

# Metody i narzędzia eksploracji danych

ASPIRANT  
KATEDRY  
STUDIÓW DOKTORANCKICH  
Wydziału Elektroniki WAT  
*D. Dobrowolski*  
dr hab. inż. Andrzej DOBROWOLSKI

## Informacje ogólne

Kod przedmiotu: WELXXCXD-MNED

Kod Erasmus: (brak danych) 2014-106-27

Nazwa przedmiotu: Metody i narzędzia eksploracji danych

Jednostka: Wydział Elektroniki

Grupy:

Strona przedmiotu: <http://zese/wel.wat.edu.pl/sosowski/MNED.htm>

Punkty ECTS i **2.00** LUB 3.00 (w zależności od programu) **i**

inne: → zobacz reguły punktacji

Język prowadzenia: polski

Rodzaj studiów: III stopnia

Rodzaj przedmiotu: wybieralny

Forma zajęć liczba W30/+

godzin/rygor:

Przedmioty wprowadzające: Brak przedmiotów wprowadzających

Programy: Dyscyplina naukowa studiów: Elektronika, Telekomunikacja

Autor: prof. dr hab. inż. Stanisław Osowski

Skrócony opis: W ramach przedmiotu przedstawione zostaną następujące treści kształcenia:

Podstawowe pojęcia eksploracji danych, przegląd algorytmów optymalizacyjnych, metody regresji liniowej i logistycznej, klasyfikatory Bayes'a i drzewa decyzyjne, sztuczne sieci neuronowe MLP i RBF, sieć wektorów nośnych SVM, zespoły klasyfikatorów i predyktorów, metody oceny jakości klasyfikatorów i predyktorów, transformacje statystyczne PCA i ICA, wybrane metody generacji i selekcji cech diagnostycznych, algorytmy grupowania danych wielowymiarowych, zależności asocjacyjne między danymi, systemy rozmyte Takagi-Sugeno-Kanga, metody wizualizacja danych wielowymiarowych.

Pełny opis: Wykłady realizowane są w formie werbalno-wizualnej prezentacji następujących treści:

### 1. Podstawowe pojęcia eksploracji danych

Formy akwizycji danych, normalizacja danych, dane odstające i brakujące. przegląd zadań eksploracji danych i stosowanych narzędzi.

## 2. Algorytmy optymalizacyjne

Przegląd algorytmów gradientowych, algorytm symulowanego wyżarzania, algorytmy genetyczne i ewolucyjne.

## 3. Regresja liniowa i logistyczna

Regresja w zastosowaniach klasyfikacji i aproksymacji. Algorytmy regresji i ich implementacja w Matlabie.

## 4. Klasyfikatory Bayes'a i drzewa decyzyjne

Pełny i naiwny klasyfikator Bayes'a, implementacja naiwnego klasyfikatora Bayes'a w Matlabie, drzewa decyzyjne, algorytm tworzenia drzewa, implementacja drzewa decyzyjnego w Matlabie.

## 5. Sztuczne sieci neuronowe MLP

Sieć perceptronu wielowarstwowego MLP, algorytmy uczące MLP, program uczenia sieci MLP. Problem generalizacji sieci neuronowych.

## 6. Sztuczne sieci neuronowe RBF

Sieć RBF: struktura sieci i algorytmy uczące, program komputerowy do uczenia sieci RBF. Problemy praktyczne wykorzystania sieci.

## 7. Sieć wektorów nośnych SVM

Sieć do klasyfikacji, algorytmy uczenia sieci, sieć SVM do zadania regresji, sieć jednoklasowa SVM.

## 8. Zespoły klasyfikatorów i predyktorów

Struktura zespołu, metody integracji zespołu klasyfikatorów i predyktorów. Przykłady integracji.

## 9. Metody oceny jakości klasyfikatorów i predyktorów

Miary jakości predykcji i klasyfikacji, macierze rozkładu klas, miary wrażliwości i specyficzności, krzywa ROC, metody poprawy jakości klasyfikatorów, różnicowanie kosztów klasyfikacji, metoda równoważenia klas.

## 10. Transformacja PCA i ICA

Zastosowanie transformacji w kompresji, generacji cech diagnostycznych i wizualizacji danych wielowymiarowych. Toolbox ICALAB. Rzutowania nieliniowe danych wielowymiarowych. Zastosowania.

## 11. Wybrane metody generacji i selekcji cech diagnostycznych

Metody generacji cech dla sygnałów i obrazów. Metody selekcji cech: metoda Fishera, korelacji danych wejściowych z klasą, zastosowanie sieci liniowej SVM, zastosowanie hipotez statystycznych. Zastosowanie metod w klasyfikacji i predykcji.

## 12. Algorytmy grupowania danych

Algorytm K-means i rozmyty c-means. Algorytmy WTM. Ocena jakości grupowania danych. Zastosowania metod grupowania.

## 13. Zależności asocjacyjne między danymi

Podstawowe definicje asocjacji. Generacja zbiorów najczęściej pojawiających się w transakcjach. Generacja reguła asocjacyjnych. Metody oceny reguł asocjacyjnych.

## 14. Systemy rozmyte

Podstawy matematyczne systemów rozmytych. Wnioskowanie rozmyte Mamdaniego-Zadeha i Takagi-Sugeno-Kanga. Metody uczenia systemów rozmytych. Sieci neuro-fuzzy.

## 15. Wizualizacja danych wielowymiarowych

Metody wizualizacji danych wielowymiarowych. Transformacje liniowe i nieliniowe danych. Przykłady zastosowań narzędzi eksploracji danych.

Literatura: Literatura podstawowa:

1. S. Osowski, Metody i narzędzia eksploracji danych, BTC, Warszawa, 2012
2. S. Osowski, Sieci neuronowe do przetwarzania informacji, Oficyna Wydawnicza PW, Warszawa, 2006

Literatura uzupełniająca:

1. P. N. Tan., M. Steinbach, V. Kumar V., Introduction to data mining, Pearson Education Inc., Boston, 2006

Efekty uczenia: KAUE\_W1 / Doktorant ma zaawansowaną wiedzę z matematyki, niezbędną do:

- 1) modelowania i analizy zaawansowanych urządzeń i systemów elektronicznych, a także zjawisk fizycznych w nich występujących,
- 2) opisu i analizy działania oraz syntezy złożonych systemów elektronicznych,
- 3) opisu, analizy i syntezy algorytmów przetwarzania sygnałów i informacji / EiT\_W01

KAUE\_W2 / Doktorant ma zaawansowaną wiedzę w zakresie metod sztucznej inteligencji oraz eksploracji danych / EiT\_W07

KAUE\_U1 / Doktorant ma zaawansowaną wiedzę w zakresie metod sztucznej inteligencji oraz eksploracji danych / EiT\_U02

KAUE\_U2 / Doktorant potrafi dokonać analizy i syntezy złożonych

sygnałów i systemów przetwarzania sygnałów, stosując techniki analogowe i cyfrowe oraz odpowiednie narzędzia / EiT\_U05

KAUE\_K1 / Doktorant rozumie potrzebę uczenia się przez całe życie; potrafi inspirować i organizować proces uczenia się innych osób / EiT\_K01

KAUE\_K2 / Doktorant potrafi odpowiednio określić priorytety służące realizacji określonego przez siebie lub innych zadania / EiT\_K04

Kryteria Warunkiem zaliczenia przedmiotu jest uzyskanie oceny pozytywnej z oceniania: egzaminu pisemnego i ustnego.

Efekty: KAUE\_W1, KAUE\_W2, KAUE\_U1 i KAUE\_U2 sprawdzane są poprzez egzamin sprawdzający organizowany pod koniec semestru oraz poprzez dyskusje na wykładach;

Efekty KAUE\_K1 i KAUE\_K2 weryfikowane są częściowo poprzez dyskusje z doktorantami nad różnymi aspektami problemów poruszanych na wykładzie.



DYREKTOR  
Instytutu Systemów Elektronicznych  
Wydziału Elektroniki WAT



dr hab. inż. Tadeusz DĄBROWSKI