

Analýza a klasifikace dat – úvodní informace



RNDr. Eva Koritáková, Ph.D.

Přínos předmětu

- orientace v principech rozpoznávání a klasifikace dat s důrazem na zpracování medicínských a biologických dat
- schopnost zvolit a aplikovat adekvátní metodu analýzy a klasifikace dat k dosažení požadovaných výsledků
- schopnost správné interpretace dosažených výsledků včetně vyhodnocení úspěšnosti klasifikace

Forma výuky předmětu

- tzv. „flipped classroom“ – teoretické učivo (prezentace, nahrané přednášky a výukové texty) si studenti prochází sami a při kontaktní výuce probíhá opakování formou Kahootu a cvičení na tabuli či v softwaru Matlab

Datum	Téma	Domácí příprava	Délka videa	Výuka	Vyučující
20.09.2023	Organizační informace + CCA	Ano		Distanční	EK
27.09.2023	Úvod	Ano	1:36:50	-	
04.10.2023	Bayesův klasifikátor	Ano	1:22:10	Prezenční	TJ
11.10.2023	Minimální vzdálenost	Ano	2:13:09	Prezenční	TJ
18.10.2023	FLDA	Ano	28:24 + 2:43 + 3:22	Prezenční	TJ
25.10.2023	SVM a sekvenční klasifikace	Ano	49:22 + 28:50	-	
01.11.2023	Úspěšnost a ROC	Ano	1:40:09 + 2:20	Prezenční	TJ
15.11.2023	Redukce, PCA a variální učení	Ano	2:00:40	Prezenční	TJ
22.11.2023	ICA a selekce	Ano	1:23:51	Prezenční	TJ
???	Neuronové sítě	-		Prezenční	RV

Požadavky ke zkoušce

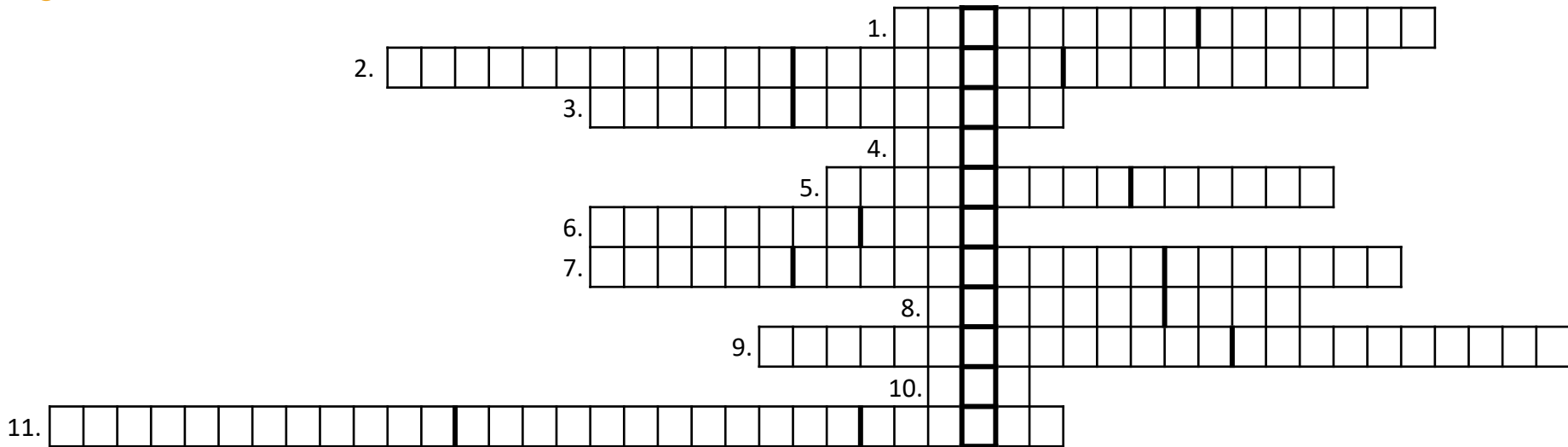
- předmět je ukončen písemnou zkouškou formou testu v ISu
- je nutné porozumět probíraným tématům, být schopen zvolit správnou metodu na daný analytický problém a umět získané výsledky interpretovat
- maximálně 2 absence na cvičeních

Doporučená literatura

- Janoušová, E. et al.: online výukové materiály Vícerozměrné metody pro analýzu a klasifikaci dat
<http://portal.matematickabiologie.cz/index.php?pg=analiza-a-hodnoceni-biologickyh-dat--vicerozmerne-metody-pro-analyzu-dat>
- Holčík, J.: Analýza a klasifikace dat. Brno, CERM 2012, 112s.
<http://www.iba.muni.cz/res/file/ucebnice/holcik-analyza-klasifikace-dat.pdf>
<http://www.iba.muni.cz/index.php?pg=vyuka--ucebnice>
- DUDA R. O., HART P. E., STORK D. G., 2000: Pattern Classification. Wiley-Interscience, New York, 680 pp.
- BISHOP C., 2006: Pattern Recognition and Machine Learning. Springer, New York, 738 pp.
- KUNCHEVA L. I., 2004: Combining Pattern Classifiers: Methods and Algorithms. Wiley-Interscience, New Jersey, 376 pp.

Vícerozměrná analýza dat - opakování

Křížovka - zadání



1. Metoda, jež umožňuje spočítat vzdálenost vzdušnou čarou mezi dvěma body.
2. Předpoklad mnoha metod vícerozměrné analýzy dat.
3. Vztah dvou spojených veličin, když s rostoucími hodnotami jedné proměnné narůstají hodnoty druhé proměnné.
4. Metoda, která umožňuje zobrazit pozorované vzdálenosti/nepodobnosti mezi objekty v euklidovském prostoru (zkratka).
5. Tabulkové znázornění vztahu proměnných či objektů.
6. Grafické znázornění vztahu většího počtu spojených proměnných.
7. Metoda vedoucí zpravidla k protáhlým shlukům v datech.
8. Grafické znázornění podobnosti objektů převedením hodnot proměnných na tvary či symboly.
9. Typ analýzy, která umožňuje vytvořit shluky stejného řádu.
10. Metoda, která umožňuje transformaci dat do nových nekorelovaných proměnných (zkratka).
11. Metoda, jež umožňuje zjistit, zda se od sebe liší dvě skupiny objektů, které jsou popsány hodnotami několika proměnných.