Bialystok University of Technology									
Field of study	**Computer Science**							**Degree level and programme type**	**Engineer's degree full-time programme**
Specialization/ diploma path	**---**							**Study profile**	academic
Course name	**Data Exploration**							**Course code**	**FCS-00091**
								Course type	obligatory
Forms and number of hours of tuition	L	C	LC	P	SW	FW	S	**Semester**	2
	30				30			**No. of ECTS credits**	**6**

Entry requirements	
Course objectives	The course prepares the student to select and apply independently methods and algorithms to data mining tasks often encountered in practice. Special attention will be paid to computational aspects related to the implementation of tasks of mining large data sets.
Course content	L: Data preprocessing. Selection of variables for analysis. Principal components method. Methods of classification (linear and nonlinear decision rules, Bayesian decision rules), regression, clustering. Evaluation of classifiers. Classification and regression trees. Methods of discovering associations in data, sequence patterns. Survival analysis. Multidimensional scaling. Linear separability of multidimensional datasets, linearization of datasets by rank layers of binary classifiers. Application of selected methods and algorithms for detecting regularities (patterns) in large data sets will also be analyzed. SW: Preparing data for analysis (normalization, standardization, discretization). Selected algorithms for data classification. Selected cluster analysis algorithms. Decision tree. Association rules. Random forests. Bagging and boosting. Evaluation of obtained models.
Teaching methods	informative lecture, programming, project method,
Assessment method	Lecture - oral evaluation SW - realization of partial tasks, evaluation of reports from partial tasks, realization and presentation of the project

Symbol of learning outcome	**Learning outcomes**	**Reference to the learning outcomes for the field of study**
LO1	knows basic methods, techniques and tools used in data exploration	K_W05
LO2	is able to use the known methods and models for the analysis and evaluation of algorithms and for data analysis	K_U05 K_U10
LO3	can construct models in the area of data exploration and skillfully use them	K_U10
LO4	can practically use basic methods, techniques and tools for data exploration	K_U10

Symbol of learning outcome	**Methods of assessing the learning outcomes**	**Type of tuition during which the outcome is assessed**
LO1	pass a lecture	L
LO2	project evaluation, project presentation, evaluation of reports	SW
LO3	project evaluation, project presentation, evaluation of reports	SW
LO4	project evaluation, project presentation	SW

Student workload (in hours)	**No. of hours**
Calculation	
1 - Participation in practical classes -	30
2 - Participation in lectures -	30
3 - Preparation of reports and carrying out homework -	25
4 - Participation in consultations -	5
5 - Implementation of project tasks (including preparation of presentations) -	35
6 - Preparation for passing -	25
TOTAL:	**150**

Quantitative indicators	**HOURS**	**No. of ECTS credits**
Student workload – activities that require direct teacher participation	65 (1)+(4)+(2)	2.6
Student workload – practical activities	90 (1)+(3)+(5)	3.6

Basic references	1. R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification, John Wiley, wydanie drugie, New York, 2001.
Supplementary references	1. R. A. Johnson, D. W. Wichern: Applied Multivariate Statistical Analysis, Prentiice-Hall, Upper Saddle River 2002.

Organisational unit conducting the course	Software Department	**Date of issuing the programme**
Author of the programme	dr inż. Magdalena Topczewska	Feb. 17, 2022

L – lecture, C – classes, LC – laboratory classes, P – project, SW – specialization workshop, FW – field work,S – seminar