Module/Course Syllabus
Program: COMPUTER SCIENCE
Full-time master degree program

Course:	**Advanced data mining**
Type of the course:	directional
Course code:	I2S1.06
Year:	I
Semester:	I
Form of the degree program:	full-time
Form of classes and number of hours per semester:	60
Lecture	30
Classes	0
Laboratory	30
Project	0
Number of ECTS credits:	4
Form of assessment:	exam
Course language:	English

Course objective (CO)	
CO1	Introduction to advanced methods of statistical data analysis.
CO2	Presentation of selected methods of categorization and prediction.
CO3	Introduction to advanced methods of artificial intelligence used in data analysis.

Prerequisites in terms of knowledge, skills and other competencies	
1	Structural and object-oriented programming.
2	Fundamentals of probability and statistics.
3	

Learning outcomes (LO)	
	In terms of knowledge:
LO 1	Has basic knowledge of statistical measures used in data analysis.
LO 2	Has basic knowledge of artificial intelligence methods in data analysis.
LO 3	Knows selected methods of categorization and prediction.
	In terms of skills:
LO 4	He can examine data integrity and detect anomalies.
LO 5	He can use selected artificial intelligence tools in classification and prediction problems.
	In terms of social competence:
LO 6	He is able to make responsible design decisions, also in the non-technical area.

Course content	
Form of classes – lectures (L)	
	Course content
L1	Basic statistical measures used in data analysis. Interdependence measures, ANOVA and MANOVA.
L2	Basic techniques of statistical reasoning. Data integrity.
L3	Selected anomaly detection techniques. Basic categorization and prediction techniques.

L4	Comparison of categorization and prediction methods, linear regression. Generalized linear and non-linear models.
L5	Overview of artificial intelligence methods used in data analysis.
L6	Decision trees, random forest, artificial neural networks.
L7	Data dimension reduction methods, principal component analysis (PCA).
L8	Association rules. Visualization and reporting.
Form of classes – laboratories (Lab)	
	Course content
Lab1	Introduction, discussion of the rules of passing.
Lab2	Discussion of analytical tools used in classes.
Lab3	Downloading data and preparing for analysis.
Lab4	Determination of basic statistical measures using statistical packages and modules.
Lab5	Data integrity testing.
Lab6	Individual project in the field of data preparation, integrity testing, determination and interpretation of statistical measures.
Lab7	Presentation of individual projects.
Lab8	Categorization and prediction.
Lab9	Generalized linear, polynomial and logistic regression, advanced methods of building models.
Lab10	Advanced techniques of artificial intelligence in data analysis using available software.
Lab11	Dimensional reduction and basket analysis.
Lab12	Basic techniques of graphical presentation of results, types of charts, selection of variables.
Lab13	Team project - data acquisition and processing.
Lab14	Team project - setting models, creating static and dynamic reports.
Lab15	Presentation of group projects.

Didactic methods	
1	Lecture with multimedia presentation
2	Laboratory exercises: implementation of tasks using the known methods of data analysis
3	Laboratory exercises: thematic discussion

Assessment methods and criteria		
Assessment method symbol	**Assessment method description**	**Passing threshold**
A1	Individual and group projects	**75%**
A2	Exam	**51%**

Required textbooks and other course materials	
1	Cuesta, H. (2013). Practical data analysis. Packt Publishing Ltd.
2	Dean, J. (2014). Big data, data mining, and machine learning: value creation for business leaders and practitioners. John Wiley& Sons.
Recommended textbooks and other course materials	
1	Aggarwal C. (2015). Data Mining. Springer

Student workload	
Form of activity	**Average number of hours to complete the activity**
Contact hours with the lecturer, including:	60

participation in lectures	30
participation in laboratories	30
Student's own work, including:	40
preparation for the exam	10
preparation for the laboratory	15
preparation project	15
Total student workload	100
Total number of ECTS credits	4

Learning outcomes matrix					
Learning outcome	**Reference to learning outcomes defined for the master's program**	**Course objectives**	**Course content**	**Didactic methods**	**Assessment methods**
L01	I2A_W06 +++	*CO1*	*L1, L2*	1	*A2*
L02	I2A_W08 ++	*CO3*	*L5-L8*	1	*A2*
L03	I2A_W10 ++	*CO2*	*L3, L4*	1	*A2*
L04	I2A_U04 ++	*CO1*	*Lab1-Lab7*	2, 3	*A1*
L05	I2A_U15 ++	*CO3*	*Lab8 -Lab 14*	2, 3	*A1*
L06	I2A_K01 +	*CO1-CO3*	*Lab 7- Lab 14*	2, 3	*A1*

The author of the program:	dr Adam Kiersztyn
E-mail address:	a.kiersztyn@pollub.pl
Organizational unit:	Department of Computer Science