

**Kvíz BI-ZNS navazující na cv. 11-B (10/10)**  
Autor kvízu: Mgr. Ing. Ladislava Smítková Janků, Ph.D.

Jméno studenta: JAN ŠMERDA  
Studijní skupina: 7.

1. Formulace úlohy shlukové analýzy.
2. Vysvětlete rozdíl mezi hierarchickými a nehierarchickými shlukovacími metodami.
3. Jak se provádí standardizace dat před aplikací některého z algoritmů shlukové analýzy?
4. Popište metodu K-means. *viz 2. z odpovědí*
5. Popište heuristický prosívací algoritmus pro redukci množiny pravidel.
6. Popište algoritmus TREPAN.
7. Vyjmenujte kroky algoritmu CRED a stručně je vysvětlete.

Otázky 1-4 řešte s pomocí materiálů ke Cvičení 11 a k Přednášce 9 (odpřednášená v 10. týdnu ZS 2020/2021).

Otázky 5-7 řešte s pomocí materiálů k Přednášce 11 (bude odpřednášena ve 12. týdnu ZS 2020/2021).

- 1) Cíl je nalézt služby (cluster) bodů, které si jsou blízké.
- 2) - hierarchické metody více připomínají rozhodování stromy.  
 Začíná se tím, že najdeme všechny prvky služby a postupně se slučují podle vzdálenosti.
- nehierarchické metody (např. k-means) fungují iterativně.  
 Hierarchické metody se při poi. těžiště a body se přivádí k tomu těžišti, ke kterému jsou blíže.  
 Do středu těchto služeb se ~~načítá~~ <sup>přemísť</sup> těžiště a proces se opakuje.  
 Proces končí, pokud již žádná těžiště neexistovala přičtením bodů do služeb.

3) Rozhodnutí dat se přestěhují tak, aby všechny dimenze sdílely stejný rozsah.

5) Zároveň na postupném odstraňování pravidel, která nepřispívají k dobré klasifikaci.

Pro každý vzor  $x$  hledáme množinu  $T$  a sady pravidel  $R$  z množiny pravidel  $R$  se pohybuje na  $x$ , ~~že~~ do které by byl vzor  $x$  řazen na soubor pravidel  $R$  a redukované množiny pravidel  $R - \{r\}$ .

Společná výchozí - počet vzorů, které byly správně přičteny pomocí

$R$  - počet vzorů, které nebyly správně přičteny pomocí  $R - \{r\}$

Pokud je výchozí nulová nebo záporná, pravidlo odstraníme

6) Vlastní testovací a trénovací data v tomto případě  
slouží k ověření již naučené neuronové sítě.  
Jinak se rozhodovací strom vyvíjí podle stejných pravidel.

- 7)
- 1) Převod spojitého výstupní proměnné na diskrétní hodnoty  
pomocí nejvhodnějších prahových konstant
  - 2) vytvoření stromu, který mapuje diskretizovaný výstup  
dvouvrstvý k 1. kroku na výstupní výstupy.  
Z tohoto stromu extrahují pravidla (tzv. pravidla).  
Tito rozhodují na dříve funkce.
  - 3) Analogicky postupují v případě výstupní výstupy,  
čímž získávají pravidla.
  - 4) Substitují pravidla do meziklasických pravidel.  
Závěry výstupních pravidel jsou předpoklady meziklasických pravidel.  
Tím získávají celková pravidla.
  - 5) Pravidla zjednoduší pomocí vhodného slovníku.