**Nazwa przedmiotu:**

Advanced Machine Learning

**Koordynator przedmiotu:**

Prof. dr hab. Jan Mielniczuk

**Status przedmiotu:**

Obowiązkowy

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Data Science

**Grupa przedmiotów:**

Współny

**Kod przedmiotu:**

.

**Semestr nominalny:**

2 / rok ak. 2022/2023

**Liczba punktów ECTS:**

6

**Liczba godzin pracy studenta związanych z osiągnięciem efektów uczenia się:**

1. godziny kontaktowe – 83 h; w tym
a) obecność na wykładach – 30 h
b) obecność na laboratoriach – 30 h
c) obecność na zajęciach projektowych – 15 h
d) konsultacje – 5 h |
e) obecność na egzaminie – 3 h
2. praca własna studenta – 70 h; w tym
a) zapoznanie się z literaturą – 5 h
b) przygotowanie projektu – 25 h
c) rozwiązanie zadań domowych – 10 h
d) przygotowanie do zajęć projektowych – 10 h
e) przygotowanie prezentacji projektu – 5 h
f) przygotowanie do egzaminu – 10 h
g) przygotowanie do zajęć laboratoryjnych –5 h
Razem 153 h, co odpowiada 6 pkt. ECTS

**Liczba punktów ECTS na zajęciach wymagających bezpośredniego udziału nauczycieli akademickich:**

1. obecność na wykładach – 30 h
2. obecność na laboratoriach – 30 h
3. obecność na zajęciach projektowych – 15 h
4. konsultacje – 5 h
5. obecność na egzaminie – 3 h
Razem 83 h, co odpowiada 3pkt. ECTS

**Język prowadzenia zajęć:**

angielski

**Liczba punktów ECTS, którą student uzyskuje w ramach zajęć o charakterze praktycznym:**

1. obecność na laboratoriach – 30 h
2. obecność na zajęciach projektowych – 15 h
3. przygotowanie projektu – 25 h
4. rozwiązanie zadań domowych – 10 h
5. przygotowanie do zajęć laboratoryjnych –5 h
6. przygotowanie do zajęć projektowych – 10 h
7. przygotowanie prezentacji projektu – 5h
Razem 100 h, co odpowiada 4 pkt. ECTS

**Formy zajęć i ich wymiar w semestrze:**

|  |  |
| --- | --- |
| Wykład:  | 30h |
| Ćwiczenia:  | 0h |
| Laboratorium:  | 30h |
| Projekt:  | 15h |
| Lekcje komputerowe:  | 0h |

**Wymagania wstępne:**

Introduction to Machine Learning

**Limit liczby studentów:**

Number of groups: no limits Project – the number of students in a group matches the limits defined by the Warsaw University of Technology Laboratory – the number of students in a group matches the lim

**Cel przedmiotu:**

Course objective: The aim of the course is to acquiant listeners with advanced machine learning techniques and to teach them, via hands-on approach, how to model and analyse regression data with the aid the state-of-the art statistical and data mining techniques,, including in particular modeling methods for regression curves, classification methods and analysis of time series.

**Treści kształcenia:**

Lecture:
1. Linear regression; Least Squares Method, regression diagnostics nd tests
2. Linear regression: departures from the linear model and its adaptation
3. Bayesian approach to linear modelling: Bayesian model comparison
4. Regularisation in regression: ridge regression and Lasso
5. High-dimensional regression, variable selection, filters, wrappers, criterion methods, greedy solutions to variable selection
6. Generalized linear models: Poisson and negative binomial responses
7. Loglinear models, contingency tables
8. Nonlinear regression: nonlinear parametric methods, nonparametric methods, regression tress CART, MARS method
9. Linear classification methods: empirical Bayes rules, density estimation
10. Nonlinear classification methods: QDA, classification TREES CART, kernel SVM. Multilabel classification.
11. Committees of classifiers: bagging, boosting, gradient algorithms, random forests.
12. Graphical models: Bayesian networks
13. Graphical models: random Markov fields, inference for graphical models
14. Practical construction of learning system for high-dimensional data: Multisplit, Random Subspace Methods, Shrunken Centroids
15. Time series: time series characteristics, forecasting, Yule-Walker method, innovation algorithm
16. Linear processes: ARMA(p,q) modelling
17. Functional Data Analysis
18. General kernel method
19. Mixture models and EM algorithm
Laboratory: Analysis and modelling of data using classification, regression and time series analytical methods
Project classes: Practical and multipurpose analysis of given data sets and construction of automatic modelling system meeting specific requirements ( bound on mean squared error of a forecast, LIFT criterion on test data)

**Metody oceny:**

50% based on project grading, 50%: exam performance

**Egzamin:**

tak

**Literatura:**

1..Bishop, Pattern Recognition and Machine Learning, Springer 2006
2. Hastie, Tibshirani, Friedman, Elements of Statistical Learning, wydanie drugie, Springer 2009
3. Izenman, Modern Multivariate Statistical Techniques, Springer 2008

**Witryna www przedmiotu:**

.

**Uwagi:**

.

## Charakterystyki przedmiotowe

### Profil ogólnoakademicki - wiedza

**Charakterystyka W01:**

He/She knows how to fit linear model using Least Squares regression, Lasso, and Ridge method,knows main types of model inadequacy and remedial methods. He/She has knowledge on variable selection, also in high-dimensional case.

Weryfikacja:

Exam, grading of the project

**Powiązane charakterystyki kierunkowe:** DS2\_W02, DS2\_W04

**Powiązane charakterystyki obszarowe:**

**Charakterystyka W02:**

He/She knows nonlinear regression estimation methods and when they are useful. Knows how to assess goodness of fit for regression and adequacy of classifier. Knows how to construct commmittess of classifiers and how to use them to rank predictors. Knows basic characteristics of statationary processes.

Weryfikacja:

Exam, grading of the projec

**Powiązane charakterystyki kierunkowe:** DS2\_W03

**Powiązane charakterystyki obszarowe:**

### Profil ogólnoakademicki - umiejętności

**Charakterystyka U01:**

He/She is able to assess goodness of fit of a linear model, to identify possible departures from the linear model and modify the linear model

Weryfikacja:

Exam, grading of the project

**Powiązane charakterystyki kierunkowe:** DS2\_U14, DS2\_U03, DS2\_U08

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U02:**

He/She is able to construct nonlinear classifier and assess its performance

Weryfikacja:

Exam, grading of the project

**Powiązane charakterystyki kierunkowe:** DS2\_U04, DS2\_U08, DS2\_U14

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U03:**

He/She is able to construct nonlinear regression estimator and assess its performance

Weryfikacja:

 Exam, grading of the project

**Powiązane charakterystyki kierunkowe:** DS2\_U03, DS2\_U08, DS2\_U14

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U04:**

He/She is able to construct a forecast based on linear/nonlinear time series models

Weryfikacja:

Exam, grading of the project

**Powiązane charakterystyki kierunkowe:** DS2\_U03, DS2\_U14

**Powiązane charakterystyki obszarowe:**

### Profil ogólnoakademicki - kompetencje społeczne

**Charakterystyka K01:**

He/She understands consequences of data mining based conclusions, in particular of reporting false discovery signals

Weryfikacja:

Grading of the project

**Powiązane charakterystyki kierunkowe:** DS2\_K02, DS2\_K03

**Powiązane charakterystyki obszarowe:**