

**AGH**AGH UNIVERSITY OF SCIENCE  
AND TECHNOLOGY

Nazwa modułu zajęć: Data mining and machine learning in civil engineering problems

Rok akademicki: 2019/2020 Kod: GBUD-2-304-GT-n Punkty ECTS: 3

Wydział: Górnictwa i Geoinżynierii

Kierunek: Budownictwo Specjalność: Geotechnika i budownictwo specjalne

Poziom studiów: Studia II stopnia Forma studiów: Niestacjonarne

Język wykładowy: Polski Profil: Ogólnoakademicki (A) Semestr: 3

Strona www: —

Prowadzący moduł: dr hab. inż. prof. AGH Rusek Janusz (rusek@agh.edu.pl)

**Treści programowe zapewniające uzyskanie efektów uczenia się dla modułu zajęć**

Problems of classification and regression as well as dimensionality reduction in data sets on building objects and operational impacts using Data-Mining and Machine Learning methods.

**Opis efektów uczenia się dla modułu zajęć**

Kod MEU	Student, który zaliczył moduł zajęć zna i rozumie/potrafi/jest gotów do	Powiązania z KEU	Sposób weryfikacji i oceny efektów uczenia się osiągniętych przez studenta w ramach poszczególnych form zajęć i dla całego modułu zajęć
Wiedza: zna i rozumie			
M_W001	<ul style="list-style-type: none"><li>- formulation of the objective function used in the process of learning artificial neural networks (ANN) and related systems,</li><li>- regularization and stabilizing methods of the learning process,</li><li>- methods of gradient optimization and derivative-free optimization on the example of Pattern Search and genetic algorithms and related methods (PSO etc.),</li><li>- construction of the objective function for the Support Vector Machine (SVM) method in the classification and regression problem,</li><li>- methods for assessing the effectiveness of learning in terms of generalization.</li></ul>	BUD2A_W01, BUD2A_W04	Kolokwium

M_W002	- methods for assessing the quality of data for further processing, - Data-Mining methods used to reduce the dimensionality of data and extract important features (feature selection), - methods of preliminary analysis of the linear relationship between variables. Basic concepts from the area of classification and regression models.	BUD2A_W02, BUD2A_W04	Kolokwium
Umiejętności: potrafi			
M_U001	- create regression and classification models based on available data using the Python programming environment and the Matlab platform, - use models for individual simulation purposes and integration with other parts of the code, - assess the significance of the impact of individual input variables on the final neural network or SVM system.	BUD2A_W01, BUD2A_K04, BUD2A_U03	Projekt
M_U002	- form classification and regression problems based on data, - separate the training, validation and test set for artificial neural network and SVM models, - write optimization algorithms used to build neural networks.		Kolokwium
Kompetencje społeczne: jest gotów do			
M_K001	work in a team while doing design exercises and exchange information.	BUD2A_K01, BUD2A_K04, BUD2A_U04	Projekt

### Liczba godzin zajęć w ramach poszczególnych form zajęć

Suma	Forma zajęć dydaktycznych										
	Wykład	Ćwiczenia audytoryjne	Ćwiczenia laboratoryjne	Ćwiczenia projektowe	Konwersatorium	Zajęcia seminaryjne	Zajęcia praktyczne	Zajęcia terenowe	Zajęcia warsztatowe	Prace kontrolne i przejściowe	Lektorat
15	9	0	0	6	0	0	0	0	0	0	0

### Matryca kierunkowych efektów uczenia się w odniesieniu do form zajęć i sposobu zaliczenia, które pozwalają na ich uzyskanie

Kod MEU	Student, który zaliczył moduł zajęć zna i rozumie/potrafi/jest gotów do	Forma zajęć dydaktycznych
---------	---	---------------------------

		Wykład	Ćwiczenia audytoryjne	Ćwiczenia laboratoryjne	Ćwiczenia projektowe	Konwersatorium	Zajęcia seminaryjne	Zajęcia praktyczne	Zajęcia terenowe	Zajęcia warsztatowe	Prace kontrolne i przejściowe	Lektorat
Wiedza: zna i rozumie												
M_W001	- formulation of the objective function used in the process of learning artificial neural networks (ANN) and related systems, - regularization and stabilizing methods of the learning process, - methods of gradient optimization and derivative-free optimization on the example of Pattern Search and genetic algorithms and related methods (PSO etc.), - construction of the objective function for the Support Vector Machine (SVM) method in the classification and regression problem, - methods for assessing the effectiveness of learning in terms of generalization.	+	-	-	+	-	-	-	-	-	-	-
M_W002	- methods for assessing the quality of data for further processing, - Data-Mining methods used to reduce the dimensionality of data and extract important features (feature selection), - methods of preliminary analysis of the linear relationship between variables. Basic concepts from the area of classification and regression models.	+	-	-	+	-	-	-	-	-	-	-
Umiejętności: potrafi												
M_U001	- create regression and classification models based on available data using the Python programming environment and the Matlab platform, - use models for individual simulation purposes and integration with other parts of the code, - assess the significance of the impact of individual input variables on the final neural network or SVM system.	-	-	-	+	-	-	-	-	-	-	-
M_U002	- form classification and regression problems based on data, - separate the training, validation and test set for artificial neural network and SVM models, - write optimization algorithms used to build neural networks.	-	-	-	+	-	-	-	-	-	-	-

Kompetencje społeczne: jest gotów do													
M_K001	work in a team while doing design exercises and exchange information.	-	-	-	+	-	-	-	-	-	-	-	-

## Nakład pracy studenta (bilans punktów ECTS)

Forma aktywności studenta	Obciążenie studenta
Udział w zajęciach dydaktycznych/praktyka	15 godz
Przygotowanie do zajęć	10 godz
przygotowanie projektu, prezentacji, pracy pisemnej, sprawozdania	20 godz
Samodzielne studiowanie tematyki zajęć	30 godz
Egzamin lub kolokwium zaliczeniowe	2 godz
Dodatkowe godziny kontaktowe	1 godz
Sumaryczne obciążenie pracą studenta	78 godz
Punkty ECTS za moduł	3 ECTS

## Pozostałe informacje

### Szczegółowe treści kształcenia w ramach poszczególnych form zajęć (szczegółowy program wykładów i pozostałych zajęć)

#### Wykład

The course will discuss methods in the field of Data Mining and Machine Learning. The mathematical foundations of these methods and examples of their application will be presented. The problems of classification and regression in a linear and nonlinear approach will be discussed. The methods of reducing the size of the data originally collected will be presented. Examples will be presented of the use of these methods in the area of information about civil engineering construction objects.

#### Ćwiczenia projektowe

Execution of design exercises within which, for an available set of data on the technical characteristics of construction objects and operational impacts, classification and regression problems will be formulated. Methods of artificial neural networks and the SVM method will be used for this. Finished models will be tested for correctness and generalization. Different optimization methods for teaching artificial neural networks will be tested. The SVM method will analyze the selection of parameters resulting from the original form of the objective function.

#### Metody i techniki kształcenia:

Wykład: Classes conducted with the use of multimedia presentation tools.

Ćwiczenia projektowe: Classes conducted with the use of multimedia presentation tools and collective work in specific programming environments in the classroom – the need to take classes in the computer

room.

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

attendance at classes: obligatory (2 unjustified absences)

attendance at lectures: indicated (optional)

### **Zasady udziału w poszczególnych zajęciach, ze wskazaniem, czy obecność studenta na zajęciach jest obowiązkowa:**

Wykład:

- Obecność obowiązkowa: Nie

- Zasady udziału w zajęciach: Not specified because lectures are not mandatory

Ćwiczenia projektowe:

- Obecność obowiązkowa: Tak

- Zasady udziału w zajęciach: Presence and active participation in the performance of individual programming parties included in the syllabus content

### **Sposób obliczania oceny końcowej**

Degree of the test (lecture) X 0.5 + Degree of the project X 0.25 + Degree of the test (classes) X 0.25

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

in the student's own scope

### **Wymagania wstępne i dodatkowe, z uwzględnieniem sekwencyjności modułów**

- basic knowledge of mathematical analysis, algebra and probabilistics,

- ability to act on vectors and matrices.

### **Zalecana literatura i pomoce naukowe**

OSOWSKI, Stanisław. Metody i narzędzia eksploracji danych. Legionowo: Wydawnictwo BTC, 2013.

OSOWSKI, S. Sieci neuronowe do przetwarzania informacji, Oficyna Wydawnicza Pol. Warszawskiej, Warszawa, 2006.

ŁĘSKI, Jacek. Systemy neuronowo-rozmyte. Wydawnictwa Naukowo-Techniczne, 2008.

CICHOSZ, Paweł. Systemy uczące się. Wydawnictwa Naukowo-Techniczne, 2000.

### **Publikacje naukowe osób prowadzących zajęcia związane z tematyką modułu**

RUSEK, Janusz. Modelowanie stopnia zużycia technicznego budynków na terenach górniczych z wykorzystaniem wybranych metod sztucznej inteligencji. Wydawnictwa AGH, 2013.

RUSEK, Janusz. Creating a model of technical wear of building in mining area, with utilization of regressive SVM approach. Archives of Mining Sciences, 2009, 54.3: 455-466.

FIREK, K.; RUSEK, Janusz; WODYŃSKI, Aleksander. Wybrane metody eksploracji danych i uczenia maszynowego w analizie stanu uszkodzeń oraz zużycia technicznego zabudowy terenów górniczych. Przegląd Górniczy, 2016, 72.1: 50-55.

### **Informacje dodatkowe**

-