GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	Deep Processing Text and Speech Signal, PG_00064489		
Field of study	Informatics		
Date of commencement of studies	February 2025	Academic year of realisation of subject	2025/2026
Education level	second-cycle studies	Subject group	Optional subject group Specialty subject group 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	2	ECTS credits	3.0
Learning profile	general academic profile	Assessment form	exam
Conducting unit	Department of Multimedia Systems -> Faculty of Electronics, Telecommunications and Informatics		
Name and surname of lecturer (lecturers)	Subject supervisor	dr hab. inż. Piotr Szczuko	
	Teachers	dr hab. inż. Piotr Szczuko dr inż. Arkadiusz Harasimiuk mgr inż. Szymon Zaporowski dr inż. Sebastian Cygert	

Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	15.0	0.0	30.0	0.0	0.0	45
	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	45	6.0	24.0	75

Subject objectives	The course is dedicated to important theoretical and practical aspects of text analysis and speech sound processing. The included topics are: modelling meaning and context, graph neural networks, attention in neural networks, GPT networks. For speech processing following will be presented: emission and perception of speech, registration and processing aimed at machine learning applications, AI methods for text and speech, analysis and recognition, feature extraction, applications of wavenet and tacotron architectures. Example tools and libraries in Python will be presented and used in practice

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	The student is able to critically analyze the results of training and application of machine learning methods dedicated to text and speech signal processing. The student is able to formulate conclusions from the work done, from the observation of the training process and the inference of selected deep models.	[SU1] Assessment of task fulfilment [SU2] Assessment of ability to analyse information [SU3] Assessment of ability to use knowledge gained from the subject [SU5] Assessment of ability to present the results of task
[K7_W02] knows and understands, to an increased extent, selected laws of physics and physical phenomena, as well as methods and theories explaining the complex relationships between them, constituting advanced general knowledge in the field of technical sciences related to the field of study	Student is able to describe phenomena related to speech emission and perception, chooses tools and methods for correct registration and processing of the signals aimed at machine learning applications.	[SW1] Assessment of factual knowledge
[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	Student is able to correctly choose tools for text and speech preprocessing; properly defines goals and research problems and applies data processing adequate to the defined aims.	[SU3] Assessment of ability to use knowledge gained from the subject [SU1] Assessment of task fulfilment

Subject contents

- Tools, best practices
- Digital signals, speech representation, parameterisation
- Fundamentals of speech acoustics, signal acquisition and processing, practical issues in ML
- NLP applications with GPT models
- Principles and applications of Wavenet architecture
- Deep learning-based voice synthesisers
- Speaker recognition
- Style transfer for speech
- Graph neural networks
- Boosted decision trees in speech processing
- Voice Biometrics

Prerequisites and co-requisites

Assessment methods and criteria

Subject passing criteria	Passing threshold	Percentage of the final grade
Laboratory reports	51.0%	50.0%
Colloquium	51.0%	50.0%

Recommended reading

Basic literature	Bengio Yoshua, Courville Aaron, Goodfellow Ian, Deep Learning, 2018 Lane Hobson, Cole Howard, Hannes Hapke, Natural Language Processing in Action: Understanding, Analyzing, and Generating Text with Python, Manning 2019
Supplementary literature	Teaching materials for sound and vision processing: https://sound.eti.pg.gda.pl/student/materialy.html
eResources addresses	Adresy na platformie eNauczanie:

Example issues/ example questions/ tasks being completed

Work placement

Not applicable

Document generated electronically. Does not require a seal or signature.