

**AGH**AGH UNIVERSITY OF SCIENCE
AND TECHNOLOGY

Nazwa modułu: Statystyczna analiza danych internetowych

Rok akademicki: 2016/2017 Kod: ZIPM-3-015-s Punkty ECTS: 3

Wydział: Zarządzania

Kierunek: Inżynieria Produkcji Metali Nieżelaznych Specjalność: —

Poziom studiów: Studia III stopnia Forma i tryb studiów: Stacjonarne

Język wykładowy: Polski Profil kształcenia: Ogólnoakademicki (A) Semestr: 0

Strona www: <http://www.zarz.agh.edu.pl>

Osoba odpowiedzialna: Duda Jan Tadeusz (jtduda@zarz.agh.edu.pl)

Osoby prowadzące: Duda Jan Tadeusz (jtduda@zarz.agh.edu.pl)

Opis efektów kształcenia dla modułu zajęć

Kod EKM	Student, który zaliczył moduł zajęć wie/umie/potrafi	Powiązania z EKK	Sposób weryfikacji efektów kształcenia (forma zaliczeń)
Wiedza			
M_W001	Słuchacz kursu ma wiedzę na temat dostępnych źródeł informacji internetowych o charakterze ilościowym i jakościowym.	IPM3A_W01, IPM3A_W03	Odpowiedź ustna, Projekt, Aktywność na zajęciach, Udział w dyskusji
M_W002	Potrafi dokonać analizy poprawności i spójności danych oraz ocenić ich przydatność dla rozwiązania postawionego zadania z zakresu Inżynierii Produkcji. Zna odpowiednie narzędzia statystyki.	IPM3A_W02	Egzamin
Umiejętności			
M_U001	Potrafi przygotować w środowisku Matlab program przetwarzający wybrane dane internetowe. Potrafi ocenić poprawność wyników przetwarzania danych. Potrafi zastosować odpowiednie narzędzia statystyki.	IPM3A_U05, IPM3A_U03	Odpowiedź ustna, Projekt, Aktywność na zajęciach, Udział w dyskusji
Kompetencje społeczne			
M_K001	Ma świadomość znaczenia i odpowiedzialności poprawnego przetwarzania danych dla potrzeb inżynierii produkcji.	IPM3A_K01, IPM3A_K03	Aktywność na zajęciach, Odpowiedź ustna, Udział w dyskusji

Matryca efektów kształcenia w odniesieniu do form zajęć

Kod EKM	Student, który zaliczył moduł zajęć wie/umie/potrafi	Forma zajęć										
		Wykład	Ćwiczenia audytorijne	Ćwiczenia laboratoryjne	Ćwiczenia projektowe	Konwersatorium	Zajęcia seminaryjne	Zajęcia praktyczne	Zajęcia terenowe	Zajęcia warsztatowe	Inne	E-learning
Wiedza												
M_W001	Słuchacz kursu ma wiedzę na temat dostępnych źródeł informacji internetowych o charakterze ilościowym i jakościowym.	-	-	-	-	+	-	-	-	-	-	-
M_W002	Potrafi dokonać analizy poprawności i spójności danych oraz ocenić ich przydatność dla rozwiązania postawionego zadania z zakresu Inżynierii Produkcji. Zna odpowiednie narzędzia statystyki.	-	-	-	-	+	-	-	-	-	-	-
Umiejętności												
M_U001	Potrafi przygotować w środowisku Matlab program przetwarzający wybrane dane internetowe. Potrafi ocenić poprawność wyników przetwarzania danych. Potrafi zastosować odpowiednie narzędzia statystyki.	-	-	-	-	+	-	-	-	-	-	-
Kompetencje społeczne												
M_K001	Ma świadomość znaczenia i odpowiedzialności poprawnego przetwarzania danych dla potrzeb inżynierii produkcji.	-	-	-	-	+	-	-	-	-	-	-

Treść modułu zajęć (program wykładów i pozostałych zajęć)**Konwersatorium**

1. Pojęcia podstawowe – formy i źródła danych, Internet jako źródło danych, kwestia wiarygodności i aktualności danych, prawa autorskie,

Cele przetwarzania – wydobywanie wiedzy z danych.

Typy danych – jakościowe (lingwistyczne), ilościowe – całkowite, rzeczywiste.

Metody wydobywania wiedzy – grupowanie i klasyfikacja (wiedza jakościowa), modelowanie matematyczne (wiedza ilościowa).

Szeregi czasowe (problem miary czasu i kompletności).

Formaty danych – binarne (nie dostępne) HTML, XML, csv.

Narzędzia: algorytmy grupowania i klasyfikacji, statystyka opisowa, ilościowe analizy statystyczne jedno i wielowymiarowe (kwestia losowości danych),

Narzędzia przetwarzania: MS Excel, Visual Basic – problem eksportu i importu danych,

prezentacja graficzna: szeregi czasowe a wykresy przekrojowe – interpretacja.
2.Statystyka jako narzędzie wydobywania wiedzy: rozkłady prawdopodobieństwa jedno i wielowymiarowe, analizy korelacyjne, autokorelacji – problem kompletności.
3.Modelowanie nieparametryczne i parametryczne.
4.Regresja jako narzędzie wydobywania wiedzy.
5.Grupowanie szeregów czasowych – analizy podobieństwa (definicje podobieństwa), prefiltracja jako narzędzie dekompozycji szeregów.

Sposób obliczania oceny końcowej

Ocena końcowa wystawiana jest przez prowadzącego na podstawie oceny z egzaminu, z uwzględnieniem obecności i aktywności na konwersatoriach.
Ocenę zaokrągla się zależnie od obecności, z możliwością podniesienia (maks. o 1.0) na podstawie dodatkowego kolokwium ustnego.

Wymagania wstępne i dodatkowe

Znajomość podstawowych metod statystyki matematycznej i algebry, umiejętność programowania w środowisku MATLAB

Zalecana literatura i pomoce naukowe

J. Koronacki, J. Mielniczuk: Statystyka dla studentów kierunków technicznych i przyrodniczych. WNT 2001.
T. J. Hastie, R. J. Tibshirani, J. Friedman: The Elements of Statistical Learning, Springer 2001
J. Koronacki, J. Ćwik, Statystyczne systemy uczące się, WNT 2005.
M. Sobczyk: Statystyka, PWN, 2014
M. Cieślak: Prognozowanie gospodardze. Metody i zastosowania, PWN, 2011

Publikacje naukowe osób prowadzących zajęcia związane z tematyką modułu

Duda J.T., Pełech-Pilichowski T.: Enhancements of moving trend based filters aimed at time series prediction. Advances in systems science: proceedings of the International Conference on Systems Science 2013 (ICSS 2013) / eds. Jerzy Świątek, [et al.], Advances in Intelligent Systems and Computing. Springer International Publishing, 2014, s. 747-756
Duda J.T., Pełech-Pilichowski T.: Wykorzystanie trendu pełzającego do analizy i prognozowania szeregów finansowych. W: Uwarunkowania, metody i strategie rozwoju przedsiębiorstw, red. nauk. Howaniec H., Madyda A., Waszkielewicz W., Wydawnictwo Naukowe ATH, 2013, s. 79-101
Pełech-Pilichowski T., Duda J.T.: A two-level detector of short-term unique changes in time series based on a similarity method. Expert Systems 2015 vol. 32 no. 4, s. 555-561
Pełech-Pilichowski T., Duda J.T.: Low-frequency signal reconstruction and abrupt change detection in non-stationary time series by enhanced moving trend based filters. Advances in ICT for business, industry and public sector, eds. Mach-Król M., Olszak C.M., Pełech-Pilichowski T. (Studies in Computational Intelligence, vol. 579), Springer, 2015, s. 111-125

Informacje dodatkowe

Brak

Nakład pracy studenta (bilans punktów ECTS)

Forma aktywności studenta	Obciążenie studenta
Dodatkowe godziny kontaktowe z nauczycielem	2 godz
Egzamin lub kolokwium zaliczeniowe	2 godz
Udział w konwersatoriach	14 godz
Samodzielne studiowanie tematyki zajęć	18 godz
Przygotowanie do zajęć	14 godz
Wykonanie projektu	25 godz
Sumaryczne obciążenie pracą studenta	75 godz
Punkty ECTS za moduł	3 ECTS