



Karta przedmiotu

Nazwa i kod przedmiotu	Przetwarzanie mowy, PG_00058857						
Kierunek studiów	Informatyka						
Data rozpoczęcia studiów	luty 2023 r.	Rok akademicki realizacji przedmiotu			2022/2023		
Poziom kształcenia	II stopnia	Grupa zajęć			Grupa zajęć fakultatywnych 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	Wydział Elektroniki, Telekomunikacji i Informatyki -> Katedra Inteligentnych Systemów Interaktywnych						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot	dr hab. inż. Jan Daciuk					
	Prowadzący zajęcia z przedmiotu	dr hab. inż. Jan Daciuk					
Formy zajęć i metody nauczania	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		2.0		18.0	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_U42] potrafi rozwiązywać problemy inżynierskie i badawcze w zakresie projektowania, oceny i utrzymania systemów i aplikacji informacyjnych z wykorzystaniem metod eksperymentalnych i technik zarządzania	Student potrafi rozwiązywać problemy inżynierskie i badawcze w zakresie projektowania, oceny i utrzymania systemów z dziedziny biometrii i przetwarzania mowy.	[SU3] Ocena umiejętności wykorzystania wiedzy uzyskanej w ramach przedmiotu
	[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_U06] potrafi analizować działanie elementów, układów i systemów związanych z kierunkiem studiów oraz mierzyć ich parametry i badać charakterystyki techniczne, interpretować uzyskane wyniki i wyciągać wnioski	Student potrafi mierzyć charakterystyki i analizować działanie systemów biometrycznych i systemów przetwarzania mowy.	[SU3] Ocena umiejętności wykorzystania wiedzy uzyskanej w ramach przedmiotu
	[K7_W04] zna i rozumie w zaawansowanym 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
	[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
Treści przedmiotu	<p>Biometria (wykład):</p> <ol style="list-style-type: none"> <li>1. Podstawowe wiadomości</li> <li>2. Miary</li> <li>3. Cechy biometryczne i wykorzystujące je metody i systemy (częściowo na seminarium)</li> </ol> <p>Przetwarzanie mowy (wykład):</p> <ol style="list-style-type: none"> <li>1. Cechy mowy (relacja z ortografią, fonetyka i fonologia)</li> <li>2. Rozpoznawanie mowy</li> <li>3. Synteza mowy</li> <li>4. Standardy oznaczania mowy</li> <li>5. Konkretny systemy przetwarzania mowy (seminarium)</li> </ol>		
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
	seminarium (ocena wystąpienia i przygotowania)	50.0%	30.0%
	wykład (kolokwium)	50.0%	70.0%

Zalecana lista lektur	Podstawowa lista lektur	
		<ol style="list-style-type: none"> <li>1. Marek Wiśniewski, Zarys fonetyki i fonologii współczesnego języka polskiego, Wydawnictwo Uniwersytetu Mikołaja Kopernika, wydanie IV, Toruń 2001.</li> <li>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.</li> <li>3. Bartosz Ziółko, Mariusz Ziółko, Przetwarzanie mowy, Wydawnictwa AGH, Kraków 2011.</li> <li>4. Ruud M. Bolle, Jonathan H. Connell, Sharath Pankanti, Nalini K. Ratha, Andrew W. Senior, Biometria, Wydawnictwa Naukowo-Techniczne, Warszawa, 2008.</li> <li>5. Krzysztof Ślot, Wybrane zagadnienia biometrii, Wydawnictwa Komunikacji i Łączności, Warszawa, 2008.</li> <li>6. Krzysztof Ślot, Rozpoznawanie biometryczne. Nowe metody ilościowej reprezentacji obiektów, Wydawnictwa Komunikacji i Łączności, Warszawa, 2010.</li> <li>7. Zygmunt Ciota, Metody przetwarzania sygnałów akustycznych w komputerowej analizie mowy, Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2010.</li> <li>8. Mariusz Kubanek, Wybrane metody i systemy biometryczne bazujące na ukrytych modelach Markowa, Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2013.</li> <li>9. Deep Learning in Biometrics, Mayank Vatsa, Richa Singh, Angshul Majumdar (eds.), CRC Press, 2018.</li> <li>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.</li> </ol>

	<p>Uzupełniająca lista lektur</p>	<ol style="list-style-type: none"> <li>1. Danuta Ostaszewska, Jolanta Tambor, Fonetyka i fonologia współczesnego języka polskiego, PWN, Warszawa, 2002.</li> <li>2. Alicja Nagórko, Podręczna gramatyka języka polskiego, PWN, Warszawa, 2010.</li> <li>3. Zygmunt Ciota, Metody przetwarzania sygnałów akustycznych w komputerowej analizie mowy, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2010.</li> <li>4. Ryszard Tadeusiewicz, Sygnał mowy, Wydawnictwa Komunikacji i Łączności, Warszawa 1988, książka dostępna pod adresem: <a href="http://winntbg.bg.agh.edu.pl/skrypty/0004/">http://winntbg.bg.agh.edu.pl/skrypty/0004/</a></li> <li>5. Tomasz Zieliński, Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań, Wydawnictwa Komunikacji i Łączności, 2009.</li> <li>6. Grażyna Demenko, Korpusowe badania języka mówionego, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2015.</li> <li>7. Xuedong Huang, Alex Acero, Hsiao-Wuen Hon, Spoken Language Processing: A Guide to Theory, Algorithm and System Development, Prentice Hall, 2001.</li> <li>8. Paul Taylor, Text-to-speech synthesis, Cambridge University Press, 2009.</li> <li>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: <a href="http://hss.ulb.uni-bonn.de/2009/1650/1650.pdf">http://hss.ulb.uni-bonn.de/2009/1650/1650.pdf</a></li> <li>10. Peter Birkholz, 3D-Artikulatorische Sprachsynthese, rozprawa doktorska, Universität Rostock, 2005. Dostępna pod adresem: <a href="http://www.vocaltractlab.de/publications/birkholz-2005-dissertation.pdf">http://www.vocaltractlab.de/publications/birkholz-2005-dissertation.pdf</a></li> <li>11. Takashi Masuko, HMM-Based Speech Synthesis and Its Applications, rozprawa doktorska, Tokyo Institute of Technology, 2002. Dostępna pod adresem: <a href="http://www.kbys.ip.titech.ac.jp/masuko/masuko-doctor.pdf">http://www.kbys.ip.titech.ac.jp/masuko/masuko-doctor.pdf</a></li> <li>12. Sercan Ö. Arik et al., Deep Voice: Real Time Neural Text-To-Speech, ICML 2017. Dostępny pod adresem: <a href="https://arxiv.org/abs/1702.07825">https://arxiv.org/abs/1702.07825</a></li> <li>13. Sercan Ö. Arik et al., Deep Voice: Multi-Speaker Real Time Neural Text-To-Speech, NIPS 2017. Dostępny pod adresem: <a href="https://arxiv.org/abs/1705.08947">https://arxiv.org/abs/1705.08947</a></li> <li>14. Wei Ping et al., Deep Voice 3: Scaling Text-To-Speech With Convolutional Sequence Learning, ICLR 2018. Dostępny pod adresem: <a href="https://arxiv.org/abs/1710.07654">https://arxiv.org/abs/1710.07654</a></li> <li>15. Sercan Ö. Arik et al., Neural Voice Cloning with a Few Samples, NIPS 2018. Dostępny pod adresem: <a href="https://arxiv.org/abs/1802.06006">https://arxiv.org/abs/1802.06006</a></li> <li>16. Aäron van den Oord et al., WaveNet: A Generative Model for Raw Audio, arXiv preprint, Dostępny pod adresem: <a href="https://arxiv.org/abs/1609.03499">https://arxiv.org/abs/1609.03499</a></li> <li>17. Jose Sotelo et al., Char2wav: End-to-End Speech Synthesis, 2017. Dostępny pod adresem: <a href="https://openreview.net/pdf?id=B1VWyySKx">https://openreview.net/pdf?id=B1VWyySKx</a></li> <li>18. Alex Graves, Abdel-rahman Mohamed, Geoffrey Hinton, Speech Recognition with Recurrent Neural Networks. Dostępny pod adresem: <a href="http://www.cs.toronto.edu/~fritz/absps/RNN13.pdf">http://www.cs.toronto.edu/~fritz/absps/RNN13.pdf</a></li> <li>19. Dario Amodei et al., Deep Speech 2: End-to-End Speech Recognition in English and Mandarin, ICML 2016. Dostępny pod adresem: <a href="http://proceedings.mlr.press/v48/amodei16.pdf">http://proceedings.mlr.press/v48/amodei16.pdf</a></li> <li>20. Alex Graves, Navdeep Jaitly, Towards End-To-End Speech Recognition with Recurrent Neural Networks, ICML14, pp. 1764-1772. 2014. Dostępny pod adresem: <a href="http://proceedings.mlr.press/v32/graves14.pdf">http://proceedings.mlr.press/v32/graves14.pdf</a></li> <li>21. Haşim Sak, Andrew Senior, Kanishka Rao, Françoise Beaufays, Fast and Accurate Neural Network Acoustic Models for Speech Recognition, Interspeech 2015. Dostępny pod adresem: <a href="https://arxiv.org/abs/1507.06947">https://arxiv.org/abs/1507.06947</a></li> </ol>
	<p>Adresy eZasobów</p>	<p>Adresy na platformie eNauczanie: Biometria i przetwarzanie mowy 2023 - Moodle ID: 28413 <a href="https://enauzanie.pg.edu.pl/moodle/course/view.php?id=28413">https://enauzanie.pg.edu.pl/moodle/course/view.php?id=28413</a></p>
<p>Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania</p>		
<p>Praktyki zawodowe w ramach przedmiotu</p>	<p>Nie dotyczy</p>	