

C – Charakteristika studijního předmětu nebo tematického bloku	
Název studijního předmětu	Získávání znalostí z databází
Způsob zakončení	Zkouška
Další požadavky na studenta	
Základní znalosti databází a programování v rozsahu standardních magisterských kurzů oborů zaměřených na informatiku	
Přednášející	
Doc. Ing. Ladislav Beránek, MBA, CSc.	
Stručná anotace předmětu	
<p>Cílem předmětu je seznámit studenty s metodami automatizovaného získávání znalostí z databází. V předmětu bude podán přehled problematiky (která je klíčovým faktorem rozvoje informačních technologií), používaných metod a systémů vyvíjených ve světě a u nás. Budou probrány metodiky (5A, CRISP a další), databáze jako zdroje dat (OLAP, datové sklady) a dále metody jako jsou rozhodovací pravidla, neuronové sítě ale i metody zpracování dat z distribuovaných zdrojů a vyhledávání na webu. Základními tématy jsou zejména: reprezentace znalostí, příprava dat, čištění dat a jejich vizualizace, klasifikační metody (OneR, naiveBayes), rozhodovací stromy, klasifikace (metoda nejbližšího souseda), asociační pravidla, metody shlukové analýzy (K-means) a další. Vyhodnocení metod – odhad chyby, ROC, princip MDL, křivky cen.</p> <p>Součástí předmětu bude seminární projekt, kde studenti zpracují zadanou úlohu v některém z nástrojů používaných pro řešení úloh data miningu, např. se systémem Weka.</p>	
Odborná literatura	
<p>Základní:</p> <p>Berka, P. Dobývání znalostí z databází. Praha : Academia, 2003.</p> <p>Witten, H., I., Eibe, F. Data Mining : Practical Machine Learning Tools and Techniques. 2nd edition. San Francisco : Morgan Kaufman, 2005.</p> <p>Markov, Z., Larose, T., D. Data Mining the Web : Uncovering Patterns in Web Content, Structure, and Usage. New Jersey : John Wiley & Sons, 2007.</p> <p>Fayyad, U. Advances in Knowledge Discovery and Data Mining, Boston : MIT Press, 1996.</p>	