GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

| | | | |
|---|---|---|---|
| Subject name and code | Biometry and Speech Processing, PG_00063902 | | |
| Field of study | Informatics | | |
| Date of commencement of studies | February 2027 | Academic year of realisation of subject | 2026/2027 |
| Education level | second-cycle studies | Subject group | Optional subject group<br>Specialty subject group<br>Subject group related to scientific research in the field of study |
| Mode of study | Full-time studies | Mode of delivery | at the university |
| Year of study | 1 | Language of instruction | Polish |
| Semester of study | 1 | ECTS credits | 2.0 |
| Learning profile | general academic profile | Assessment form | assessment |
| Conducting unit | Department of Intelligent Interactive Systems -> Faculty of Electronics Telecommunications and Informatics -> Faculties of Gdańsk University of Technology | | |
| Name and surname of lecturer (lecturers) | Subject supervisor | dr inż. Jerzy Dembski | |
| | Teachers | dr inż. Jerzy Dembski | |

| Lesson types | Lesson type | Lecture | Tutorial | Laboratory | Project | Seminar | SUM |
|---|---|---|---|---|---|---|---|
| | Number of study hours | 15.0 | 0.0 | 0.0 | 0.0 | 15.0 | 30 |
| | E-learning hours included: 0.0 | | | | | | |

| Learning activity and number of study hours | Learning activity | Participation in didactic classes included in study plan | Participation in consultation hours | Self-study | SUM |
|---|---|---|---|---|---|
| | Number of study hours | 30 | 4.0 | 16.0 | 50 |

| | |
|---|---|
| Subject objectives | The course concerns issues in biometrics and in speech processing. In biometrics, the aim is to familiarize students with methods and devices used in that domain. In speech processing, the aim is to get students aquainted with properties of speech, and to familiarize them with methods and particular solutions used in speech recognition and in speech synthesis. |

## Learning outcomes

| Course outcome | Subject outcome | Method of verification |
|---|---|---|
| [K7_U12] is able, to an increased extent, to analyze the operation of components and systems related to the field of study, as well as to measure their parameters and study their technical characteristics, and to plan and carry out experiments related to the field of study, including computer simulations, interpret the obtained results and draw conclusions | can deeply analyze, measure parameters, and examine technical characteristics of biometric and speech recognition systems, as well as to plan and conduct experiments in biometrics and speech processing, interpret the results and to draw conslusions | [SU1] Assessment of task fulfilment |
| [K7_U01] can apply mathematical knowledge to formulate and solve complex and non-typical problems related to the field of study by: - appropriate selection of source information and its critical analysis, synthesis, creative interpretation and presentation, - application of appropriate methods and tools | The student can use the acquired mathematical knowledge in solving biometrics and speech processing problems. | [SU2] Assessment of ability to analyse information |
| [K7_W03] knows and understands, to an increased extent, the construction and operating principles of components and systems related to the field of study, including theories, methods and complex relationships between them and selected specific issues - appropriate for the curriculum | The student knows the theory and methods in biometrics and speech processing as well as the foundations of functioning of the systems used in those domains. | [SW2] Assessment of knowledge contained in presentation<br>[SW3] Assessment of knowledge contained in written work and projects |
| [K7_W04] knows and understands, to an increased extent, the principles, methods and techniques of programming and the principles of computer software development or programming devices or controllers using microprocessors or other elements or programmable devices specific to the field of study, and organization of work of systems using computers or such devices | The student knows and understands in advanced degree the foundations, methods, and programming techniques as well as rules for developing software for biometrics and for speech processing. | [SW1] Assessment of factual knowledge<br>[SW2] Assessment of knowledge contained in presentation |

## Subject contents

Course content – lecture

Biometrics (lecture):

1. Foundations
2. Measures
3. Biometric features as well as methods and systems that use them (partially at a seminar)

Speech processing (lecture):

1. Speech features (relation with ortography, phonetics nad phonology)
2. Speech recognition
3. Speech synthesis
4. Standards for speech markup
5. Examples of speech processing systems (seminar)

## Prerequisites and co-requisites

Basic knowledge of machine learning, and neural networks ijn particular.

## Assessment methods and criteria

| Subject passing criteria | Passing threshold | Percentage of the final grade |
|---|---|---|
| lecture (exam) | 50.0% | 70.0% |
| seminar (evaluation of a presentation and its preparation) | 50.0% | 30.0% |

| | | |
|---|---|---|
| Recommended reading | Basic literature | 1. Marek Wiśniewski, Zarys fonetyki i fonologii współczesnego języka polskiego, Wydawnictwo Uniwersytetu Mikołaja Kopernika, wydanie IV, Toruń 2001.<br>2. Daniel Jurafsky, James H. Martin, Speech and Language Processing.<br>An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, second edition, Prentice-Hall, 2008.<br>3. Bartosz Ziółko, Mariusz Ziółko, Przetwarzanie mowy, Wydawnictwa AGH, Kraków 2011.<br>4. Ruud M. Bolle, Jonathan H. Connell, Sharath Pankanti, Nalini K. Ratha, Andrew W. Senior, Biometria, Wydawnictwa Naukowo-Techniczne, Warszawa, 2008.<br>5. Krzysztof Ślot, Wybrane zagadnienia biometrii, Wydawnictwa Komunikacji i Łączności, Warszawa, 2008.<br>6. Krzysztof Ślot, Rozpoznawanie biometryczne. Nowe metody ilościowej reprezentacji obiektów, Wydawnictwa Komunikacji i Łączności, Warszawa, 2010.<br>7. Zygmunt Ciota, Metody przetwarzania sygnałów akustycznych w komputerowej analizie mowy, Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2010.<br>8. Mariusz Kubanek, Wybrane metody i systemy biometryczne bazujące<br>na ukrytych modelach Markowa, Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2013.<br>9. Deep Learning in Biometrics, Mayank Vatsa, Richa Singh, Angshul Majumdar (eds.), CRC Press, 2018.<br>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. |

| | | |
|---|---|---|
| | Supplementary literature | 1. Danuta Ostaszewska, Jolanta Tambor, Fonetyka i fonologia współczesnego języka polskiego, PWN, Warszawa, 2002.<br>2. Alicja Nagórko, Podręczna gramatyka języka polskiego, PWN, Warszawa, 2010.<br>3. Zygmunt Ciota, Metody przetwarzania sygnałów akustycznych w komputerowej analizie mowy, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2010.<br>4. Ryszard Tadeusiewicz, Sygnał mowy, Wydawnictwa Komunikacji i Łączności, Warszawa 1988, książka dostępna pod adresem: http://winntbg.bg.agh.edu.pl/skrypty/0004/<br>5. Tomasz Zieliński, Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań, Wydawnictwa Komunikacji i Łączności, 2009.<br>6. Grażyna Demenko, Korpusowe badania języka mówionego, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2015.<br>7. Xuedong Huang, Alex Acero, Hsiao-Wuen Hon, Spoken Language Processing: A Guide to Theory, Algorithm and System Development,<br>Prentice Hall, 2001.<br>8. Paul Taylor, Text-to-speech synthesis, Cambridge University Press, 2009.<br>9. Stefan Breuer, Multifunktionale und Multilinguale Unit-Selection-Sprachsynthese. Designprinzipien für Architektur und<br>Sprachbausteine, rozprawa doktorska, Rheinischen Friedrich-Wilhelms-Universität Bonn, 2009. Dostępna pod adresem: http://hss.ulb.uni-bonn.de/2009/1650/1650.pdf<br>10. Peter Birkholz, 3D-Artikulatorische Sprachsynthese, rozprawa doktorska, Universität Rostock, 2005. Dostępna pod adresem: http://www.vocaltractlab.de/publications/birkholz-2005-dissertation.pdf<br>11. Takashi Masuko, HMM-Based Speech Synthesis and Its Applications,<br>rozprawa doktorska, Tokyo Institute of Technology, 2002. Dostępna pod adresem:<br>http://www.kbys.ip.titech.ac.jp/masuko/masuko-doctor.pdf<br>12. Sercan Ö. Arik et al., Deep Voice: Real Time Neural Text-To-Speech,<br>ICML 2017. Dostępny pod adresem:<br>https://arxiv.org/abs/1702.07825<br>13. Sercan Ö. Arik et al., Deep Voice: Multi-Speaker Real Time Neural Text-To-Speech, NIPS 2017. Dostępny pod adresem:<br>https://arxiv.org/abs/1705.08947<br>14. Wei Ping et al., Deep Voice 3: Scaling Text-To-Speech With Convolutional Sequence Learning, ICLR 2018. Dostępny pod adresem:<br>https://arxiv.org/abs/1710.07654<br>15. Sercan Ö. Arik et al., Neural Voice Cloning with a Few Samples, NIPS 2018. Dostępny pod adresem:<br>https://arxiv.org/abs/1802.06006<br>16. Aäron van den Oord et al., WaveNet: A Generative Model for Raw Audio, arXiv preprint, Dostępny pod adresem:<br>https://arxiv.org/abs/1609.03499<br>17. Jose Sotelo et al., Char2wav: End-to-End Speech Synthesis, 2017. Dostępny pod adresem:<br>https://openreview.net/pdf?id=B1VWyySKx<br>18. Alex Graves, Abdel-rahman Mohamed, Geoffrey Hinton, Speech Recognition with Recurrent Neural Networks. Dostępny pod adresem:<br>http://www.cs.toronto.edu/~fritz/absps/RNN13.pdf<br>19. Dario Amodei et al., Deep Speech 2: End-to-End Speech Recognition<br>in English and Mandarin, ICML 2016. Dostępny pod adresem:<br>http://proceedings.mlr.press/v48/amodei16.pdf<br>20. Alex Graves, Navdeep Jaitly, Towards End-To-End Speech Recognition with Recurrent Neural Networks, ICML14, pp. 1764-1772. 2014. Dostępny pod adresem:<br>http://proceedings.mlr.press/v32/graves14.pdf<br>21. Haşim Sak, Andrew Senior, Kanishka Rao, Françoise Beaufays, Fast<br>and Accurate Neural Network Acoustic Models for Speech Recognition, Interspeech 2015. Dostępny pod adresem:<br>https://arxiv.org/abs/1507.06947 |
| | eResources addresses | |
| Example issues/ example questions/ tasks being completed | | |
| Practical activites within the subject | Not applicable | |