1.	Course		*Mining massive datasets*	
2.	Code		KNI_E30	
3.	Study programme		Computer Science and Engineering PhD study programme	
4.	Study programme organized by		FCSE	
5.	Cycle		Third – PhD	
6.	Academic year / semester winter/summer/elective		7. ECTS credits 7,5	
8.	Teacher		Prof. d-r Dejan Gjorgjevikj, Prof. d-r Gjorgji Madzarov	
9.	Prerequisites		None	
10.	Course programme goals (competences): The students will attain in depth understanding of the machine learning and data mining techniques for massive data sets. They will be able to successfully apply machine learning algorithms when solving real problems concerning business intelligence, social networks, web data description. They will be able to concept, analyze, realize and evaluate the developed system performances.			
11.	Course syllabus: The MapReduce model, associative rules, nearest neighbor search in high dimensional data, reducing dimensions, locality sensitive hashing), recommendation systems, massive data sets clustering techniques, link analysis, machine learning techniques for massive data sets, data stream mining, structural data relations retrieval, web advertising, massive data industry examples.			
12.	Teaching methods: Classes supported with slide presentations, interactive teaching, lab equipment and other software packages, teamwork, case studies, invited guest lecturers, presentations of project works, e-learning materials, forums and consultations.			
13.	Total fund of work hours		7,5 EKTC x 30 h = 225 h	
14.	Available hours distribution		45+30+150 = 225	
15.	Teaching activities	15.1.	Theoretical classes	45 h
		15.2.	Practical classes (labs, exercises), seminars, team work	30 h
16.	Other activities	16.1.	Project tasks	50 h
		16.2.	Self study	50 h
		16.3.	Homework	50 h
17.	Grading			
	17.1.	Tests		40 points
	17.2.	Seminar work/ project (presentation: written and oral)		50 points
	17.3.	Active participation		10 points
18.	Grading criteria (points/grade)		to 59 points	5 (five) (F)
			from 60 to 68 points	6 (six) (E)
			from 69 to 76 points	7 (seven) (D)

		from 77 to 84 points	8 (eight) (C)
		from 85 to 92 points	9 (nine) (B)
		from 93 to 100 points	10 (ten) (A)
19.	Conditions for attending the final exam	Successful completion of activities 15.1 and 15.2	
20.	Language	Macedonian or English	
21.	Quality assessment	Internal evaluation and student pools	

22. Literature

22.1. Compulsory

No.	Author	Title	Publisher	Year
1.	Anand Rajaraman, Jeffrey David Ullman	Mining of Massive Datasets	Cambridge University Press	2011
2.	Michael Minelli, Michele Chambers, Ambiga Dhiraj	Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses	Wiley CIO	2013
3.	Thomas C. Redman	Data Driven: Profiting from Your Most Important Business Asset	Harvard Business Press	2008

22.2. Additional

No.	Author	Title	Publisher	Year
1.				
2.				
3.				