



Karta przedmiotu

Nazwa i kod przedmiotu	Optymalizacja struktur i obliczeń w sieciach neuronowych, PG_00064490						
Kierunek studiów	Informatyka						
Data rozpoczęcia studiów	luty 2025 r.		Rok akademicki realizacji przedmiotu		2025/2026		
Poziom kształcenia	II stopnia		Grupa zajęć		Grupa zajęć fakultatywnych Grupa zajęć specjalnościowych Grupa zajęć powiązanych z prowadzonymi badaniami naukowymi w dziedzinie nauki związanej z kierunkiem - profil ogólnoakademicki		
Forma studiów	stacjonarne		Sposób realizacji		na uczelni		
Rok studiów	2		Język wykładowy		polski		
Semestr studiów	3		Liczba punktów ECTS		3.0		
Profil kształcenia	ogólnoakademicki		Forma zaliczenia		egzamin		
Jednostka prowadząca	Wydział Politechniki Gdańskiej -> Wydział Elektroniki, Telekomunikacji i Informatyki -> Katedra Systemów Multimedialnych						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot		dr hab. inż. Piotr Szczuko				
	Prowadzący zajęcia z przedmiotu		dr hab. inż. Piotr Szczuko				
Formy zajęć	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	15.0	0.0	15.0	15.0	0.0	45
	W tym liczba godzin zajęć na odległość: 0.0						
	Adres kursu na platformie eNauczanie: https://enauczanie.pg.edu.pl/moodle/enrol/index.php?id=47727 Moodle ID: 47727 Optymalizacja struktur i obliczeń w sieciach neuronowych 2026 https://enauczanie.pg.edu.pl/moodle/course/view.php?id=47727						
Aktywność studenta i liczba godzin pracy	Aktywność studenta	Udział w zajęciach dydaktycznych, objętych planem studiów		Udział w konsultacjach		Praca własna studenta	RAZEM
	Liczba godzin pracy studenta	45		6.0		24.0	75
Cel przedmiotu	Celem przedmiotu jest zapoznanie studenta w sposób teoretyczny, praktyczny (laboratorium) i problemowy (projekt) z zagadnieniami optymalizacji modeli w uczeniu głębokim. Przedstawione zostają techniki optymalizacji, upraszczania, struktury modeli, optymalizacji i przyspieszania obliczeń. Omawiane są algorytmy i podejścia do efektywnego uczenia modeli, zapewniania wysokiej skuteczności i dokładności działania w praktycznych przypadkach, m.in. ograniczonej pamięci, zaszumienia danych.						

Efekty uczenia się przedmiotu	Efekt kierunkowy	Efekt z przedmiotu	Sposób weryfikacji i oceny efektu
	[K7_U07] potrafi wykorzystać zaawansowane metody wspomagania procesów i funkcji, specyficzne dla kierunków studiów	Student zrealizował praktycznie model uczenia maszynowego i zoptymalizował go w wybrany sposób - uzasadniony realizowanym zadaniem, dopasowany do specyfiki modelu, stosując funkcje wybranej biblioteki programistycznej.	[SU4] Ocena umiejętności korzystania z metod i narzędzi [SU1] Ocena realizacji zadania
	[K7_W10] zna i rozumie w pogłębionym stopniu podstawowe procesy zachodzące w cyklu życia urządzeń, obiektów i systemów technicznych oraz metody wspomagania procesów i funkcji, specyficzne dla kierunku studiów	Student zna i rozumie zaawansowane metody optymalizacji modeli głębokiego uczenia w cyklu życia systemów AI, od projektowania architektury, przez trening, po wdrożenie, w tym techniki kompresji, przyspieszania obliczeń oraz zapewniania skuteczności działania w warunkach ograniczonych zasobów i zaszumionych danych.	[SW1] Ocena wiedzy faktograficznej [SW3] Ocena wiedzy zawartej w opracowaniu tekstowym i projektowym
	[K7_W11] zna i rozumie w pogłębionym stopniu ogólne zasady tworzenia i rozwoju form indywidualnej przedsiębiorczości oraz ekonomiczne, prawne i inne uwarunkowania różnych rodzajów działań związanych z nadaną kwalifikacją, w tym zasady ochrony własności przemysłowej i prawa autorskiego	Student zna i rozumie uwarunkowania ekonomiczne, prawne oraz etyczne związane z wdrażaniem i komercjalizacją zoptymalizowanych modeli głębokiego uczenia, w tym zasady ochrony własności intelektualnej dotyczące algorytmów, architektur sieci neuronowych oraz oprogramowania	[SW3] Ocena wiedzy zawartej w opracowaniu tekstowym i projektowym [SW1] Ocena wiedzy faktograficznej
[K7_U12] potrafi w pogłębionym stopniu analizować działanie elementów, układów i systemów związanych z kierunkiem studiów oraz mierzyć ich parametry i badać charakterystyki techniczne, a także planować i przeprowadzać eksperymenty związane z kierunkiem studiów, w tym symulacje komputerowe, interpretować uzyskane wyniki i wyciągać wnioski	Student potrafi analizować działanie i mierzyć parametry zoptymalizowanych modeli głębokiego uczenia, w tym badać charakterystyki techniczne architektur sieci neuronowych pod kątem efektywności obliczeniowej i dokładności, a także planować i przeprowadzać eksperymenty, dotyczące technik kompresji, przyspieszania obliczeń oraz uczenia w warunkach ograniczonych zasobów i zaszumionych danych, oraz interpretować uzyskane wyniki i formułować wnioski.	[SU1] Ocena realizacji zadania [SU4] Ocena umiejętności korzystania z metod i narzędzi [SU3] Ocena umiejętności wykorzystania wiedzy uzyskanej w ramach przedmiotu	
Treści przedmiotu	<p>Treści przedmiotu - wykład Metody redukcji rozmiaru sieci i przyspieszania obliczeń. Kwantyzacja, upraszczanie modeli i destylacja wiedzy. Przerzedzanie danych i architektury modeli. Uczenie modeli na zaszumionych danych. Poszukiwanie architektury modeli NAS. Trening samo-nadzorowany, pre-trenowanie modeli. Oszacowanie niepewności w sieciach neuronowych (metody kalibracji sieci neuronowych, techniki dropout, grupowanie modeli, sieci Bayesowskie). Odporność modeli, metody przeciwstawne (adversarial). Modele hybrydowe, pozbawione wag, kapsułkowe.</p> <p>Treści przedmiotu - laboratoria Eksperymentalna weryfikacja możliwości redukcji rozmiaru sieci. Praktyczne zapoznanie z uczeniem modeli na zaszumionych danych. Symulowane i rzeczywiste błędy etykiet w znanych zbiorach. Wykorzystanie treningu samo-nadzorowanego i pre-trenowanie modeli, do zwiększenia zdolności generalizacji. Zapoznanie ze sposobami szacowania niepewności w sieciach neuronowych i kalibracji modeli.</p> <p>Treści przedmiotu - projekt Krytyczna analiza stanu wiedzy i wybór modelu i zbioru danych. Analiza stanu wiedzy pod kątem wyboru metody optymalizacji modeli. Przygotowanie środowiska badawczego. Eksperymenty optymalizacji modeli, dostosowania danych. Sformułowanie wniosków, dokumentacja i publiczna prezentacja projektu.</p>		
Wymagania wstępne i dodatkowe			
Sposoby i kryteria oceniania osiągniętych efektów uczenia się	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej
	Laboratorium	51.0%	35.0%
	Kolokwium	51.0%	35.0%
	Projekt	51.0%	30.0%

Zalecana lista lektur	Podstawowa lista lektur	<p>Torsten Hoefler, Dan Alistarh, Tal Ben-Nun, Nikoli Dryden, Alexandra Peste. (2021) Sparsity in Deep Learning: Pruning and growth for efficient inference and training in neural networks. [2102.00554] (arxiv.org)</p> <p>Yu Cheng, Duo Wang, Pan Zhou, and Tao Zhang. 2020. A Survey of Model Compression and Acceleration for Deep Neural Networks. (2020). arXiv:cs.LG/1710.09282</p> <p>Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. 2019. Neural Architecture Search: A Survey. (2019). arXiv:stat.ML/1808.05377</p> <p>Manish Gupta and Puneet Agrawal. 2020. Compression of Deep Learning Models for Text: A Survey. (2020). arXiv:cs.CL/2008.05221</p> <p>V. Sze, Y. Chen, T. Yang, and J. S. Emer. 2017. Efficient Processing of Deep Neural Networks: A Tutorial and Survey. Proc. IEEE 105, 12 (2017), 22952329. https://doi.org/10.1109/JPROC.2017.2761740</p>
	Uzupełniająca lista lektur	<p>Tensorflow model optimization (2022) https://www.tensorflow.org/model_optimization.</p> <p>Yi Tay, Mostafa Dehghani, Dara Bahri, and Donald Metzler. 2020. Efficient transformers: A survey. (2020). arXiv:cs.LG/2009.06732.</p>
	Adresy eZasobów	
Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania	<ul style="list-style-type: none"> • Opisz, skomentuj, uzasadnij przydatność jednej z wybranych metod poznanych w ramach przedmiotu: redukcji rozmiaru sieci i przyspieszania obliczeń, kwantyzacji, upraszczanie modeli i destylacji wiedzy. • Uzasadnij potrzebę przeredzania danych i architektury modeli i korzyści płynące z tych operacji. • Opisz w jaki sposób realizować można uczenie modeli na zaszumionych danych. • Omów jak działa trening samo-nadzorowany, pre-trenowanie modeli oraz jakie są korzyści tych etapów. • Opisz metody kalibracji sieci neuronowych, techniki dropout, grupowania modeli. • Omów jak szacować odporność modeli. • Podaj przykład modelu pozbawionego wag oraz zastosowania sieci kapsułkowych. 	
Zajęcia praktyczne w ramach przedmiotu	Nie dotyczy	

Dokument wygenerowany elektronicznie. Nie wymaga pieczęci ani podpisu.