COURSE OUTLINE

(1) GENERAL

SCHOOL	SCHOOL OF SCIENCES		
ACADEMIC UNIT	DEPARTMENT OF STATISTICS		
LEVEL OF STUDIES	POSTGRADUATE PROGRAM		
COURSE CODE	333-1007	SEMESTER	B
COURSE TITLE	STATISTICAL TECHNIQUES IN BIG DATA MINING		

INDEPENDENT TEACHING ACTIVITIES	WEEKLY TEACHING HOURS	CREDITS
	2	6

COURSE TYPE	SPECIALISED GENERAL KNOWLEDGE
PREREQUISITE COURSES:	NO
LANGUAGE OF INSTRUCTION and EXAMINATIONS:	GREEK
IS THE COURSE OFFERED TO ERASMUS STUDENTS	YES
COURSE WEBSITE (URL)	http://www.samos.aegean.gr/samos_actuar/modules_eng.html

(2) LEARNING OUTCOMES

Learning outcomes
Students will be able to: extract information from a data set and transform it into an understandable structure for further use. Also the students after the end of the course would have the opportunity to understand, analyze and apply data mining techniques in big data
General Competences
Search for, analysis and synthesis of data and information, with the use of the necessary technology Decision-making Working independently and Team work Working in an interdisciplinary environment

(3) SYLLABUS

What is data mining, Types of data, Data quality, Data preprocessing. Classification the general approach, Decision trees induction, rule based classifiers, classical and Bayesian and neural nets classifiers, nearest neighbour classifiers. Association patterns. Cluster analysis, Similarity and Distance, characteristics of clustering algorithms. Hierarchical clustering.

(4) TEACHING and LEARNING METHODS - EVALUATION

DELIVERY	Face-to-face
USE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY	Communication with students via e-mail Posting slides and course material on the website
TEACHING METHODS	

Activity	*Semester workload*
Lectures	24
Independent study	63

	assignments	63
	Course total (25 per ECTS)	***150***
STUDENT PERFORMANCE EVALUATION	Student evaluation is done in Greek through a written examination which includes short-answer questions and problem solving. For students with disabilities, evaluation takes place via oral exams.	

(5) ATTACHED BIBLIOGRAPHY

- Suggested bibliography:

1. Tan P.N., Steinbach M., Kumar V., Introduction to Data Mining (2005), ISBN: 0321321367.
2. Han J., Kamber M., Data Mining: Concepts and Techniques (2000), ISBN: 1558604898.