Z1-PU7	WYDANIE N1

(*faculty stamp*)

COURSE DESCRIPTION

1. Course title: DATA MINING, Data mining in practice	**2. Course code DM_DMIP**
3. Validity of course description: 2018/2019	
4. Level of studies: MSc programme	
5. Mode of studies: intramural studies	
6. Field of study: CONTROL, ELECTRONIC AND INFORMATION ENGINEERING (MACRO)	(FACULTY SYMBOL) RAU-2
7. Profile of studies: ACADEMIC	
8. Programme: DATA SCIENCE	
9. Semester: 2	
10. Faculty teaching the course: Faculty of Automatic Control, Electronics and Computer Science	
11. Course instructor: Dr hab. Marek Sikora prof. nzw.	
12. Course classification: common courses	
13. Course status: compulsory ~~/elective~~	
14. Language of instruction: English	
15. Pre-requisite qualifications: Discrete Mathematics, Algorithms and data structures, Machine Learning, Soft Computing, Statistics	
16. Course objectives: The aim of the course is to make the students familiar with the methodology of the data exploration process, particularly with respect to complex-structure data. Use cases analysis will be presented, along with weak and strong points of particular analytical methods. The selected analytical platforms will be discussed.	

17. Description of learning outcomes:

Nr	Learning outcomes description	Method of assessment	Teaching methods	Learning outcomes reference code
1.	Student understands methodology of data exploration within the CRISP DM standard.	Credit	Lecture	K2A_W15, KW_16, K2A_W_17
2.	Student knows methods of preparation, cleaning, and improving quality of data.	Credit	Lecture	K2A_W15, K2A_W18
3.	Student knows methods of data modeling adequate to analytical task type. Student can estimate quality of models and can choose the one most adequate to the realized task.	Credit	Lecture	K2A_W17, K2A_W21, K2A_W27, K2A_W28
4.	Student knows basic methods of social networks analysis.	Credit	Lecture	K2A_W22, K2A_W23, K2A_W27, K2A_W28
5.	Student knows general principles of actions of internet browsing tools and recommendation systems.	Credit	Lecture	K2A_W22, K2A_W24, K2A_W26

6.	Student can analyze exemplary data set according to CRISP DM methodology.	Laboratory tasks	Laboratory	K2A_U01, K2A_U02
7.	Student is able to use results of analyses in the business process. Student can design methodology of monitoring quality of analytical models and their modifications in the course of operation.	Laboratory tasks	Laboratory	K2A_U09, K2A_U10, K2A_K07
8.	Student can use main platforms of analytics and data exploration (RapidMiner, TensorFlow).	Laboratory tasks	Laboratory	K2A_U17, K2A_U18
9.	Student is able to perform, with the use of the chosen tool, the analysis of the social network, can transform the network, identify societies, point out critical nodes and can visualize the network to support the analysis.	Laboratory tasks	Laboratory	K2A_U15, K2A_K01, K2A_K02

18. Teaching modes and hours

Lecture 15 / ~~BA /MA Seminar / Class / Project~~ / Laboratory 15

19. Syllabus description:

Lecture:

1. Cross Industry Standard Process for Data Mining data mining methodology.
2. Data preparation and cleaning (feature selection, feature extraction, missing values, unbalanced data)
3. Analytical model developing, selection and evaluation (classification, regression, survial/reliability analysis)
4. Model deployment and maintenance (analytical model deployment, batch mode, incremental mode, concept drift)
5. Social network analysis I (network characteristics and measures, network extraction)
6. Social network analysis II (network visualisation, community identification)
7. Beyond PageRank Intelligent search and recommendation engines

Laboratory:

1. Data mining - use cases (RapidMiner, TensorFlow: classification – seismic and methane hazard a
2. assessment, regression – gas and fuel consumption forecasting).
3. Data mining – use cases (RapidMiner, TensorFlow – churn analysis, predictive maintenance, market basket analysis).
4. Social Network analysis – use cases (Pajek, Gephi – network transformation, identification

20. Examination: semester NO

21. Primary sources:

Witten I. H., Frank E., Hail M.A: Data Mining. Practical Machine Learning Tools and Techniques. Fourth Edition. Theory and its applications, Morgan Kaufmann / Elsevier 2017.
Wasserman, S., Faust, K. Social network analysis: Methods and applications. Cambridge university press. 1994

22. Secondary sources:

Ahlemeyer-Stubbe A., Coleman S.: A practical guide to data mining for business and industry, Wiley 2014.

23. Total workload required to achieve learning outcomes

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

24. Total hours: 60

25. Number of ECTS credits:2

26. Number of ECTS credits allocated for contact hours: 1

27. Number of ECTS credits allocated for in-practice hours (laboratory classes, projects):1

26. Comments:

Approved:

…….
(date, Instructor's signature)

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