



Elements of statistical machine learning

Course description sheet

Basic information

Field of study Computer Science Major - Organisational unit Faculty of Computer Science Study level First-cycle (engineer) programme Form of study Full-time studies Profile General academic		Didactic cycle 2022/2023 Course code WIINF5.II8.08694.22 Lecture languages Polish Mandatoriness Obligatory Block General Modules Course related to scientific research Yes
Course coordinator	Marcin Kurdziel	
Lecturer	Marcin Kurdziel, Rafał Grzeszczuk	
Period Semester 4	Method of verification of the learning outcomes Completing the classes Activities and hours Lectures: 14 Laboratory classes: 14	Number of ECTS credits 3

Course's learning outcomes

Code	Outcomes in terms of	Learning outcomes prescribed to a field of study	Methods of verification
Knowledge - Student knows and understands:			

Code	Outcomes in terms of	Learning outcomes prescribed to a field of study	Methods of verification
W1	Student ma wiedzę z zakresu wnioskowania Bayesowskiego oraz wybranych metod estymacji parametrów modeli uczących się.	INF1A_W07	Activity during classes, Oral answer, Completion of laboratory classes
W2	Student ma wiedzę z zakresu wybranych modeli uczenia maszynowego dla problemów regresji i klasyfikacji danych.	INF1A_W07	Activity during classes, Oral answer, Completion of laboratory classes
Skills - Student can:			
U1	Student potrafi zaimplementować wybrane modele statystycznego uczenia maszynowego.	INF1A_U05, INF1A_U07	Execution of laboratory classes, Completion of laboratory classes
U2	Student zna wybrane biblioteki przydatne w implementacji metod statystycznego uczenia maszynowego.	INF1A_U05, INF1A_U07	Execution of laboratory classes, Completion of laboratory classes
U3	Student potrafi wytrenować wybrane modele statystycznego uczenia maszynowego i ocenić ich skuteczność.	INF1A_U05, INF1A_U07	Execution of laboratory classes, Completion of laboratory classes

Student workload

Activity form	Average amount of hours* needed to complete each activity form
Lectures	14
Laboratory classes	14
Preparation for classes	25
Realization of independently performed tasks	25
Contact hours	5
Student workload	Hours 83
Workload involving teacher	Hours 28

* hour means 45 minutes

Program content

No.	Program content	Course's learning outcomes	Activities
1.	1. Częstościowa i Bayesowska interpretacja prawdopodobieństwa. Twierdzenie Bayesa. Wiarygodność, rozkład a priori i rozkład a posteriori. Wnioskowanie Bayesowskie. Pojęcie rozkładu sprzężonego. Przykład: model Beta-dwumianowy. (2h) 2. Wielowymiarowy rozkład normalny i jego własności. Parametryzacja: wartość przeciętna i macierz kowariancji. Rozkład łączny, brzegowy i warunkowy zmiennych normalnych. Liniowe modele Gaussowskie (2h) 3. Modele liniowe dla problemu regresji. Regresja liniowa w ujęciu klasycznym (MLE). Bayesowska regresja liniowa. (2h) 4. Wprowadzenie do procesów Gaussowskich. Regresja procesem Gaussowskim. (2h) 5. Modele liniowe dla problemu klasyfikacji. Regresja logistyczna i wieloklasowa regresja logistyczna. Estymator MLE dla regresji logistycznej. Stochastyczny spadek wzdłuż gradientu. (2h) 6. Metody Monte Carlo we wnioskowaniu Bayesowskim. Algorytmy generowania próbek. Metody MC: importance sampling. Estymacja parametrów metodą Monte Carlo. (2h) 7. Metody MCMC: Metropolis-Hastings. Hierarchiczne modele Bayesowskie. Model generujący dla obserwacji. Programowanie probabilistyczne. (2h)	W1, W2	Lectures

No.	Program content	Course's learning outcomes	Activities
2.	1. Wprowadzenie do pakietów obliczeniowych w języku Python. (2h) 2. Wielowymiarowy rozkład normalny i liniowe modele Gaussowskie. (2h) - Liniowy model Gaussowski dla zagadnienia estymacji położenia obiektu. 3. Regresji liniowa. (2h) - Implementacja klasycznej regresji liniowej z estymatorem MLE. - Implementacja Bayesowskiej regresji liniowej. 4. Procesy Gaussowskie. (2h) - Implementacja regresji procesem Gaussowskim. 5. Regresja logistyczna. (2h) - Implementacja wieloklasowej regresji logistycznej dla zdjęć ręcznie pisanych cyfr. 6. Wnioskowanie metodami Monte Carlo. (2h) - Wprowadzenie do pakietu TensorFlow Probability. - Implementacja algorytmu Importance Sampling. 7. Programowanie probabilistyczne w TensorFlow Probability. (2h) - Opracowanie modelu generującego dla wskazanego zbioru obserwacji. - Estymacja parametrów opracowanego modelu i wnioskowanie o postawionych hipotezach dotyczących obserwacji.	U1, U2, U3	Laboratory classes

Extended information/Additional elements

Teaching methods and techniques :

Lectures

Activities	Methods of verification	Credit conditions
Lectures	Activity during classes, Oral answer, Completion of laboratory classes	Zaliczenie przedmiotu wymaga uzyskania oceny pozytywnej z ćwiczeń laboratoryjnych.
Lab. classes	Execution of laboratory classes, Completion of laboratory classes	Zaliczenie przedmiotu wymaga uzyskania oceny pozytywnej z ćwiczeń laboratoryjnych.

Rules of participation in given classes, indicating whether student presence at the lecture is obligatory

Lectures: Studenci uczestniczą w zajęciach poznając kolejne treści nauczania zgodnie z sylabusem przedmiotu. Studenci winni na bieżąco zadawać pytania i wyjaśniać wątpliwości. Rejestracja audiowizualna wykładu wymaga zgody prowadzącego. Laboratory classes: Studenci wykonują ćwiczenia laboratoryjne zgodnie z materiałami udostępnionymi przez prowadzącego. Student jest zobowiązany do przygotowania się w przedmiocie wykonywanego ćwiczenia, co może zostać zweryfikowane kolokwium w formie ustnej lub pisemnej. Zaliczenie zajęć odbywa się na podstawie zaprezentowania rozwiązania postawionego problemu.

Literature

Obligatory

1. Kevin P. Murphy "Machine Learning: A probabilistic perspective", MIT Press, 2012

Optional

1. Christopher M. Bishop "Pattern Recognition and Machine Learning", Springer, 2007

Scientific research and publications

Publications

1. M. Jamroz, M. Kurdziel, M. Opala. "A Bayesian nonparametrics view into deep representations." Advances in Neural Information Processing Systems 33 (2020): 1440-1450.
2. 1. K. Grzegorzcyk, M. Kurdziel "Disambiguated Skip-gram model", In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP2018, pp 1445-1454, 2018
3. 2. P.I. Wójcik, M. Kurdziel "Training neural networks on high-dimensional data using random projection", Pattern Analysis and Applications, <https://doi.org/10.1007/s10044-018-0697-0>, 2018
4. 3. K. Grzegorzcyk, M. Kurdziel "Binary Paragraph Vectors", In: Proceedings of the 2nd Workshop on Representation Learning for NLP, ACL2017, pp 121-130, 2017

Learning outcomes prescribed to a field of study

Code	Content
INF1A_U05	Potrafi wykorzystać poznane metody i modele do tworzenia programów o charakterze użytkowym, a także potrafi adekwatnie wykorzystać znane algorytmy i struktury danych w budowie systemu komputerowego.
INF1A_U07	Potrafi ocenić, dobrać i stosować właściwe metody i narzędzia stosowane przy realizacji zadań związanych z budową systemu komputerowego, potrafi ocenić przydatność i korzystać z dostępnych bibliotek oraz komponentów oprogramowania.
INF1A_W07	Ma podstawową wiedzę z zakresu technik i zastosowań inżynierskich metod obliczeniowych oraz zagadnień sztucznej inteligencji