

Course name: Big data & data mining		
Field of study: Computer science	Type of study: Full-time	Source code: CIDM2_03 (61E25)
Course characteristics: Mandatory within the additional content	Level: Second (M.Sc.)	Year: I Semester: II
Type of classes: lectures, laboratories, project	Hours per week: 2 lect, 2 lab, 1 proj	ECTS points amount: 5 ECTS

COURSE GUIDE

I. GENERAL INFORMATION OF THE COURSE

AIMS OF THE COURSE

- A1. Introducing the students to the basics of data analysis and exploration.
- A2. Application of analytical databases, OLAP cubes and big data programming model.
- A3. Obtaining by the students the practical skills in state-of-the-art solutions for data mining and big data.

PRELIMINARY REQUIREMENTS FOR THE KNOWLEDGE, SKILLS AND OTHER COMPETENCIES

1. Knowledge of the design and administration of databases.
2. Personal computer operating systems literacy.
3. Ability to program in any high level language.
4. Basic knowledge of programming in a database environment.
5. Ability to use different sources of information and technical documentation.
6. Ability to work independently and in a group.
7. Ability to correctly interpret and present their own activities.

II. LEARNING OUTCOMES

- EU 1– has basic theoretical knowledge in the field of knowledge discovery methods in data structures,
- EU 2– has a basic knowledge of data warehouse,
- EU 3– has basic knowledge about the design of Business Intelligence,
- EU 4– knows the technology and tools for the tasks related to the mining of the knowledge contained in analytical databases,
- EU 5– able to select appropriate algorithms for data analysis depending the problem and to implement them,
- EU 6– can design appropriate diagrams to organize information using known tools,
- EU 7– is able to propose a solution to a specific issue related to data mining and big data.

LEARNING CONTENT

Lectures	Hours
----------	-------

Lect. 1	Introduction to the analysis and data mining.	2
Lect. 2	Data Warehousing - architecture.	2
Lect. 3	OLAP Technology - OLAP cubes.	2
Lect. 4	Introduction to MDX and MDX expressions	2
Lect. 5	Server SSAS - the basics of working in the environment and automate administrative tasks, part 1.	2
Lect. 6	Server SSAS - the basics of working in the environment and automate administrative tasks, part 2.	2
Lect. 7	Introduction to the basic techniques of data mining.	2
Lect. 8	Application of data mining techniques - classification.	2
Lect. 9	Application of data mining techniques - regression.	2
Lect. 10	Application of data mining techniques - segmentation.	2
Lect. 11	Application of data mining techniques – association	2
Lect. 12	Application of data mining techniques - sequential analysis and forecasting	2
Lect. 13	Big data programming model – part 1	2
Lect. 14	Big data programming model – part 2	2
Lect. 15	Reading and evaluation of outcomes - visualization and reporting	2
Laboratories		Hours
Lab. 1	Introduction to the SQL Server environment and tools used in the data analysis process.	2
Lab. 2	Project of simple analytical databases and analytical cubes.	2
Lab. 3	Installing and getting to know the structure of the sample data warehouse.	2
Lab. 4	Basic MDX expressions used during data processing.	2
Lab. 5	MDX - use of additional built-in functions of language.	2
Lab. 6	Server SSAS - monitoring the work and safety - roles, permissions, etc. SQL Server Profiler.	2
Lab. 7	Tasks SQL Server Agent, XMLA.	2
Lab. 8	Analysis of the data using Excel.	2
Lab. 9	Practical use of the classification.	2
Lab. 10	Practical use of regression.	2
Lab. 11	Practical use of segmentation.	2
Lab. 12	Practical use of association.	2
Lab. 13	Practical use of sequential analysis.	2
Lab. 14	Practical use of forecasting.	2
Lab. 15	Methods of presentation and evaluation of results.	2
Project		Hours
Proj. 1	SQL databases 1	2
Proj. 2	SQL databases 2	2
Proj. 3	Mining multimedia data 1	2
Proj. 4	Mining multimedia data 2	2
Proj. 5	Mining multimedia data 3	2
Proj. 6	Big data programming model 1	2
Proj. 7	Big data programming model 2	2
Proj. 8	Big data programming model 3	1

DIDACTIC TOOLS

1.	– lectures using multimedia presentations
2.	– blackboard and chalk or whiteboards and pens
3.	– laboratory guides
4.	– reports from laboratory activities

5. – computer stations with software

EVALUATION OF SUBJECT EDUCATIONAL EFFECTS ACHIEVEMENT SPOSOBY OCENY (F – FORMING, P – SUMMARIZING)

F1. – assessment of preparation for laboratory exercises

F2. – assessment of the ability to apply acquired knowledge while performing exercises

F3. – assessment of reports on the implementation of exercises

F4. – assessment of activity during classes

P1. – assessment of the ability to solve the problems posed and the method of presentation of the results obtained - credit for grade *

P2. – assessment of mastery of the teaching material of the lecture - passing the lecture (or exam)

*) the condition for obtaining credit is to receive positive grades from all laboratory exercises.

STUDENT WORKLOAD

Forma aktywności	Średnia liczba godzin na zrealizowanie aktywności
Hours with the teacher	30Lect 30Lab 15P → 75 h
Consultation hours	1 h
Exam	74 h
Student's independent workload	
Total	Σ 125 h
Total number of ECTS POINTS for the COURSE	6 ECTS
Number of ECTS points obtained by a student in classes that require direct participation of the teacher	3,04 ECTS

BASIC AND ADDITIONAL LITERATURE

Mark Hall, Ian Witten, Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann 2011.

D. J. Hand, Heikki Mannila, Padhraic Smyth, Principles of Data Mining, MIT Press, 2001

Jamie MacLennan, ZhaoHui Tang, Bogdan Crivat, Data Mining with Microsoft SQL Server 2008, John Wiley & Sons

PROWADZĄCY PRZEDMIOT (IMIĘ, NAZWISKO, ADRES E-MAIL)

1. dr hab. inż. Rafał Scherer, prof. PCz, rafal.scherer@iisi.pcz.pl

MACIERZ REALIZACJI EFEKTÓW UCZENIA SIĘ

Efekt uczenia się	Odniesienie danego efektu do efektów zdefiniowanych dla całego programu (PEK)	Charakterystyki I stopnia PRK	Charakterystyki II stopnia PRK		Cele przedmiotu	Treści programowe	Narzędzia dydaktyczne	Sposób oceny
		uniwersalne	O charakterze ogólnym	Dla kwalifikacji obejmujących kompetencje inżynierskie				
EU1	KAB2_W01 KAB2_W02 KAB2_U10	P7U_W P7U_U	P7S_WG P7S_UW	P7S_WG P7S_UW	A1	W1-W12 L1-12 P1-5	1,2,3,4,5	F1-F4 P1,P2
EU2	KAB2_W01 KAB2_W02 KAB2_U10	P7U_W P7U_U	P7S_WG P7S_UW	P7S_WG P7S_UW	A1, A2	W1-W12 L1-12 P1-5	1,2,3,4,5	F1-F4 P1,P2
EU3	KAB2_W01 KAB2_W02 KAB2_U11 KST_U22	P7U_W P7U_U	P7S_WG P7S_UW	P7S_WG P7S_UW	A1, A2	W1-W12 L1-12 P1-5	1,2,3,4,5	F1-F4 P1,P2
EU4	KAB2_W01 KAB2_W02 KIF_W08 KST_W21 KAB2_U11 KIF_U17	P7U_W P7U_U	P7S_WG P7S_UW	P7S_WG P7S_UW	A2, A3	W1-W12 L1-12 P1-5	1,2,3,4,5	F1-F4 P1,P2
EU5	KAB2_W01 KAB2_W02 KIF_W08 KST_W21 KIF_U22	P7U_W P7U_U	P7S_WG P7S_UW	P7S_WG P7S_UW	A2, A3	W1-W12 L1-12 P1-5	1,2,3,4,5	F1-F4 P1,P2
EU6	KIF_W08 KST_W21 KIF_U19 KIF_U22 KIO2_U10 KIO2_U18 KST_U22	P7U_W P7U_U	P7S_WG P7S_UW	P7S_WG P7S_UW	A2, A3	W1-W12 L1-12 P1-5	1,2,3,4,5	F1-F4 P1,P2
EU7	KAB2_U02 KAB2_U03	P7U_U	P7S_UO P7S_UW	P7S_UW	A1, A2, A3	W1-W12 L1-12 P1-5	1,2,3,4,5	F1-F4 P1,P2

III. FORMY OCENY - SZCZEGÓŁY

	Na ocenę 2	Na ocenę 3	Na ocenę 4	Na ocenę 5
Efekt 1	Student nie opanował podstawowej wiedzy z zakresu teorii eksploracji danych	Student częściowo opanował wiedzę z zakresu teorii eksploracji danych	Student opanował wiedzę z zakresu teorii eksploracji danych, potrafi wskazać zakres jej zastosowania	Student bardzo dobrze opanował wiedzę z zakresu materiału objętego programem nauczania, samodzielnie

				zdo­bywa i poszerza wiedzę przy użyciu róż­nych źródeł
Efekt 2	Student nie opanował wiedzy z zakresu budowy hurtowni danych	Student częściowo opanował wiedzę z zakresu budowy hurtowni danych	Student opanował wiedzę z zakresu budowy hurtowni danych, potrafi wskazać zakres ich zastosowania	Student bardzo dobrze opanował wiedzę z zakresu materiału objętego programem nauczania, samodzielnie zdobywa i poszerza wiedzę przy użyciu róż­nych źródeł
Efekt 3	Student nie opanował podstawowej wiedzy z zakresu inteligencji biznesowej	Student częściowo opanował podstawowej wiedzy z zakresu inteligencji biznesowej	Student opanował wiedzę z zakresu inteligencji biznesowej	Student bardzo dobrze opanował wiedzę z zakresu materiału objętego programem nauczania, potrafi wskazać problemy związane z każdą z nich
Efekt 4	Student nie potrafi posługiwać się eksploracją baz analitycznych	Student nie potrafi wykorzystać zdobytej wiedzy, zadania wynikające z realizacji ćwiczeń wykonuje z pomocą prowadzącego	Student poprawnie wykorzystuje wiedzę oraz samodzielnie rozwiązuje problemy wynikające w trakcie realizacji ćwiczeń	Student potrafi posługiwać się eksploracją baz analitycznych, potrafi dokonać oceny oraz uzasadnić trafność przyjętych metod
Efekt 5	Student nie potrafi dobrać odpowiedniego algorytmu analizy danych	Student nie potrafi wykorzystać zdobytej wiedzy, zadania wynikające z realizacji ćwiczeń wykonuje z pomocą prowadzącego	Student poprawnie wykorzystuje wiedzę oraz samodzielnie rozwiązuje problemy wynikające w trakcie realizacji ćwiczeń	Student potrafi dobrać odpowiedni algorytm analizy danych, potrafi dokonać oceny oraz uzasadnić trafność przyjętych metod
Efekt 6	Student nie potrafi posługiwać się diagramami organizacji informacji	Student nie potrafi wykorzystać zdobytej wiedzy, zadania wynikające z realizacji ćwiczeń wykonuje z pomocą prowadzącego	Student poprawnie wykorzystuje wiedzę oraz samodzielnie rozwiązuje problemy wynikające w trakcie realizacji ćwiczeń	Student potrafi posługiwać się diagramami organizacji informacji, potrafi dokonać oceny oraz uzasadnić trafność przyjętych metod
Efekt 7	Student nie opracował sprawozdania/ Student nie potrafi zaprezentować wyników swoich badań	Student wykonał sprawozdanie z wykonanego ćwiczenia, ale nie potrafi dokonać interpretacji oraz analizy wyników własnych badań	Student wykonał sprawozdanie z wykonanego ćwiczenia, potrafi prezentować wyniki swojej pracy oraz dokonuje ich analizy	Student wykonał sprawozdanie z wykonanego ćwiczenia, potrafi w sposób zrozumiały prezentować, oraz dyskutować osiągnięte wyniki

IV. INNE PRZYDATNE INFORMACJE O PRZEDMIOCIE

1. Wszelkie informacje dla studentów (prezentacje do zajęć, instrukcje do ćwiczeń)

laboratoryjnych, przykładowe aplikacje) dostępne są na stronie internetowej <http://iisi.pcz.pl/ClaDM/>, w zakładce Dydaktyka.

2. Informacja na temat konsultacji przekazywana jest studentom podczas pierwszych zajęć danego z przedmiotu.