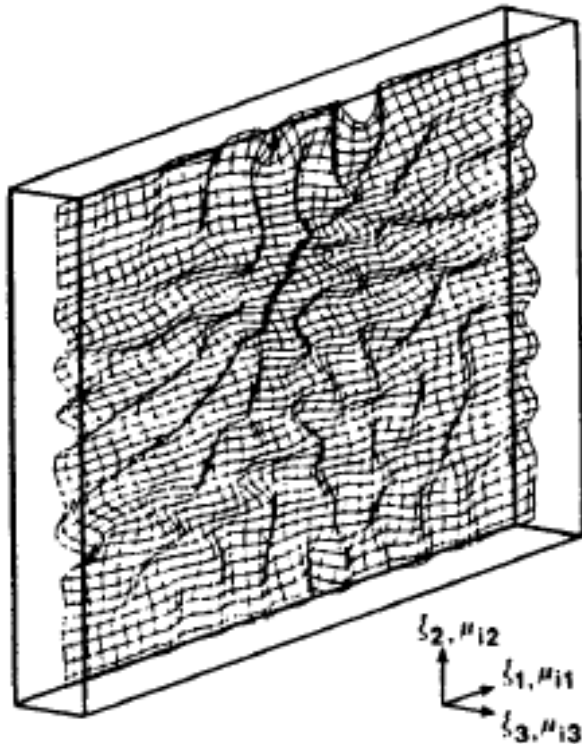


Redukcia dimenzie dát

- Dáta chceme vizualizovať
 - Maximálne 3 dimenzie, ideálne 2
- Potrebujeme skomprimovať dáta
- Odstrániť šum
- Odhaliť nadbytočné atribúty

Redukcia dimenzie dát

- Metóda hlavných komponentov
 - Principal Component Analysis (PCA)



Metóda hlavných komponentov

- Odčítať strednú hodnotu
- Vypočítať kovariančnú maticu $\frac{1}{n-1} \mathbf{XX}^T$
- Vypočítať vlastné vektory (jednotkové) a vlastné čísla kovariančnej matice
- Projekcia dát - vynásobiť každý dátový vektor maticou obsahujúcou vlastné vektory s najväčšími vlastnými číslami
- Bezparametrická štatistická metóda