

Štátne skúšky – okruhy tém k štátnicovému predmetu

Akademický rok:	2019/2020
Garantujúce pracovisko:	Katedra kybernetiky a umelej inteligencie FEI TU - KKUI
Študijný program:	Hospodárska informatika
Študijný odbor:	Hospodárska informatika
Stupeň štúdia:	2. – inžiniersky
Štátnicový predmet:	Hlavné poznatky odboru hospodárska informatika a ich využitie

Okruhy otázok z predmetu:

1. Vzťahy a rozdiely medzi dátami, informáciami a znalosťami, význam znalostí pre organizácie. Rôzne uhly pohľadu na manažment znalostí. Podmieneny pohľad na MZ: vplyv charakteristík riešených úloh, typu používaných znalostí, charakteristík organizácie a vonkajšieho prostredia na MZ.
2. Vyhľadávanie informácií, základné modely, práca s textovými dokumentmi, spôsoby vyhodnotenia kvality vyhľadávania. Vyhľadávanie na webe, využitie štruktúry liniek, algoritmy PageRank a HITS.
3. Proces objavovania znalostí z databáz, popis jeho základných krokov. Základné typy znalostí, ktoré je možné získať z dát. Metodológia CRISP-DM, popis jej jednotlivých krokov. Popisné dolovanie v dátach, ciele, príklady metód a používaných algoritmov (zhlukovanie, asociačné pravidlá, detekcia anomálií).
4. Prediktívne dolovanie v dátach, ciele strojového učenia, prediktívna úloha klasifikačného a regresného typu, regresná analýza, kontrolované učenie v úlohách klasifikačného typu (prehľadávanie priestoru pojmov, produkčné pravidlá, rozhodovacie stromy a zoznamy),
5. Kontrolované učenie s prvkami kvantitatívneho usudzovania (naivný Bayesov klasifikátor, etalóny, tabuľka kritérií, lineárna a sférická prahová jednotka a SVM metóda podporných vektorov). Nekontrolované učenie (zhlukovacie techniky), lenivé učenie, učenie súborom metód (bagging, boosting a náhodné stromy), aplikácie v praxi.
6. Základné spôsoby prehľadávania. Lokálne prehľadávanie a jeho komponenty. Iteračné zlepšovanie a únik z lokálneho extrému. Prírodne inšpirované algoritmy a algoritmy pre odhad distribúcie. Výpočtová zložitosť, základné zložitosťné triedy.
7. Stacionárne náhodné postupnosti, lineárny regresný model prvého rádu, analytické riešenie, lineárny trend, minimum kvadrátu odchýlok. Lineárne regresné modely vyššieho rádu, redukcia rozmernosti sústavy rovníc pomocou aproximácie, Almonovej model. Redukcia rozmernosti sústavy rovníc pomocou prvých diferencií prvkov náhodnej postupnosti, Koyckov model. Modely s fixovaným prvkom, Davidsonov model. Princíp a využitie Choleskyho dekompozície.
8. Systémy a manažérske informačné systémy (MIS), charakteristika systémov MIS, metódy a prostriedky realizácie manažérskej úrovne riadenia. Trojvrstvová architektúra systémov MIS (relačná vrstva, multidimenzionálna vrstva, prezentačná vrstva - prostriedky OLAP).

9. Relačná vrstva, systémy OLTP ako produkčné systémy a zdroje dát pre MIS, relačný databázový model, charakteristika relačných databáz, podmienky minimálnej relačnosti, pravidlá pre relačné databázové systémy, databázové tabuľky-entity, relačné vzťahy medzi entitami, primárny kľúč a cudzí kľúč, relácie, tri normálne formy.
10. Multidimenzionálna vrstva a multidimenzionálny dátový model, definícia a charakteristika dátových skladov, prvky dátových skladov a ich charakteristika (dimenzie, fakty, metadáta, repozitáre). Metódy, postupy a prostriedky budovania a napĺňania dátových skladov. Prezentačná vrstva, OLAP analýza, definícia a pravidlá OLAP, operácie a spôsoby prezentácie výstupov OLAP analýzy.
11. Definícia veľkých dát, charakteristika veľkých dát – 3V/5V model, princípy distribuovaných súborových systémov a databáz, typy NoSQL databáz a ich charakteristika – stĺpcovo orientované databázy, úložiska typu kľúč : hodnota, dokumentové databázy, grafové databázy.
12. Architektúry systémov pre spracovanie veľkých dát. Distribuované prostredie pre dávkové spracovanie dát – základný model MapReduce. Distribuované prostredie pre spracovanie prúdových dát – dátové toky, fronty správ. Kombinovaná architektúra. Prehľad technológií pre spracovanie veľkých dát.

Košice, 02.03.2020

prof. Ing. Peter Sinčák, CSc.
vedúci KKUI