(faculty stamp)

COURSE DESCRIPTION

1. Course title: DATA ANALYSIS AND COMPUTATIONAL INTELLIGENCE	**2. Course code:** DAaCI
3. Validity of course description: 2018/2019	
4. Level of studies: BA, BSc programme	
5. Mode of studies: INTRAMURAL STUDIES	
6. Field of study: INFORMATICS	
7. Profile of studies: comprehensive	
8. Programme: ALL	
9. Semester: VII	
10. Faculty teaching the course: Faculty of Automatic Control, Electronics and Computer Science	
11. Course instructor: Marek SIKORA, Ph.D., D.Sc.	
12. Course classification: common courses	
13. Course status: obligatory	
14. Language of instruction: English	
15. Pre-requisite qualifications: Mathematics and Mathematical Logic, Algorithms and Data Structures	
16. Course objectives: The aim of the course is to show students basic algorithms of data analysis and exploration. The course raises elementary analitical tasks: outlier analysis, classification, regression, grouping and association analysis. Each of them is presented in the context of symbolic data representation (decision trees and rules) non-symbolic (support vector machines – coming from computational intelligence) and collective approaches (boosting, bagging).	
17. Description of learning outcomes:	

Nr	Learning outcomes description	Method of assessment	Teaching methods	Learning outcomes reference code
1	Student knows different algorithms of classification, regression, grouping, association analysis and outlier detection.	Final test	Lecutre, Laboratory	K1A_W02 K1A_W12
2	Student can apply algorithms of outlier and anomaly detection properly, even if they need some modifications.	Final test	Laboratory	K1A_U03 K1A_U08 K1A_U10 K1A_U14
3	Student can apply algorithms of data grouping properly, even if they need some modifications. Student can evaluate the grouping results.	Final test	Laboratory	K1A_U03 K1A_U08 K1A_U10 K1A_U14
4	Student can apply (and if it necessary – to modify) algorithms of association analysis and evaluate results on the background of training data.	Final test	Laboratory	K1A_U03 K1A_U08 K1A_U10 K1A_U14

5	Student can apply (and if it necessary – to modify) algorithms of classification and evaluate results on the background of training data.	Final test	Laboratory	K1A_U03 K1A_U08 K1A_U10 K1A_U14
6	Student can apply (and if it necessary – to modify) algorithms of regression and evaluate results on the background of training data.	Final test	Laboratory	K1A_U03 K1A_U08 K1A_U10 K1A_U14

18. **Teaching modes and hours**
Lecture / BA /MA Seminar / Class / Project / Laboratory
30 h (Lecture), 15 h (Laboratory),

19. **Syllabus description:**
Lectures and laboratories are leader in a traditional way. During the lecture the definitions, block diagrams of data analysis algorithms and computational intelligence methods are presented, together with sample applications. Students can observe the reasoning process, ask questions, participate and cooperate in algorithms modifications. Laboratories are assumed to be the place of practical analytical tasks applications, illustrating possibilities and purposefulness of using of methods presented at lectures.

Lectures:

1. Introduction to data analysis and exploration (typical tasks, evaluation of obtained analytical models, methods of experimental evaluation and analytical models comparison).
2. Introduction to RapidMiner environment.
3. Initial data quality evaluation. Data preprocessing.
4. Cluster analysis and clustering evaluation.
5. Association analysis (contingence tables, association rules, results evaluation).
6. Building classifiers of symbolic data representation (rules, trees).
7. Building classifiers of non-symbolic data representation (suport vector machines, naive Bayes, k-NN)
8. Building regression models of symbolic data representation (rules, trees).
9. Building regession models of non-symbolic data representation (suport vector machines, linear and non-linear regression, multiple regression, k-NN)

Collective approaches to classification and regression tasks (bagging, boosting).

Laboratory:

1. Data preprocessing
2. Clustering the data.
3. Association analysis.
4. Building and validation of classification models.
5. Building and validation of regression models.

20. Examination: NO

21. Primary sources:

1. Morzy T. Eksploracja danych. Wydawnictwo Naukowe PWN, Warszawa 2013.
2. Krawiec K., Stefanowski J. Uczenie maszynowe i sieci neuronowe. Wyd. Pol. Poznańskiej, 2003.
3. Cichosz P.: Systemy uczące się. WNT, Warszawa 2000.
4. Sikora M.: Wybrane metody oceny i przycinania reguł decyzyjnych. Monografia – Wydawnictwo Politechniki Śląskiej. Studia Informatica 33(3B), 2012 (dostępna jako pdf.).

22. Secondary sources:

1. Witten I.H., Frank E.: Data mining: practical machine learning tools and techniques. Morgan Kaufmann, 2011.
2. Han J., Kamber M.: Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, 2001.
3. M. North. Data Mining for the masses. A Global Text Project Book, 2012

23. Total workload required to achieve learning outcomes

Lp.	Teaching mode :	Contact hours / Student workload hours
1	Lecture	30/30
2	Classes	-/-
3	Laboratory	15/45
4	Project	-/-
5	BA/ MA Seminar	-/-
6	Other	10/20
	Total number of hours	55/95

24. Total hours: 150

25. Number of ECTS credits: 5

26. Number of ECTS credits allocated for contact hours: 2

27. Number of ECTS credits allocated for in-practice hours (laboratory classes, projects): 3

26. Comments:

Approved:

…….....
(date, Instructor's signature)

…….....
(date, the Director of the Faculty Unit signature)