

10. - 11. cvičení

Shlukování, učení bez učitele

- 1 algoritmus k středů a kompetiční učení - vizualizace průběhu algoritmů
- 2 srovnání algoritmu k středů s kompetitivním učením na různých datech
- 3 sofistikovanější porovnávání shlukovacích algoritmů, nalezení optimálního počtu shluků
- 4 aplikace shlukování: vektorová kvantizace
- 5 hierarchické shlukování
- 6 Kohonenova mapa (SOM)
- 7 Algoritmus LVQ

Shlukování

Ukázky v Matlabu:

- **clustering_kmeans.m** - vizualizace učení u algoritmu k středů
- **clustering_competitive_layer.m** -vizualizace učení u kompetitivní neuronové sítě
- **clustering_kmeans_evaluate.m** - ukázka, jak použít v Matlabu implementovaný algoritmus k středů a jak vyhodnotit kvalitu shlukování
- **clustering_bitmap.m** - ukázka využití shlukování pro vektorovou kvantizaci
- **clustering_hierarchy_countries.m** - hierarchické shlukování států
- **clustering_hierarchy.m** - hierarchické shlukování - další

Shlukování - kmeans a competlayer

- Nejprve se podíváme na shlukování metodou k středů a pomocí kompetitivního modelu pro různá data.
clustering_kmeans.m, clustering_competitive_layer.m

Otázky:

- Jaký má na učení vliv inicializace centroidů?
- Liší se výsledky pro různé inicializace?
- Jak dlouho trvá učení pro větší data?
- Je některá z metod (alg. k středů, competlayer) výrazně lepší nebo horší?
- Jak zhodnotit, zda jsou vytvořené shluky kvalitní?

Shlukování - kmeans a competlayer

- Pak se podíváme, jak je implementovaný algoritmus k středů v Matlabu, jak nastavit jeho parametry a jak zhodnotit výsledky **clustering_kmeans_evaluate.m**

Otázky:

- Zkuste si pohrát s nastavením parametrů metody kmeans
- Zkuste pro jednotlivá data najít optimální počet shluků

Shlukování a vektorová kvantizace

- Podívejte se, jak lze využít shlukování k vektorové kvantizaci.
clustering_bitmap.m
 - Zkuste různé počty centroidů a porovnejte výsledky
 - Jak dlouho trvá učení?
 - Úkol: vyzkoušejte vlastní obrázky - fotografie.

Hierarchické shlukování

- Ukázka hierarchického shlukování':
clustering_hierarchy_countries.m **clustering_hierarchy.m**
 - Nejprve se podíváme na to, jak si s daty poradí kmeans.
 - Pak si ukážeme hierarchické shlukování

Úkol

- Pohrajte si s nastavením parametrů: metrika, počet klastrů, prohlédněte si dendrogramy
- Srovnejte jednotlivé metriky hierarchického shlukování mezi sebou.
- Pokuste se nalézt optimální dvojice: metrika + počet shluků.

Kohonenovy mapy (SOMs)

- budeme pracovat s knihovnou SOM Toolbox
<http://www.cis.hut.fi/projects/somtoolbox/>

Ukázky v Matlabu:

- **SOM_pohadkove_postavy.m** - ukázka práce s knihovnou, vizualizace,
- **SOM_colors.m** - samoorganizace barev, vizualizace
- **SOM_bitmap.m** - samoorganizace barev v případě konkrétního obrázku
- **SOM_clustering.m** - ukázka využití SOM pro shlukování a pro vektorovou kvantizaci
- **SOM_clustering.m** - algoritmus LVQ