

INFORMAČNÝ LIST PREDMETU

Vysoká škola: Univerzita Konštantína Filozofa v Nitre	
Fakulta: Fakulta prírodných vied	
Kód predmetu: KI/ WM/15	Názov predmetu: Web Mining
Druh, rozsah a metóda vzdelávacích činností: Forma výučby: Prednáška / Cvičenie Odporúčaný rozsah výučby (v hodinách): Týždenný: 2 / 2 Za obdobie štúdia: 26 / 26 Metóda štúdia: prezenčná	
Počet kreditov: 6	
Odporúčaný semester/trimester štúdia: 2.	
Stupeň štúdia: II.	
Podmieňujúce predmety:	
Podmienky na absolvovanie predmetu: Úspešné absolvovanie predmetu je podmienené skúškou, samostatným riešením projektu a úloh zadaných v kurze. Ústna skúška pozostáva z teoretických otázok (30%), z otázok týkajúcich sa riešených úloh v kurze (30%) a z obhajoby projektu (40%). Hodnotenie predmetu je dané výsledkom skúšky. Hodnotenie: A - výborne: 100% - 93% B - veľmi dobre: 92,99% - 85% C - dobre: 84,44% - 77% D - uspokojivo: 76,99% - 69% E - dostatočne: 68,99% - 60% FX - nedostatočne: 59,99% - 0%	
Výsledky vzdelávania: Študent rozumie procesu objavovania znalostí. Študent rozlišuje domény web mining-u, je schopný získať a predspracovať relevantné dáta o obsahu, štruktúre a používaní webu. Dôraz je kladený na predspracovanie dát (Data Preparation, DP), ktoré je časovo najnáročnejšou fázou celého procesu objavovania znalostí (Knowledge Discovery, KD). Študent diskutuje o možnostiach predspracovania dát. Pri riešení projektu sa kladie dôraz na oblasť objavovania znalostí na základe používania webu- web usage mining. Práve na tejto oblasti si študent ozrejní princíp a jednotlivé fázy procesu objavovania znalostí. Študent po absolvovaní predmetu je schopný riadiť proces objavovania znalostí.	
Stručná osnova predmetu: 1. Proces objavovania znalostí; metodiky (Six Sigma, SEMMA, CRISP-DM). 2. Domény objavovania znalostí z webu (web content mining, web structure mining, web usage mining); zdroje dát (kolekcia dokumentov/web stránok, mapa webu, web crawling, common log file, extended log file, cookies). 3. Riadenie procesu web usage mining-u; definícia cieľovej úlohy, určenie typu problému; získanie relevantných dát o používaní webu, prieskum dát; predspracovanie dát (čistenie dát, identifikácia používateľov/sedení, rekonštrukcia aktivít používateľov webu/dopĺňanie ciest); dolovanie z dát/	

aplikácia analytických metód; interpretácia a evalvácia nájdených znalostí; aplikácia získaných znalostí.

4. Projekt: metodika evalvácie prípravy dát v procese objavovania vzorov správania sa používateľov webu (predspracovanie dát na rôznych úrovniach prípravy dát, hľadanie vzorov správania sa používateľov webu v jednotlivých súboroch, porozumenie výstupným dátam, porovnanie získaných znalostí z jednotlivých súborov).

Odporúčaná literatúra:

Berka, P. 2003. Dobývání znalostí z databází. Praha : Academia, 2003. ISBN 80-200-1062-9.

Liu, B. 2007. Web data mining: Exploring hyperlinks, contents and usage data. Springer, 2007. ISBN 978-3-540-37881-5.

Munk M. - Drlik M. 2011. Impact of different pre-processing tasks on effective identification of users' behavioral patterns in web-based educational system. In Procedia Computer Science. Elsevier, ISSN 1877-0509, 2011, vol. 4, p. 1640-1649.

Munk, M. - Drlík, M. - Kapusta, J. - Munková, D. 2013. Methodology Design for Data Preparation in the Process of Discovering Patterns of Web Users Behaviour. In Applied Mathematics & Information Sciences. ISSN 1935-0090, 2013, vol. 7, no. 1L, p. 27-36.

Munk, M. - Drlík, M. 2011. Influence of Different Session Timeouts Thresholds on Results of Sequence Rule Analysis in Educational Data Mining. In Communications in Computer and Information Science. Springer, ISSN 1865-0929, 2011, vol. 166, p. 60-74.

Munk, M. - Kapusta, J. - Švec, P. - Turčáni, M. 2010. Data Advance Preparation Factors Affecting Results of Sequence Rule Analysis in Web Log Mining. In E & M EKONOMIE A MANAGEMENT: Informační management. ISSN 1212-3609, 2010, vol. 13, no. 4, p. 143-160.

Munk, M. - Kapusta, J. - Švec, P. 2010. Data Preprocessing Evaluation for Web Log Mining: Reconstruction of Activities of a Web Visitor. In Procedia Computer Science. Elsevier, ISSN 1877-0509, 2010, vol. 1, no. 1, p. 2273-2280.

Munk, M. 2011. Počítačová analýza dát. Nitra : UKF, 2011. 361 s. ISBN 978-80-8094-895-5.

Jazyk, ktorého znalosť je potrebná na absolvovanie predmetu:

slovenský

Poznámky:

Hodnotenie predmetov

Celkový počet hodnotených študentov: 339

A	B	C	D	E	FX
61.95	8.26	12.98	6.49	9.14	1.18

Vyučujúci: prof. RNDr. Michal Munk, PhD.

Dátum poslednej zmeny: 17.02.2020

Schválil: prof. RNDr. Michal Munk, PhD.