**SYLLABUS**

**regarding the qualification cycle FROM 2023TO 2024**

1. Basic Course/Module Information

|  |  |
| --- | --- |
| Course/Module title | *Methods of Data Mining* |
| Course/Module code \* |  |
| Faculty (name of the unit offering the field of study) | College of Natural Sciences |
| Name of the unit running the course | Institute of Computer Science |
| Field of study | Computer Science & Computer Science and Econometrics |
| Qualification level | First degree |
| Profile | Academic |
| Study mode | Full-time |
| Year and semester of studies | 3rd year, 5th semester |
| Course type | obligatory at the IDSS specialty |
| Language of instruction | English |
| Coordinator | Dr Krzysztof Balicki |
| Course instructor | Dr Wojciech Rząsa |

\* - as agreed at the faculty

* 1. Learning format – number of hours and ECTS credits

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Semester  (n0.) | Lectures | Classes | Colloquia | Lab classes | Seminars | Practical classes | Internships | others | **ECTS credits** |
| 5 |  | 30 |  |  |  |  |  |  | 5 |

1.2. Course delivery methods:

- conducted in a traditional way

- involving distance education methods and techniques

1.3. Course/Module assessment: pass with a grade

2. Prerequisites

|  |
| --- |
| Handling MS Excel program on basic level  Basic knowledge of statistical parameters, exhaustive algorithms and greedy ones |

3. Objectives, Learning Outcomes, Course Content, and Instructional Methods

3.1. Course/Module objectives

|  |  |
| --- | --- |
| O1 | A student should acquire a basic knowledge about the most fundamental and useful notions, concepts of data mining. |
| O2 | A student will know some methods and classical algorithms of data mining regarding supervised learning and unsupervised one. She/he understands how parameters of algorithms influence the output. |
| O3 | She/he will be able to simulate work of a few algorithms during exploration of sample data. |

3.2. Course/Module Learning Outcomes (to be completed by the coordinator)

|  |  |  |
| --- | --- | --- |
| Learning Outcome | The description of the learning outcome  defined for the course/module | Relation to the degree programme outcomes |
| LO\_01 | A student knows some models of knowledge discovery from databases |  |
| LO\_02 | A student understands the aim of the 3 following aspects of data mining: preprocessing, clustering and classification. She/he knows some methods and algorithms of the mentioned above three areas of data mining, including Cross‐Validation technique for classifier’s quality estimation |  |
| LO\_03 | A student can use MS Excel to import some real-life data, visualize data as 2D graphics. |  |
| LO\_04 | A student can make single imputation in case of data with missing values and normalize values of attributes |  |
| LO\_05 | A student can join cases into clusters according to k‐means algorithm and classify cases by means of k‐nn algorithm |  |
| LO\_06 | A student can simulate ID3 and C4.5 algorithms for simple data sets |  |
| LO\_o7 | A student is able to see possibility and advantage of using data mining for different data sets. |  |

**3.3. Course content (to be completed by the coordinator)**

1. Lectures

|  |
| --- |
| Content outline |
| Three models of data mining: academic, CRISP, hybrid ones |
| Scales of measure: nominal, ordinal, interval, ratio ones |
| Information system vs decision system |
| Dealing with some data imperfections: missing values, outliers, data inconsistency, undiscretized values |
| Clustering as an unsupervised learning technique (k-means algorithm) |
| Classifying as a supervised learning technique (kNN, ID3, C4.5) |
| Classifier quality estimation |

1. Classes, tutorials/seminars, colloquia, laboratories, practical classes

|  |
| --- |
| Content outline |
| Scales of measure: nominal, ordinal, interval, ratio ones |
| Information system vs decision system |
| Dealing with some data imperfections: missing values, outliers, data inconsistency, undiscretized values |
| Clustering as an unsupervised learning technique (k-means algorithm) |
| Classifying as a supervised learning technique (kNN, ID3, C4.5) |
| Classifier quality estimation |

3.4. Methods of Instruction

Lecture supported by a multimedia presentation / distance learning

Classes: project work (implementation project, practical project) / distance learning

4. Assessment techniques and criteria

4.1 Methods of evaluating learning outcomes

|  |  |  |
| --- | --- | --- |
| Learning outcome | Methods of assessment of learning outcomes (e.g. test, oral exam, written exam, project, report, observation during classes) | Learning format (lectures, classes,…) |
| LO-01 - LO-o4 | observation during classes | classes |
| LO-o5 – LO-o7 | project | lecture |

4.2 Course assessment criteria

|  |
| --- |
| Final project’s quality is a base for the assessment.  Entirely done, excellent project or with very minor imperfections – grade A  Partly done project with some minor errors – grade C  Partly done project with incidental major errors – grade E  Undone project or partly done project with some major errors – grade F |

5. Total student workload needed to achieve the intended learning outcomes

– number of hours and ECTS credits

|  |  |
| --- | --- |
| Activity | Number of hours |
| Scheduled course contact hours | 30 |
| Other contact hours involving the teacher (consultation hours, examinations) | 3 |
| Non-contact hours - student's own work (preparation for classes or examinations, projects, etc.) | 92 |
| Total number of hours | 125 |
| Total number of ECTS credits | 5 |

\* One ECTS point corresponds to 25-30 hours of total student workload

6. Internships related to the course/module

|  |  |
| --- | --- |
| Number of hours | *-----------------------------* |
| Internship regulations and procedures | *-----------------------------* |

7. Instructional materials

|  |
| --- |
| Compulsory literature:  Pointed parts of book: K.J. Cios, W. Pedrycz, R.W. Swiniarski, L.A Kurgan: Data Mining. A Knowledge Discovery Approach, Springer 2007 |
| Complementary literature: |

Approved by the Head of the Department or an authorised person