

**SYLLABUS**  
Academic year 2025-2026

Dean,  
Prof. dr. eng. Vasile-Ion Manta

**1. Program data**

1.1 Higher education institution	“Gheorghe Asachi” Technical University of Iași
1.2 Faculty	Automatic Control and Computer Engineering
1.3 Department	Computers
1.4 Field of studies	Computers and Information Technology
1.5 The cycle of studies <sup>1</sup>	Master
1.6 Study program	Artificial Intelligence

**2. Subject data**

2.1 Name of the subject / Code	<b>Voice and Speech Recognition (Recunoașterea vocii și a vorbirii) / AI.211</b>						
2.2 Course coordinator	Prof. dr. eng. Florina Ungureanu						
2.3 Application instructor	Lect. dr. eng. Alexandru-Gabriel Tudorache						
2.4 Year of study <sup>2</sup>	2	2.5 Semester <sup>3</sup>	1	2.6 Type of assessment <sup>4</sup>	colloquium	2.7 Type of subject <sup>5</sup>	DA

**3. Estimated total time of daily activities (hours per semester)**

3.1 Number of hours per week	2	3.2 lectures	1	3.3a sem.		3.3b laboratory		3.3c project	1
3.4 Total hours in curriculum <sup>6</sup>	28	3.5 lectures	7	3.6a sem.		3.6b laboratory		3.6c project	7
Distribution of the time fund <sup>7</sup>									No. hours
Study by textbook, course support, bibliography and notes									30
Additional documentation in the library, on specialist electronic platforms and in the field									20
Preparation of seminars/labs/projects, assignments, reports and portfolios									20
Tutorial <sup>8</sup>									-
Examinations <sup>9</sup>									2
Other activities:									-
3.7 Total hours of individual study <sup>10</sup>	72								
3.8 Total hours per semester <sup>11</sup>	100								
3.9 Number of credits	4								

**4. Prerequisites (where applicable)**

4.1 curriculum <sup>12</sup>	
4.2 competences	

**5. Conditions (where applicable)**

5.2 conducting the project <sup>13</sup>	<ul style="list-style-type: none"> <li>Laboratory room with computers and Internet access</li> <li>Open-source or free IDEs that allow students to develop code in C++ or Python (Visual Studio Community Edition / PyCharm)</li> </ul>
--	---

**6. Specific competences accumulated<sup>14</sup>**

<sup>1</sup> Bachelor / Master

<sup>2</sup> 1-4 for Bachelor's, 1-2 for Master's

<sup>3</sup> 1-8 for Bachelors, 1-3 for Masters

<sup>4</sup> Exam, colloquium or VP A/R – from the curriculum

<sup>5</sup> DF - fundamental subject, DID - subject in the field, DS - specialized subject or DC - complementary subject - from the education plan

<sup>6</sup> It is equal to 14 weeksx number of hours from point 3.1 (similar for 3.5, 3.6abc)

<sup>7</sup> The lines below refer to the individual study; the total is completed at point 3.7.

<sup>8</sup> Between 7 and 14 hours

<sup>9</sup> Between 2 and 6 hours

<sup>10</sup> The sum of the values on the previous lines, which refer to the individual study.

<sup>11</sup> The sum of the number of hours of direct teaching activity (3.4) and the number of hours of individual study (3.7); must be equal to the number of credits allocated to the subject (point 3.9)x 24 hours per credit.

<sup>12</sup> Mention the subjects that must be passed previously or equivalent

<sup>13</sup> Computing technique, software packages, experimental stands, etc.

Number of credits assigned to the subject <sup>15</sup> :		4	Distribution of credits per competences <sup>16</sup>
<b>Professional competences</b>	CP1	Knowledge of advanced concepts of computer science and information technology and the ability to work with these concepts.	0.7
	CP2	Scientific and practical research in the field of artificial intelligence.	0.7
	CP3	Problem solving using artificial intelligence methods and techniques.	1.6
	CP4	Design and development of artificial intelligence systems.	
	CP5	Utilization of artificial intelligence tools and technologies.	0.7
	CP6		
	CPS1		
	CPS2		
<b>Transversal competences</b>	CT1	Legislation compliant application of the intellectual property rights and of the principles, norms and values of the professional ethics code within their own strategies for rigorous, effective and responsible work.	0.2
	CT2	Application of communication techniques and effective group work; developing emphatic interpersonal communication skills and assuming leadership roles/functions in a multi-specialized team.	0.1
	CT3	Creating opportunities for continuous training and the effective utilization of learning resources and techniques for personal development.	
	CTS		

### 7. Objectives of the subject (resulting from the grid of specific competences accumulated)

7.1 General objective of the subject	A better understanding of the AI methods that are used to recognize (analyse and process) audio signals and translate them into a familiar language.
7.2 Specific objectives	The project aims to achieve a hands-on approach to audio processing, similar to the ones found in different corporations that work on projects with devices that capture and interpret the human voice.

### 8. Contents

8.1 Course	Teaching methods <sup>17</sup>	Remarks
<p><b>Week 1: Introduction.</b></p> <ul style="list-style-type: none"> <li>- recent developments</li> <li>- applications / impact / legal factors</li> <li>- dialogue / error rates</li> <li>- TTS (text-to-speech) concepts</li> <li>- extraction of Social Meaning from Speech</li> </ul> <p><b>Week 3: Phonetics</b></p> <ul style="list-style-type: none"> <li>- speech production: flow, resonance and articulation</li> <li>- waveforms, spectrogram</li> </ul> <p><b>Dialog systems</b></p> <ul style="list-style-type: none"> <li>- chatbots and (goal-based) dialog agents (conversational interface)</li> <li>- dialogue + task management</li> <li>- dialogue system design: finite-state dialog manager, frame-based systems</li> <li>- Natural language understanding (NLU) for filling slots + natural language generation (NLG)</li> <li>- Neural network review</li> </ul> <p><b>Week 5: ASR (Automatic Speech Recognition) Overview</b></p> <ul style="list-style-type: none"> <li>- architecture</li> <li>- the noisy channel model</li> <li>- the Viterbi algorithm</li> <li>- HMMs for speech</li> <li>- evaluation with word error rate</li> </ul>	Powerpoint presentations, answers to questions, explanations, case-studies	

<sup>14</sup> Competencies from the G1 and G1bis Grids of the study program, adapted to the specifics of the subject, for which credits are allocated (www.rncis.ro or the faculty website)

<sup>15</sup> From the education plan

<sup>16</sup> The credits allocated to the subject are distributed on professional and transversal competences according to the specifics of the subject

<sup>17</sup> Presentations, demonstrations, exercises, etc.

<p><b>Week 7: Acoustic modelling</b></p> <ul style="list-style-type: none"> <li>- <b>Hidden Markov Models and Gaussian Mixture Models (HMM-GMM); generative HMM-GMM ASR model</b></li> <li>- <b>acoustic modelling for Large Vocabulary Continuous Speech Recognition (LVCSR)</b></li> <li>- <b>Mel-Frequency Cepstral Coefficient (MFCC)</b></li> </ul> <p><b>Week 9: Recent ASR approaches</b></p> <ul style="list-style-type: none"> <li>- <b>RNN-Transducer</b></li> <li>- <b>Google on-device ASR, enabled by RNN-T</b></li> <li>- <b>conformer: Convolution-augmented Transformer for Speech Recognition</b></li> <li>- <b>dual mode ASR</b></li> <li>- <b>analysis of negative impacts of ML systems: harmful system and biased system</b></li> </ul> <p><b>Week 11: Foundation models</b></p> <ul style="list-style-type: none"> <li>- <b>Wav2Vec 2.0</b></li> <li>- <b>HuBERT</b></li> <li>- <b>XLS-R cross-lingual features</b></li> </ul> <p><b>Text-to-Speech introduction</b></p> <ul style="list-style-type: none"> <li>- <b>text analysis and waveform synthesis</b></li> <li>- <b>text normalisation</b></li> <li>- <b>TTS modelling history and overview; types of synthesis</b></li> <li>- <b>prosody and intonation</b></li> </ul> <p><b>Week 13: TTS</b></p> <ul style="list-style-type: none"> <li>- <b>evaluation and measurement: listening test paradigm, mean opinion score, AB tests; intelligibility tests, overall quality tests</b></li> <li>- <b>diagnostic rhyme test (DRT)</b></li> <li>- <b>data collection for TTS</b></li> <li>- <b>fundamental frequency (F0) generation by regression, declination</b></li> <li>- <b>unit selection intuition: target + join cost</b></li> <li>- <b>Parametric Synthesis: predicted parameters, vocoding / combination issues, prediction algorithms</b></li> <li>- <b>extracting social meaning with supervised ML (analysis tool); case studies</b></li> </ul>		
<p><b>Course references:</b></p> <ol style="list-style-type: none"> <li>1. Stanford University, CS224S: Spoken Language Processing, <a href="https://web.stanford.edu/class/cs224s/syllabus/">https://web.stanford.edu/class/cs224s/syllabus/</a></li> <li>2. MIT OpenCourseware, Automatic Speech Recognition, <a href="https://ocw.mit.edu/courses/6-345-automatic-speech-recognition-spring-2003/download/">https://ocw.mit.edu/courses/6-345-automatic-speech-recognition-spring-2003/download/</a></li> <li>3. R.G. Lyons: Understanding digital signal processing. Prentice-Hall, 2004</li> <li>4. Mert Pilanci, Signal processing for Machine Learning, <a href="https://web.stanford.edu/class/ee269/slides.html">https://web.stanford.edu/class/ee269/slides.html</a></li> <li>5. Dan Jurafsky, Introduction to Computer Speech and Language Processing, <a href="https://web.stanford.edu/class/linguist238/">https://web.stanford.edu/class/linguist238/</a></li> <li>6. Markus Kuhn, Digital Signal Processing, University of Cambridge, Computer Laboratory, <a href="https://www.cl.cam.ac.uk/teaching/0809/DSP/">https://www.cl.cam.ac.uk/teaching/0809/DSP/</a></li> <li>7. Voice recognition vs. speech recognition, <a href="https://verbit.ai/voice-recognition-vs-speech-recognition-the-difference-and-why-they-matter/">https://verbit.ai/voice-recognition-vs-speech-recognition-the-difference-and-why-they-matter/</a></li> </ol>		
8.2c Project	Teaching methods <sup>18</sup>	Remarks
<p><b>A C++ or Python project, which uses the concepts from the course in real-world applications, such as:</b></p> <ul style="list-style-type: none"> <li>- <b>automotive - infotainment, road assistant commands</b></li> <li>- <b>healthcare - log patient details, treatment, extract recovery development</b></li> <li>- <b>sales - develop an AI chatbot, capable of convincing customers / maximise sales</b></li> <li>- <b>virtual agents - develop tools similar to Google Assistant, Siri, Alexa .. that can integrate various commands (perhaps using IoT platforms as well)</b></li> </ul> <p><b>Week 2: Introduction / team selection, conceptual choices, project work</b></p> <p><b>Week 4: Diagram, project work</b></p>	General and individual explanations, individual/team work	

<sup>18</sup> Case study, demonstration, exercise, error analysis, etc.

<b>Week 6: Project work + milestone</b> <b>Week 8: Project testing, project work</b> <b>Week 10: Project testing, project work + milestone</b> <b>Week 12: Project work + final presentation</b> <b>Week 14: Final presentation</b>		
---	--	--

**Project references: see course references, up to date documentation for the actual selected platforms, tutorials**

1. Ravikiran A S, A Guide to Speech Recognition in Python: Everything You Should Know, <https://www.simplilearn.com/tutorials/python-tutorial/speech-recognition-in-python>
2. Alexa Development Tools and Skill Management APIs, Amazon, <https://developer.amazon.com/en-US/alexa/alexa-skills-kit/get-deeper/dev-tools-skill-management-api>

**9. Corroboration of the contents of the subject with the expectations of representatives of the epistemic community, professional associations and representative employers in the field related to the program<sup>19</sup>**

**10. Evaluation**

Type of activity	10.1 Evaluation criteria	10.2 Evaluation methods	10.3 Weight in the final grade
10.4a Colloquium	Acquired practical knowledge (quantity, correctness, accuracy)	Written theoretical test	30% (minimum 5)
10.4d Project	The quality of the completed project, the correctness of the project documentation, and the justification of the chosen solutions	<ul style="list-style-type: none"> <li>• Final project demonstration, test validation</li> <li>• Milestones</li> <li>• Discussions with students</li> </ul>	70% (minimum 5)
10.5 Minimum performance standard <sup>20</sup> : grade 5 in the colloquium and project work			

Date of completion,  
5 December 2023

Signature of course coordinator,  
Prof.dr.eng. Florina Ungureanu

Signature of application instructor,  
Lect. dr. eng. Alexandru-Gabriel  
Tudorache

Date of approval in the department,  
7 December 2023

Director of department,  
Assoc. prof. dr. eng. Andrei Stan

<sup>19</sup>The connection with other subjects, the usefulness of the subject on the labor market

<sup>20</sup>The minimum performance standard from the competences grid of the study program is customized to the specifics of the subject, if applicable.