



Module name: Principles of Data and Process Mining

Academic year: 2019/2020 Code: GIPZ-2-402-CP-n ECTS credits: 3

Faculty of: Mining and Geoengineering

Field of study: - Specialty: Controlling of production processes

Study level: Second-cycle studies Form and type of study: Part-time studies

Lecture language: English Profile of education: Academic (A) Semester: 4

Course homepage: —

Responsible teacher: prof. nadzw. dr hab. inż. Brzychczy Edyta (brzych3@agh.edu.pl)

Module summary

Main issues covered by lectures are related to data mining and process mining techniques as well as tools used in advanced analytics.

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)
Social competence: is able to			
M_K001	Student can perform data or process analysis in project group.	IPZ2A_K01	Activity during classes
Skills: he can			
M_U001	Student can perform data mining analysis	IPZ2A_U02, IPZ2A_U01, IPZ2A_U04	Completion of laboratory classes
M_U002	Student can perform process mining analysis.	IPZ2A_U02, IPZ2A_U01, IPZ2A_U04	Completion of laboratory classes
Knowledge: he knows and understands			
M_W001	Student knows principles of knowledge discovery process in enterprise.	IPZ2A_W02	Test
M_W002	Student knows selected methods of Data Mining and Machine Learning.	IPZ2A_W02	Completion of laboratory classes, Test
M_W003	Student knows selected issues related to process mining methods.	IPZ2A_W02	Completion of laboratory classes, Test

Number of hours for each form of classes

Suma	Form of classes										
	Lectures	Auditorium classes	Laboratory classes	Project classes	Conversation seminar	Seminar classes	Practical classes	Fieldwork classes	Workshops	Prace kontrolne i przejściowe	Lektorat
15	9	0	6	0	0	0	0	0	0	0	0

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	Prace kontrolne i przejściowe	Lektorat
Social competence: is able to												
M_K001	Student can perform data or process analysis in project group.	-	-	+	-	-	-	-	-	-	-	-
Skills: he can												
M_U001	Student can perform data mining analysis	-	-	+	-	-	-	-	-	-	-	-
M_U002	Student can perform process mining analysis.	-	-	+	-	-	-	-	-	-	-	-
Knowledge: he knows and understands												
M_W001	Student knows principles of knowledge discovery process in enterprise.	+	-	-	-	-	-	-	-	-	-	-
M_W002	Student knows selected methods of Data Mining and Machine Learning.	+	-	-	-	-	-	-	-	-	-	-
M_W003	Student knows selected issues related to process mining methods.	+	-	-	-	-	-	-	-	-	-	-

Student workload (ECTS credits balance)

Student activity form	Student workload
Udział w zajęciach dydaktycznych/praktyka	15 h
Preparation for classes	30 h
Realization of independently performed tasks	30 h
Examination or Final test	2 h
Contact hours	1 h
Summary student workload	78 h
Module ECTS credits	3 ECTS

Additional information

Module content

Lectures

1. Knowledge Discovery Process in Databases
2. Exploratory tasks in Data Mining
3. Selected method of Data Mining: cluster analysis, decision trees, association rules
4. Introduction to Process Mining
5. Process discovery methods
6. Conformance checking and enhancement of the process
7. Mining of additional process perspectives

Laboratory classes

1. Introduction to STATISTICA software and exploratory data analysis
2. Cluster analysis, decision trees, association rules
3. Introduction to ProM software
4. Process discovery from event logs
5. Conformance checking and enhancement of the process
6. Additional process perspectives analysis

Teaching methods and techniques:

Lectures: The content presented at the lecture is provided in the form of a multimedia presentation in combination with a classical lecture panel enriched with demonstrations relating to the presented issues .

Laboratory classes: During the laboratory classes, students independently solve the practical problem, choosing the right tools. The leader stimulates the group to reflect on the problem, so that the obtained results have a high substantive value.

Warunki i sposób zaliczenia poszczególnych form zajęć, w tym zasady zaliczeń poprawkowych, a także warunki dopuszczenia do egzaminu:

The final grade is based on the test and the grade from laboratory classes.
2 attempts for each element are given.

Zasady udziału w poszczególnych zajęciach, ze wskazaniem, czy obecność

studenta na zajęciach jest obowiązkowa:

Lectures:

- Attendance is mandatory: No
- Participation rules in classes: Students participate in the classes learning the next teaching content according to the syllabus of the subject. Students should constantly ask questions and explain doubts. Audiovisual recording of the lecture requires the teacher's consent.

Laboratory classes:

- Attendance is mandatory: Yes
- Participation rules in classes: Students carry out laboratory exercises in accordance with materials provided by the teacher.

Method of calculating the final grade

The final grade is based on the test and the laboratory classes (both positive grades) and is calculated according to the algorithm: $OK = 0.5 \cdot T + 0.5 \cdot L$

where: T- test grade, L - laboratory grade

Sposób i tryb wyrównywania zaległości powstałych wskutek nieobecności studenta na zajęciach:

Laboratory classes need to be completed with other group.
Additional presentation can be required on given topic.

Prerequisites and additional requirements

Basic statistical background is needed.

Recommended literature and teaching resources

Larose T.D., 2006: Odkrywanie wiedzy z danych. Wyd. Naukowe PWN, Warszawa

Hand D., Mannila H., Smyth P., 2002: Principles of Data Mining, MIT press

Van der Aalst W., 2016: Process Mining Data Science in Action, 2nd edition, Springer

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

Brzychczy E., 2007: Budowa modeli ekonometrycznych wybranych parametrów techniczno-ekonomicznych kopalni węgla kamiennego. Wiadomości Górnicze, 11

Brzychczy E., 2008: Analiza wykorzystania kombajnów chodnikowych w przodkach korytarzowych w kopalniach węgla kamiennego z zastosowaniem drzew decyzyjnych. Zarządzanie: doświadczenia i problemy. Red. W. Sitko. Wyd. System-Graf, Lublin

Brzychczy E., Stefaniak R., Maroszek Z., Siodlak Ł., 2010: Wykorzystanie wybranych technik Data Mining do analizy kompleksów ścianowych w KWK "Ziemowit". Miesięcznik WUG Bezpieczeństwo i Ochrona Pracy w Górnictwie, nr 1

Brzychczy E., 2009: Techniki eksploracji danych w zagadnieniach eksploatacji górniczej złóż węgla kamiennego. Kwartalnik AGH, Górnictwo i Geoinżynieria, nr 3

Brzychczy E., 2009: Analiza wyposażenia przodków ścianowych na podstawie reguł asocjacyjnych. Wiadomości Górnicze, R.60, nr 3

Trzcionkowska A., Brzychczy E., 2016: Wykorzystanie reguł asocjacyjnych do analizy pracy wybranego urzędnika w oddziale wydobywczym. Inżynieria Mineralna, R. 17, nr 2

Additional information

Written test - 5 questions.

Laboratory grade is based on realisation of individual data or process mining task.