



KARTA OPISU PRZEDMIOTU - SYLABUS

Nazwa przedmiotu

Etyka i badania naukowe [S1S1E>EBN]

Przedmiot

Kierunek studiów

Sztuczna inteligencja/Artificial Intelligence

Rok/Semestr

3/6

Studia w zakresie (specjalność)

–

Profil studiów

ogólnoakademicki

Poziom studiów

pierwszego stopnia

Język oferowanego przedmiotu

angielski

Forma studiów

stacjonarne

Wymagalność

obligatoryjny

Liczba godzin

Wykład

15

Laboratorium

0

Inne

0

Ćwiczenia

15

Projekty/seminaria

0

Liczba punktów ECTS

2,00

Koordynatorzy

prof. dr hab. inż. Jerzy Stefanowski
jerzy.stefanowski@put.poznan.pl

Wykładowcy

Wymagania wstępne

Wcześniejsze ukończenie przedmiotów Sztuczna Inteligencja, Uczenia Maszynowe, Eksploracji danych, Podstawy Robotyki, Sieci Neuronowe wraz z uczeniem głębokim. Podstawowe umiejętności programowania w językach Python lub Java. Dobra znajomość języka angielskiego - czytanie literatury ze zrozumieniem i umiejętności komunikowania w formie pisanej oraz prezentacji na seminarium.

Cel przedmiotu

Celem przedmiotu jest zapoznanie studentów z aspektami etycznymi związanymi z rozwojem Sztucznej Inteligencji oraz podstawowych zasad realizacji projektów badawczych. W szczególności prezentuje się możliwe zagrożenia i ryzyka związane z niekontrolowanym rozwojem zaawansowanych podejść sztucznej inteligencji; koncepcje i postulaty tzw. sztucznej inteligencji godnej zaufania; problematykę obciążeń i trudności danych dla systemów uczących, bezstronności podejmowania decyzji wobec ludzi oraz podstaw metod wyjaśniania systemów inteligentnych typu "black box". Ponadto studenci poznają podstawowe pojęcia związane z rodzajem projektów badawczych, charakterystykę wiedzy naukowej, realizację procesu badań oraz aspektów etycznych z nim związanych.

Przedmiotowe efekty uczenia się

Wiedza:

K1st_W4: zna i rozumie podstawowe techniki, metody, algorytmy oraz narzędzia wykorzystywane w procesie rozwiązywania zadań informatycznych oraz problemów sztucznej inteligencji, w tym m.in. do grupowania, klasyfikacji, optymalizacji oraz wspomagania decyzji

K1st_W9: ma wiedzę nt. bezpieczeństwa, zagrożeń oraz zagadnień etycznych związanych z tworzeniem oraz wykorzystaniem systemów informatycznych ze szczególnym uwzględnieniem sztucznej inteligencji

Umiejętności:

K1st_U1: potrafi pozyskiwać informacje z właściwie dobranych źródeł o różnej charakterystyce, dokonywać ich krytycznej analizy, interpretacji i syntezy oraz wyczerpująco uzasadniać formułowane opinie

K1st_U4: potrafi zaplanować i przeprowadzić eksperymenty, w tym pomiary i symulacje komputerowe, interpretować uzyskane wyniki oraz wyciągać z nich wnioski w kontekście problemów eksploracji danych, uczenia maszynowego oraz problemów decyzyjnych wymagających znalezienia optymalnego rozwiązania lub podzbioru najbardziej preferowanych rozwiązań

K1st_U10: potrafi pozyskiwać, analizować i przetwarzać dane różnego typu, dokonywać ich syntezy do wiedzy i wniosków przydatnych do rozwiązywania szerokiego spektrum problemów decyzyjnych

K1st_U16: potrafi planować i realizować własne uczenie się przez całe życie oraz zna możliwości dalszego dokształcania (studia II stopnia)

Kompetencje społeczne:

K1st_K2: ma świadomość istotności wiedzy i badań naukowych związanych z informatyką i sztuczną inteligencją w rozwiązywaniu praktycznych problemów o kluczowym znaczeniu dla funkcjonowania jednostek, firm, organizacji oraz całego społeczeństwa w takich przykładowych obszarach zastosowań jak transport, ochrona zdrowia, edukacja, bezpieczeństwo publiczne czy rozrywka

K1st_K3: zna przykłady wadliwie działających systemów sztucznej inteligencji, które doprowadziły do strat ekonomicznych, społecznych lub środowiskowych

K1st_K7: jest gotowy odpowiedzialnie pełnić funkcje zawodowe, kultywować i upowszechniać w prowadzonej działalności wzory właściwego postępowania oraz prawidłowo identyfikować i rozstrzygać dylematy związane z wykonywaniem zawodu informatyka - specjalisty z zakresu sztucznej inteligencji

Metody weryfikacji efektów uczenia się i kryteria oceny

Efekty uczenia się przedstawione wyżej weryfikowane są w następujący sposób:

Wykłady: zaliczenie pisemne przeprowadzone na ostatnim wykładzie. Studenci muszą rozwiązać test z pytaniami zamkniętymi oraz otwartymi. Ponadto premiuje się aktywność - dyskusje w trakcie zajęć.

Ćwiczenia seminaryjne: Studenci wykonują zadania domowe, ponadto przygotowują jedną prezentację w trakcie zajęć, która dotyczy raportu z pracy z wybranymi tekstami. Powinni także uczestniczyć w dyskusjach prowadzonych w trakcie zajęć. Ostateczna ocena jest obliczana jako średnia z ocen częściowych przyznanych za poszczególne elementy.

Treści programowe

Wprowadzenie do podstawowych pojęć z zakresu etyki. Przedstawienie obecnie identyfikowanych zagrożeń związanych z zastosowaniami metod Sztucznej Inteligencji (SI), w tym w szczególności zaawansowanych systemów uczących się, robotyki oraz automatycznego podejmowania decyzji.

Propozycje tzw. odpowiedzialnej sztucznej inteligencji oraz postulaty inteligencji godnej zaufania (tzw. Trustworthy AI). Analiza rekomendacji z zaleceń etycznych wypracowanych przez grupę ekspertów wysokiego szczebla ds. sztucznej inteligencji przy KE. Trudności ich praktycznego wdrożenia.

Analiza jakości i trudności rozkładu przykładów uczących dla systemów SI. Możliwe źródła niedoskonałości danych. Pojęcie trudności i złożoności. Miary ich eksperymentalnej oceny.

Niezbalansowanie rozkładów w danych i metody radzenia sobie z nimi.

Bezstronność w systemach sztucznej inteligencji (Fairness-aware ML). Źródła i różnorodne obciążenia danych uczących. Miary. Wybrane metody poprawy bezstronności.

Potrzeba wyjaśniania predykcji systemów uczących się typu black box. Przykłady metod.

Cele prowadzenia badań naukowych. Rodzaje projektów badawczych. Wiedza naukowa i kryteria jej oceny. Formułowanie problemu badawczego, hipotez i celów badawczych. Przebieg standardowego procesu badawczego.

Odtwarzalność metod SI i wyników jej oceny eksperymentalnej.

Aspekty etyczne związane z prowadzeniem badań

Tematyka zajęć

Wprowadzenie do podstawowych pojęć z zakresu etyki. Przedstawienie obecnie identyfikowanych zagrożeń związanych z zastosowaniami metod Sztucznej Inteligencji (SI), w tym w szczególności zaawansowanych systemów uczących się, robotyki oraz automatycznego podejmowania decyzji. Propozycje tzw. odpowiedzialnej sztucznej inteligencji oraz postulaty inteligencji godnej zaufania (tzw. Trustworthy AI). Analiza rekomendacji z zaleceń etycznych wypracowanych przez grupę ekspertów wysokiego szczebla ds. sztucznej inteligencji przy KE. Trudności ich praktycznego wdrożenia. Analiza jakości i trudności rozkładu przykładów uczących dla systemów SI. Możliwe źródła niedoskonałości danych. Pojęcie trudności i złożoności. Miary ich eksperymentalnej oceny. Niezbalansowanie rozkładów w danych i metody radzenia sobie z nimi. Bezstronność w systemach sztucznej inteligencji (Fairness-aware ML). Źródła i różnorodne obciążenia danych uczących. Miary. Wybrane metody poprawy bezstronności. Potrzeba wyjaśniania predykcji systemów uczących się typu black box. Przykłady metod. Cele prowadzenia badań naukowych. Rodzaje projektów badawczych. Wiedza naukowa i kryteria jej oceny. Formułowanie problemu badawczego, hipotez i celów badawczych. Przebieg standardowego procesu badawczego. Odtwarzalność metod SI i wyników jej oceny eksperymentalnej. Aspekty etyczne związane z prowadzeniem badań

Metody dydaktyczne

Wykład: slajdy multimedialne dotyczące różnych aspektów wykładów, ilustrowane przykładami.
Ćwiczenia: praca z tekstami naukowymi, dyskusja wybranych zagadnień, pisanie raportów-esejów oraz prezentacja ich wyników.

Literatura

Podstawowa:

S. Matthew Liao, Ethics of Artificial Intelligence and Robotics. Oxford Univ. Press 2020.
Jan Recker, Scientific Research in Information Systems: A Beginner's Guide. Springer 2021.
M.Ryan, V. Stahl: Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications, JICES, 2021.
K.Siau, W.Wang: Artificial Intelligence: Ethics of AI and Ethical AI. JDM 2020

Uzupełniająca:

B.Stahl.D.Schroeder, R.Rodrigues, Ethics of Artificial Intelligence. Springer 2022.
Wayne C. Booth, Gregory G. Colomb, Joseph M. Williams, Joseph Bizup, The craft of research. The University of Chicago Press, 2016.

Bilans nakładu pracy przeciętnego studenta

	Godzin	ECTS
Łączny nakład pracy	55	2,00
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	30	1,00
Praca własna studenta (studia literaturowe, przygotowanie do zajęć laboratoryjnych/ćwiczeń, przygotowanie do kolokwium/egzaminu, wykonanie projektu)	25	1,00