



Karta przedmiotu

Nazwa i kod przedmiotu	Biometria i przetwarzanie mowy, PG_00063902						
Kierunek studiów	Informatyka						
Data rozpoczęcia studiów	luty 2027 r.		Rok akademicki realizacji przedmiotu		2026/2027		
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	1		Język wykładowy		polski		
Semestr studiów	1		Liczba punktów ECTS		2.0		
Profil kształcenia	ogólnoakademicki		Forma zaliczenia		zaliczenie		
Jednostka prowadząca	Wydziały Politechniki Gdańskiej -> Wydział Elektroniki, Telekomunikacji i Informatyki -> Katedra Inteligentnych Systemów Interaktywnych						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot		dr inż. Jerzy Dembski				
	Prowadzący zajęcia z przedmiotu		dr inż. Jerzy Dembski				
Formy zajęć	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	15.0	0.0	0.0	0.0	15.0	30
	W tym liczba godzin zajęć na odległość: 0.0						
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	30		4.0		16.0	50
Cel przedmiotu	Przedmiot dotyczy zagadnień biometrii i przetwarzania mowy. W zakresie biometrii celem jest zapoznanie studentów z metodami i urządzeniami używanymi w biometrii. W zakresie przetwarzania mowy celem jest zapoznanie studentów z cechami mowy oraz z metodami i z konkretnymi rozwiązaniami używanymi w rozpoznawaniu i w syntezie mowy.						

Efekty uczenia się przedmiotu	Efekt kierunkowy	Efekt z przedmiotu	Sposób weryfikacji i oceny efektu
	[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	potrafi w pogłębionym stopniu analizować, mierzyć parametry i badać charakterystyki techniczne systemów biometrii i przetwarzania mowy, a także planować i przeprowadzać eksperymenty związane z biometrią i przetwarzaniem mowy, interpretować uzyskane wyniki i wyciągać wnioski	[SU1] Ocena realizacji zadania
	[K7_U01] potrafi wykorzystywać posiadaną wiedzę matematyczną przy formułowaniu i rozwiązywaniu złożonych i nietypowych problemów związanych z kierunkiem studiów, poprzez: – właściwy dobór informacji źródłowych oraz dokonywanie ich krytycznej analizy, syntezy oraz twórczej interpretacji i prezentacji tych informacji, – zastosowanie właściwych metod i narzędzi	Student potrafi wykorzystać posiadaną wiedzę matematyczną przy rozwiązywaniu problemów związanych z biometrią i przetwarzaniem mowy.	[SU2] Ocena umiejętności analizy informacji
	[K7_W03] zna i rozumie w pogłębionym stopniu budowę i zasady działania komponentów i systemów związanych z kierunkiem studiów, w tym teorie, metody i złożone zależności między nimi oraz wybrane zagadnienia szczegółowe – właściwe dla programu kształcenia	Student zna teorię i metody dotyczące biometrii i przetwarzania mowy oraz zna zasady działania systemów używanych w tych dziedzinach.	[SW2] Ocena wiedzy zawartej w prezentacji [SW3] Ocena wiedzy zawartej w opracowaniu tekstowym i projektowym
[K7_W04] zna i rozumie w pogłębionym stopniu zasady, metody i techniki programowania oraz zasady tworzenia oprogramowania komputerów albo programowania urządzeń lub sterowników wykorzystujących mikroprocesory albo inne elementy lub układy programowalne, specyficznych dla kierunku studiów, a także organizację pracy systemów wykorzystujących komputery lub te urządzenia	Student zna i rozumie w zaawansowanym stopniu zasady, metody i techniki programowania oraz zasady tworzenia oprogramowania do biometrii i przetwarzania mowy.	[SW1] Ocena wiedzy faktograficznej [SW2] Ocena wiedzy zawartej w prezentacji	
Treści przedmiotu	Treści przedmiotu - wykład Biometria (wykład): 1. Podstawowe wiadomości 2. Miary 3. Cechy biometryczne i wykorzystujące je metody i systemy (częściowo na seminarium) Przetwarzanie mowy (wykład): 1. Cechy mowy (relacja z ortografią, fonetyka i fonologia) 2. Rozpoznawanie mowy 3. Synteza mowy 4. Standardy oznaczania mowy 5. Konkretny system przetwarzania mowy (seminarium)		
Wymagania wstępne i dodatkowe	Podstawy uczenia maszynowego, w szczególności sieci neuronowych.		
Sposoby i kryteria oceniania osiągniętych efektów uczenia się	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa ocena końcowej
	wykład (kolokwium)	50.0%	70.0%
	seminarium (ocena wystąpienia i przygotowania)	50.0%	30.0%

Zalecana lista lektur	Podstawowa lista lektur	
		<ol style="list-style-type: none"> 1. Marek Wiśniewski, Zarys fonetyki i fonologii współczesnego języka polskiego, Wydawnictwo Uniwersytetu Mikołaja Kopernika, wydanie IV, Toruń 2001. 2. Daniel Jurafsky, James H. Martin, Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, second edition, Prentice-Hall, 2008. 3. Bartosz Ziółko, Mariusz Ziółko, Przetwarzanie mowy, Wydawnictwa AGH, Kraków 2011. 4. Ruud M. Bolle, Jonathan H. Connell, Sharath Pankanti, Nalini K. Ratha, Andrew W. Senior, Biometria, Wydawnictwa Naukowo-Techniczne, Warszawa, 2008. 5. Krzysztof Ślot, Wybrane zagadnienia biometrii, Wydawnictwa Komunikacji i Łączności, Warszawa, 2008. 6. Krzysztof Ślot, Rozpoznawanie biometryczne. Nowe metody ilościowej reprezentacji obiektów, Wydawnictwa Komunikacji i Łączności, Warszawa, 2010. 7. Zygmunt Ciota, Metody przetwarzania sygnałów akustycznych w komputerowej analizie mowy, Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2010. 8. Mariusz Kubanek, Wybrane metody i systemy biometryczne bazujące na ukrytych modelach Markowa, Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2013. 9. Deep Learning in Biometrics, Mayank Vatsa, Richa Singh, Angshul Majumdar (eds.), CRC Press, 2018. 10. Yaniv Taigman, Ming Yang, MarcAurelio Ranzato, Lior Wolf, Deepface: Closing the Gap to Human-Level Performance in Face Verification, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1701-1708, 2014.

	Uzupełniająca lista lektur	<ol style="list-style-type: none"> 1. Danuta Ostaszewska, Jolanta Tambor, Fonetyka i fonologia współczesnego języka polskiego, PWN, Warszawa, 2002. 2. Alicja Nagórko, Podręczna gramatyka języka polskiego, PWN, Warszawa, 2010. 3. Zygmunt Ciota, Metody przetwarzania sygnałów akustycznych w komputerowej analizie mowy, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2010. 4. Ryszard Tadeusiewicz, Sygnał mowy, Wydawnictwa Komunikacji i Łączności, Warszawa 1988, książka dostępna pod adresem: http://winntbg.bg.agh.edu.pl/skrypty/0004/ 5. Tomasz Zieliński, Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań, Wydawnictwa Komunikacji i Łączności, 2009. 6. Grażyna Demenko, Korpusowe badania języka mówionego, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2015. 7. Xuedong Huang, Alex Acero, Hsiao-Wuen Hon, Spoken Language Processing: A Guide to Theory, Algorithm and System Development, Prentice Hall, 2001. 8. Paul Taylor, Text-to-speech synthesis, Cambridge University Press, 2009. 9. Stefan Breuer, Multifunktionale und Multilinguale Unit-Selection-Sprachsynthese. Designprinzipien für Architektur und Sprachbausteine, rozprawa doktorska, Rheinischen Friedrich-Wilhelms-Universität Bonn, 2009. Dostępna pod adresem: http://hss.ulb.uni-bonn.de/2009/1650/1650.pdf 10. Peter Birkholz, 3D-Artikulatorische Sprachsynthese, rozprawa doktorska, Universität Rostock, 2005. Dostępna pod adresem: http://www.vocaltractlab.de/publications/birkholz-2005-dissertation.pdf 11. Takashi Masuko, HMM-Based Speech Synthesis and Its Applications, rozprawa doktorska, Tokyo Institute of Technology, 2002. Dostępna pod adresem: http://www.kbys.ip.titech.ac.jp/masuko/masuko-doctor.pdf 12. Sercan Ö. Arik et al., Deep Voice: Real Time Neural Text-To-Speech, ICML 2017. Dostępny pod adresem: https://arxiv.org/abs/1702.07825 13. Sercan Ö. Arik et al., Deep Voice: Multi-Speaker Real Time Neural Text-To-Speech, NIPS 2017. Dostępny pod adresem: https://arxiv.org/abs/1705.08947 14. Wei Ping et al., Deep Voice 3: Scaling Text-To-Speech With Convolutional Sequence Learning, ICLR 2018. Dostępny pod adresem: https://arxiv.org/abs/1710.07654 15. Sercan Ö. Arik et al., Neural Voice Cloning with a Few Samples, NIPS 2018. Dostępny pod adresem: https://arxiv.org/abs/1802.06006 16. Aäron van den Oord et al., WaveNet: A Generative Model for Raw Audio, arXiv preprint, Dostępny pod adresem: https://arxiv.org/abs/1609.03499 17. Jose Sotelo et al., Char2wav: End-to-End Speech Synthesis, 2017. Dostępny pod adresem: https://openreview.net/pdf?id=B1VWyySKx 18. Alex Graves, Abdel-rahman Mohamed, Geoffrey Hinton, Speech Recognition with Recurrent Neural Networks. Dostępny pod adresem: http://www.cs.toronto.edu/~fritz/absps/RNN13.pdf 19. Dario Amodei et al., Deep Speech 2: End-to-End Speech Recognition in English and Mandarin, ICML 2016. Dostępny pod adresem: http://proceedings.mlr.press/v48/amodei16.pdf 20. Alex Graves, Navdeep Jaitly, Towards End-To-End Speech Recognition with Recurrent Neural Networks, ICML14, pp. 1764-1772. 2014. Dostępny pod adresem: http://proceedings.mlr.press/v32/graves14.pdf 21. Haşim Sak, Andrew Senior, Kanishka Rao, Françoise Beaufays, Fast and Accurate Neural Network Acoustic Models for Speech Recognition, Interspeech 2015. Dostępny pod adresem: https://arxiv.org/abs/1507.06947
	Adresy eZasobów	
Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania		
Zajęcia praktyczne w ramach przedmiotu	Nie dotyczy	