

D – Charakteristika studijního předmětu				
Název studijního předmětu	Vytěžování informací			
Typ předmětu	Povinně volitelný		doporučený ročník / semestr	3/ZS
Rozsah studijního předmětu	2+2	hod. za týden	4	kreditů 5
Jiný způsob vyjádření rozsahu				
Způsob zakončení	Zkouška		Forma výuky	Přednášky, cvičení
Další požadavky na studenta				
Vyučující				
Ing. Jindřich Žďánský, Ph.D.				
Stručná anotace předmětu				
Stručná anotace předmětu Předmět je zaměřen na úvod do vytěžování dat, tj. data mining / machine learning. Důraz bude kladen na motivaci a aplikaci vytěžování, dále pak na vlastnosti a popis různých typů dat, kvalitu dat, principy jejich předzpracování, (shlukování, vzorkování, redukce rozměru, výběr příznaků, transformace), měření podobnosti datových objektů, analýzu dat, výpočet základních statistik, jejich vizualizaci. Závěrečná část bude věnována základům klasifikace dat a úvod do asociativní a shlukové analýzy. Výklad bude doplňován praktickými příklady.				
Témata přednášek Úvod do vytěžování dat: motivace a aplikace vytěžování. Vlastnosti a popis různých typů dat a datových sad, kvalita dat. Předzpracování dat, shlukování, vzorkování, redukce rozměru, výběr příznaků, transformace dat. Měření podobnosti příp. nepodobnosti dat/ datových objektů, sad. Analýza dat, základní statistiky a jejich vizualizace, analýza vícerozměrných dat. Základy klasifikace dat, aplikace pravidel, metoda nejbližšího souseda, bayesovský klasifikátor, neuronové sítě. Úvod do asociativní a shlukové analýzy.				
Náplň cvičení Lineární algebra a MATLAB. Dávková a průběžná analýza datových souborů. Vizualizace dat a jejich statistik. Základní metody klasifikace Bayes, nejbližší soused, pravidla, neuronové sítě. Klastrování dat. Příklady data miningu Google relevance scoring, kontextové reklamy, prohledávání multimediálních dat, ISO MPEG-7 ...				
Informace ke kombinované nebo distanční formě				
Rozsah konzultací (soustředění)			hodin za týden	
Rozsah a obsahové zaměření individuálních prací studentů a způsob kontroly				
Studijní literatura a studijní pomůcky				
Pang-Ning Tan, Michael Steinbach, Vipin Kumar: Introduction to Data Mining, Addison Wesley; US ed edition (May 12, 2005), ISBN: 978-0321321367				
Daniel T. Larose: Discovering Knowledge in Data: An Introduction to Data Mining. Wiley-Interscience; 1 edition (November 18, 2004), ISBN: 978-0471666578				