Course Name : Data Mining II

Course Code	Course Type	Regular Semester	Lecture (hours/week)	Seminar (hours/week)	Lab. (hours/week)	Credits	ECTS
CMP 509	C	Fall	3.00	1.00	0.00	3.50	6.00

Lecturer	Eris Zeqo, PhD
Assistant	Damiana Teliti, Msc
Course language	Albanian
Course level	Master
Description	“Data Mining II” focuses on advanced knowledge discovery techniques including big data analysis, advanced classification, clustering, anomaly detection, and mining data streams. The emphasis is on applying state-of-the-art algorithms in real-world contexts.
Objectives	To explore advanced data analysis techniques. To apply classification and clustering algorithms in complex scenarios. To analyze streaming and unstable data. To understand anomaly detection and personalization techniques.
Core Concepts	Advanced classification and performance improvement Hierarchical, fuzzy, and DBSCAN clustering Data stream mining Anomaly and outlier detection Dimensionality reduction (PCA, t-SNE) Model selection and ensemble techniques

Course Outline

Week	Topic
1	Introduction & Review of Data Mining I
2	Advanced Classification: Boosting and Bagging
3	Feature Selection & Dimensionality Reduction
4	Fuzzy Clustering and DBSCAN
5	Mining Data Streams
6	Anomaly and Outlier Detection
7	Personalization and Recommendation Systems
8	Midterm Exam
9	Text and Unstructured Data Mining
10	Ensemble Methods: Random Forests, Stacking
11	Performance Evaluation & Cross-Validation
12	Data Mining in Big Data Contexts
13	Practical Tools (Weka, Scikit-learn, RapidMiner)
14	Ethics and Privacy in Data Analysis
15	Final Project Presentations / Review
16	Final Exam

Prerequisites	The student must attend the course at a minimum rate of 75%.
Literature	• Jiawei Han, Micheline Kamber, Jian Pei – Data Mining: Concepts and Techniques, 4th Edition, Morgan Kaufmann, 2022.
References	• Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar – Introduction to Data Mining, 2nd Edition, Pearson, 2019. • Charu C. Aggarwal – Data Mining: The Textbook, Springer, 2015. • Trevor Hastie, Robert Tibshirani, Jerome Friedman – The Elements of Statistical Learning, Springer, 2017.

Course Outcome

1	Students will be able to apply advanced data mining techniques on real-world and large-scale datasets.
2	They will be able to choose appropriate algorithms for classification, clustering, or anomaly detection.
3	They will use practical tools for both structured and unstructured data analysis.
4	They will understand the importance of ethics and privacy protection in data analysis.

Course Evaluation

In-term Studies	Quantity	Percentage
Midterms	0	0
Quizzes	0	0
Projects	1	20
Term Projects	1	20
Laboratory	0	0
Class Participation	0	0
Total in-term evaluation percent		**40**
Final exam percent		**60**
Total		**100**

ECTS Workload (Based on Student Workload)

Activities	Quantity	Duration (hours)	Total (hours)
Course duration (Including the exam week: 16x Total hours of the course)	16	4	64
Study hours outside the classroom (Preparation, Practice, etc.)	14	2	28
Duties	2	24	48
Midterms	0	0	0
Final Exam	1	10	10
Other	0	0	0
Total Work Load			**150**
Total Work Load / 25 (hours)			**6.00**
ECTS			**6.00**