



Module name: Data mining

Academic year: 2015/2016 Code: BOS-1-517-s ECTS credits: 3

Faculty of: Geology, Geophysics and Environmental Protection

Field of study: Environmental Protection Specialty: —

Study level: First-cycle studies Form and type of study: Full-time studies

Lecture language: English Profile of education: Academic (A) Semester: 5

Course homepage: <http://www.adamwalanus.pl/Dyd.html>

Responsible teacher: dr inż. Chuchro Monika (chuchro@geol.agh.edu.pl)

Academic teachers: prof. dr hab. inż. Walanus Adam (a@adamwalanus.pl)

## Description of learning outcomes for module

MLO code	Student after module completion has the knowledge/ knows how to/is able to	Connections with FLO	Method of learning outcomes verification (form of completion)
Skills			
M_U001	Potrafi obliczyć i zinterpretować statystyki jednej zmiennej	OS1A_U05, OS1A_U15	Execution of laboratory classes
M_U002	Potrafi zinterpretować macierz korelacji	OS1A_U05, OS1A_U15	Execution of laboratory classes
M_U003	Potrafi wykonać regresję w MS Excel, Matlab i Statistica	OS1A_U05, OS1A_U15	Execution of laboratory classes
M_U004	Potrafi zbadać jakość danych	OS1A_U05, OS1A_U15, OS1A_U14	Execution of laboratory classes
Knowledge			
M_W001	Zna podstawy statystyki	OS1A_W07	Execution of laboratory classes

## FLO matrix in relation to forms of classes

MLO code	Student after module completion has the knowledge/ knows how to/is able to	Form of classes
----------	--	-----------------

		Lectures	Auditorium classes	Laboratory classes	Project classes	Conversation seminar	Seminar classes	Practical classes	Fieldwork classes	Workshops	Others	E-learning
Skills												
M_U001	Potrafi obliczyć i zinterpretować statystyki jednej zmiennej	+	-	-	+	-	-	-	-	-	-	-
M_U002	Potrafi zinterpretować macierz korelacji	+	-	-	+	-	-	-	-	-	-	-
M_U003	Potrafi wykonać regresję w MS Excel, Matlab i Statistica	+	-	-	+	-	-	-	-	-	-	-
M_U004	Potrafi zbadać jakość danych	+	-	-	+	-	-	-	-	-	-	-
Knowledge												
M_W001	Zna podstawy statystyki	+	-	-	+	-	-	-	-	-	-	-

## Module content

### Lectures

1. Variables, experimental data, measurements, scales, dependent vs. independent variables
2. Relations between variables
3. Statistical significance of results (p-value)
4. Normal distribution and other probability distributions
5. Basic statistics
6. Regression
7. Classification
8. Multivariate analysis
9. Data quality, data clearing, transformations
10. Multidimensional scaling
11. Machine Learning
12. Data mining in industrial engineering

### Project classes

1. Real measurements with a tape measure, organizing spreadsheet
2. Modelling random variables (Excel)
3. Searching for extreme values, estimating low probabilities
4. Modelling normal distribution
5. Calculation of basic statistics
6. Calculation of Regression
7. Performing classification (Statistica)
8. Performing multivariate analysis (Statistica)
9. Transforming data, Box-Cox
10. Statistica – Data miner I
11. Statistica – Data miner II, machine Learning
12. Industrial statistics, SPC, QCC

### **Method of calculating the final grade**

Średnia z ocen zdobywanych w trakcie ćwiczeń laboratoryjnych

### **Prerequisites and additional requirements**

Znajomość metod numerycznych i elementów programowania

### **Recommended literature and teaching resources**

WWW, Pomoc MS Excel, Pomoc i Podręcznik Statystyki Statistica  
<http://www.statsoft.pl/textbook/stathome.html>

### **Scientific publications of module course instructors related to the topic of the module**

Additional scientific publications not specified

### **Additional information**

None

### **Student workload (ECTS credits balance)**

Student activity form	Student workload
Participation in lectures	28 h
Participation in project classes	14 h
Realization of independently performed tasks	30 h
Preparation for classes	15 h
Summary student workload	87 h
Module ECTS credits	3 ECTS