#### card of course

|  |  |
| --- | --- |
| Subject name | Speech Recognition Models - Natural Language Processing Part 1 - Laboratory |

1. The placement of the subject in the study system

|  |  |
| --- | --- |
| 1.1. Field of study | Computer science |
| 1.2. Form and path of study | Full-time/Part-time |
| 1.3. Level of education | First-cycle studies |
| 1.4. Study profile | Practical |

|  |  |
| --- | --- |
| 1. 5. Specialty | Artificial intelligence |
| 1.6. Subject Coordinator | Dr Rafał Stęgierski |

2. General characteristics of the subject

|  |  |
| --- | --- |
| 2.1. Belonging to a subject group | Optional/practical |
| 2.2. Number of ECTS | 2 |
| 2.3. Language of lectures | English |
| 2.4. Semesters in which the subject is taught | IV |
| 2.5.Criteria for selecting course participants | For specialization: Artificial Intelligence |

1. Learning outcomes and course delivery
	1. Subject Objectives

|  |  |
| --- | --- |
| No. | Subject Objectives |
|
| C1 | Introduction to the basics of natural language processing (NLP), discussing key concepts and methods. |
| C2 | Developing skills in implementing simple NLP processes, using Python libraries such as NLTK and spaCy. |
| C3 | Preparation for analysis and implementation of simple speech recognition models using basic tools. |

* 1. Subject-specific learning outcomes, divided into knowledge , skills and competences , with reference to the directional learning outcomes

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Description of subject learning outcomes | Reference to directional effectslearning (symbols) | Method of implementation (mark "X") |
| ST | NST |
| Classes at the University | Activities on the platform | Classes at the University | Activities on the platform |
| After passing the course, the student knows and understands **the knowledge** |
| W1 | Knows the basic concepts of natural language processing, such as tokenization, stemming, lemmatization. | INF\_W10 INF\_W19 | X |  | X |  |
| W2 | Understands text representation methods including Bag-of-Words and TF-IDF. | X |  | X |  |
| W3 | Knows the algorithms of basic speech recognition models such as Hidden Markov Models (HMM). | X |  | X |  |
| After passing the course, the student is **able** to: |
| U1 | Can process text data, including tokenizing and lemmatizing texts in Python. | INF\_U13 INF\_U19 INF\_U27 | X |  | X |  |
| U2 | Can apply vectorization methods such as Bag-of-Words and TF-IDF to practical examples. | X |  | X |  |
| U3 | Can implement a simple speech recognition model using ready-made tools. | X |  | X |  |
| After completing the course, the student is ready to take part in **social competences.** |
| K1 | Is able to discuss in a group forum and gain knowledge from the specialist conducting the classes. | INF\_K01 | X |  | X |  |

3.3. Forms of teaching and their number of hours - Full-time studies (ST), Part-time studies (NST)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Path | Lecture | Exercises | Design | Workshop | Laboratory | Seminar | Lecturer | Classes conducted using distance learning methods and techniques in the form of ………………. | Other | **ECTS points** |
| **ST** |  |  |  |  | 20 |  |  |  |  | 2 |
| **NST** |  |  |  |  | 10 |  |  |  |  | 2 |

3.4. Content of education (separately for each form of classes: (W, ĆW, PROJ, WAR, LAB, LEK, OTHER). It should be marked (X) how the given content will be implemented (classes at the university or classes on the e-learning platform conducted using distance learning methods and techniques)

TYPE OF CLASS: LABORATORY

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Content of the course | Reference to subject-specific learning outcomes | Method of implementation (mark "X") |
| ST | NST |
| **Classes at the University** | **Activities on the platform** | **Classes at the University** | **Activities on the platform** |
| 1. | Introduction to natural language processing, basic concepts and applications. | W1 | X |  | X |  |
| 2. | Text tokenization and lemmatization, NLTK and spaCy tools. | W1, U1 | X |  | X |  |
| 3. | Text representations, Bag-of-Words and TF-IDF vectorization methods. | W2, U2 | X |  | X |  |
| 4. | Basics of speech recognition, Hidden Markov Models (HMM) algorithms. | W3, U3 | X |  | X |  |
| 5. | Building simple NLP pipelines, combining preprocessing with data analysis. | W1, U2, K1 | X |  | X |  |
| 6. | Summary of classes and discussion of grades |  | X |  | X |  |

3.5. Methods of verifying learning outcomes (indication and description of methods of conducting classes and verification of achievement of learning outcomes and method of documentation)

|  |  |  |  |
| --- | --- | --- | --- |
| Subject Effects | Teaching methods | Methods of verifying learning outcomes | Documentation methods |
| KNOWLEDGE |
| W1-W3 | Discussion and presentation of theoretical issues | Final test: Students solve tasks covering text analysis (tokenization, lemmatization), basic vectorization methods (Bag-of-Words, TF-IDF) and theoretical questions related to NLP algorithms and speech recognition models. | Colloquium archived on the platform |
| SKILLS |
| U1-U3 | Practical classes performed at computer stations | Final test: Students solve tasks covering text analysis (tokenization, lemmatization), basic vectorization methods (Bag-of-Words, TF-IDF) and theoretical questions related to NLP algorithms and speech recognition models. | Colloquium archived on the platform |
| SOCIAL COMPETENCES |
| K1 | Practical classes performed at computer stations | Final test: Students solve tasks covering text analysis (tokenization, lemmatization), basic vectorization methods (Bag-of-Words, TF-IDF) and theoretical questions related to NLP algorithms and speech recognition models. | Colloquium archived on the platform |

3.6. Assessment criteria for the achieved learning outcomes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Learning effect | For a grade of 3 or "pass."the student knows and understands/is able to/is ready to | For a grade of 3.5, the student knows and understands /is able to/is ready to | For a grade of 4, the student knows and understands/is able to/is ready to | For a grade of 4,5, the student knows and understands/is able to/is ready to | For a grade of 5, the student knows and understands/is able to/is ready to |
| W | 51-60% of knowledge indicated in learning outcomes | 61-70% of knowledge indicated in learning outcomes | 71-80% of knowledge indicated in learning outcomes | 81-90% of knowledge indicated in learning outcomes | 91-100% of knowledge indicated in learning outcomes |
| U | 51-60% of skills indicated in learning outcomes | 61-70% of skills indicated in learning outcomes | 71-80% of skills indicated in learning outcomes | 81-90% of skills indicated in learning outcomes | 91-100% of skills indicated in learning outcomes |
| K | 51-60% of skills indicated in learning outcomes | 61-70% of skills indicated in learning outcomes | 71-80% of skills indicated in learning outcomes | 81-90% of skills indicated in learning outcomes | 91-100% of skills indicated in learning outcomes |

3.7. Literature

**Basic**

Lewis Tunstall, Leandro von Werra, Thomas Wolf, "Przetwarzanie języka naturalnego z wykorzystaniem transformerów. Budowanie aplikacji językowych za pomocą bibliotek Hugging Face", Helion 2024

**Supplementary**

Valentino Zocca, Gianmario Spacagna, Daniel Slater, Peter Roelants, "Deep Learning. Uczenie głębokie z językiem Python. Sztuczna inteligencja i sieci neuronowe", Helion 2018

4. Student workload - ECTS points balance

|  |  |
| --- | --- |
| **Types of student activity** | **Student Load** |
| **ST** | **NST** |
| **Classes requiring direct contact between the student and the academic teacher at the university premises** | **20** | **10** |
| Classes included in the study plan | 20 | 10 |
| **Student's own work** | **30** | **40** |
| Ongoing preparation for classes, preparation of project work/presentations/etc. | 15 | 20 |
| Preparation for passing classes | 15 | 20 |
| **TOTAL STUDENT HOURLY LOAD** | **50** | **50** |
| **Number of ECTS points** | **2** | **2** |

|  |  |
| --- | --- |
| Last change date | 30/09/2024 |
| The changes were introduced | INF Education Quality Team |
| The changes were approved | Arkadiusz Gwarda, M.A. |