Faculty of Fundamental Problems of Technology
COURSE CARD

Name in polish	:	**Eksploracja Danych**
Name in english	:	**Data Mining**
Field of study	:	Computer Science
Specialty (if applicable)	:	
Undergraduate degree and form of	:	masters, stationary
Type of course	:	optional
Course code	:	E2_W14
Group rate	:	Yes

	Lectures	Exercides	Laboratory	Project	Seminar
Number of classes held in schools (ZZU)	30	15	15		
The total number of hours of student workload (CNPS)	70	55	55		
Assesment	pass				
For a group of courses final course mark	X				
Number of ECTS credits	2	2	2		
including the number of points corresponding to the classes of practical (P)		2	2		
including the number of points corresponding occupations requiring direct contact (BK)	2	2	2		

PREREQUISITES FOR KNOWLEDGE, SKILLS AND OTHER POWERS
It is required to pass the following modules: Introduction to the Computer Science and Programming, Data Bases and Information Managements, Logic and Formal Structures, Probabilistic Methods and Statistic.

COURSE OBJECTIVES

C1 Presentation of the methods of data mining

C2 Profound understanding of the presented data mining methods

C3 Ability to use selected algorithms in practice

COURSE LEARNING OUTCOMES
The scope of the student's knowledge:

W1 Knows the data mining algorithms

W2 Knows the applicatinon of the data mining algorithms

The student skills:

U1 Can use the data mining algorithms in practice

U2 Can use the Apache Spark platform for efficient processing of large datasets

The student's social competence:

K1 Has the ability to cooperate with other experts specialized in data mining algorithms

COURSE CONTENT		

Type of classes - lectures		
Wy1	Introduction to the Data Mining	2h
Wy2	Building and evaluating the model	2h
Wy3	Linear regression and related methods	4h
Wy4	Resampling methods	2h
Wy5	Classification algororithms	6h
Wy6	Dimensionality reduction	4h
Wy7	Unsupervised learning	2h
Wy8	Effective implementation of machine learning algorithms	4h
Wy9	Analysis of data streams	4h

Type of classes - exercises		
Ćw1	Model design and evaluation	2h
Ćw2	Linear regression	2h
Ćw3	Resampling methods	2h
Ćw4	Classification algororithms	5h
Ćw5	Dimensionality reduction	2h
Ćw6	Unsupervised learning	2h

Type of classes - laboratory		
Lab1	Preparing Data for Mining	2h
Lab2	Linear regression and related methods	2h
Lab3	Classification algororithms	4h
Lab4	Clustering algororithms	2h
Lab5	Introduction Apache Spark	5h

Applied learning tools		

1. Traditional lecture

2. Multimedia lecture

3. Solving tasks and problems

4. Solving programming tasks

5. Creating programming projects

6. Self-study students

EVALUATION OF THE EFFECTS OF EDUCATION ACHIEVEMENTS		

Value	Number of training effect	Way to evaluate the effect of education
F1	W1-W2, K1-K1	Test
F2	U1-U2, K1-K1	Activity
F3	U1-U2, K1-K1	Implementation and presentation of solutions
P=40%*F1+30%*F2+30%*F3		

BASIC AND ADDITIONAL READING
1. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, T.Hastie, R. Tibshirani, J.Friedman, 2009
2. Mining of Massive Datasets, J.Leskovec, A.Rajaraman, J. Ullman, 2010
3. Big Data Analytics with Spark, M. Guller, 2015

SUPERVISOR OF COURSE
dr inż. Jakub Lemiesz

RELATIONSHIP MATRIX EFFECTS OF EDUCATION FOR THE COURSE
Data Mining
WITH EFFECTS OF EDUCATION ON THE DIRECTION OF COMPUTER SCIENCE

Course training effect	Reference to the effect of the learning outcomes defined for the field of study and specialization (if applicable)	Objectives of the course**	The contents of the course**	Number of teaching tools**
W1	K2_W01 K2_W02	C1	Wy1-Wy9	1 2 6
W2	K2_W02	C1	Wy1-Wy9	1 2 6
U1	K2_U09_A	C2 C3	Ćw1-Ćw6 Lab1-Lab5	3 4 5 6
U2	K2_U01_A K2_U08_A K2_U09_A K2_U15	C2 C3	Ćw1-Ćw6 Lab1-Lab5	3 4 5 6
K1	K2_K14_A	C1 C2 C3	Wy1-Wy9 Ćw1-Ćw6 Lab1-Lab5	1 2 3 4 5 6