SVEUČILIŠTE U RIJECI
FAKULTET ZA MATEMATIKU

Sveučilište u Rijeci •Fakultet za matematiku
Radmile Matejčić 2 • 51 000 Rijeka • Hrvatska
T: (051) 584-650 • F: (051) 584-699
http://www.math.uniri.hr• e-adresa: math@math.uniri.hr

COURSE SYLLABUS

General information		
Course title	Optimization techniques for data mining	
Study programme	Graduate course Discrete Mathematics and Its Applications	
Year of study	1.	
Course status	Compulsory	
Course homepage	https://moodle.srce.hr	
Language of instruction	English	
Credit values and modes of instruction	***ECTS credits / student workload***	6
	Hours (L+E+S)	30+15+15
Lecturer	***Name and surname***	Daniel Hawtin
	Office	O-507
	Office hours	By appointment
	Phone number	584-668
	E-mail	dhawtin@math.uniri.hr
Teaching assistant	***Name and surname***	Daniel Šanko
	Office	O-318
	Office hours	Thursdays at 16h
	Phone number	584-676
	E-mail	daniel.sanko@uniri.hr

1. COURSE DESCRIPTION

1.1. Course objectives

The goal of the course is to acquire a basic knowledge of databases, with particular emphasis on relational databases, and to familiarize students with terms, algorithms, and mathematical techniques used in data mining, i.e., discovering patterns in large data sets. For this purpose, the course will include:

- introducing basic concepts about databases and performing simple and complex database queries,
- introducing basic concepts and algorithms related to data mining,
- illustratrating the application of the developed algorithms in data mining,
- connecting different branches of mathematics (especially probability and statistics) as a theoretical basis for most algorithms in data mining,
- introducing a programming language for data mining,
- introduce programming language associated with data mining.

1.2. Course prerequisites

None.

1.3. Learning outcomes

After completing the course, students will be able to:

O1. explain basic concepts from database theory and concepts of the relational data model (A4, B5, C5, E4, F4, G4),
O2. analyze and process a large amount of data (A5, B5, C5, E5, F5, G4),
O3. define and understand the basic concepts of data mining (A4, B5, C5, E4, F4),
O4. describe the basic techniques of data mining (A5, B5, C5, E4, F4),
O5. analyze and compare different algorithms for data mining (A5, B5, C5, E4, F4),
O6. solve problems typical for data mining (A5, B5, C6, D5, E4, F4, G7),
O7. design simple algorithms for data mining (A7, B5, C7, D4, E4, F7, G7),
O8. evaluate the effectiveness of the algorithms presented (A7, B6, C7, D5, E5, F7, G7).

1.4. Course content

Introduction to databases. Relational data model. Relational algebra. Performing database queries. Operations in the relational model. Introduction to data mining. Data warehouses. Data analysis and processing. Discovery and presentation of knowledge in mining. Algorithms in data mining: associative rule, classification, prediction. Evaluation of knowledge. Implementation of mining in real databases. Clustering. Advanced methods in data mining.

1.5. Modes of instruction	☒ lectures ☐ seminars and workshops ☒ exercises ☒ e-learning ☐ field work	☒ independent work ☒ multimedia and the internet ☐ laboratory ☐ tutorials ☒ mentoring work ☐ consultative teaching ☐ other
1.6. Comments		

1.7. Student requirements

The student's work in the course is evaluated and graded during the semester and on the final exam. Students can earn 70 grading points during the semester and 30 grading points on the final exam. To pass the course, a student must achieve at least 50% of the total points, satisfying the minimum requirements from table 2.2.

MIDTERM EXAMS (total 40 points):

During the semester, two (written) midterm exams with questions from the covered material will be held at the times indicated in table 5. It is possible to accumulate a total of 20 points on each midterm exam. There will be an opportunity for a retake or makeup exam at the scheduled time (you can retake the exam on which you got less points).

SEMINAR (total 20 points):

Each student (or group of students) will receive a topic on which they must prepare and submit a seminar paper within the specified deadline and present it at the designated time. By preparing and presenting the seminar paper, a student can earn a maximum of 20 points. Seminars cannot be submitted later and are not corrected during makeup activities.

QUIZZES (total 10 points):

Two shorter knowledge assessments will be conducted during the semester. On each knowledge assessment, a student can earn a maximum of 5 points. Students do not have the option to take these assessments at a later date, and they are not corrected during makeup activities.

FINAL EXAM (total 30 points):

On the final exam a student can earn a maximum of 30 points. A student is considered to have passed the final exam if they have achieved at least 15 points on it.

SVEUČILIŠTE U RIJECI
FAKULTET ZA MATEMATIKU

Sveučilište u Rijeci •Fakultet za matematiku
Radmile Matejčić 2 • 51 000 Rijeka • Hrvatska
T: (051) 584-650 • F: (051) 584-699
http://www.math.uniri.hr• e-adresa: math@math.uniri.hr

1.8. Constructive Connection

LEARNING OUTCOMES	CONTENT	TEACHING ACTIVITIES	GRADING METHODS
Q1	Introduction to databases. Relational data model. Relational algebra. Performing database queries. Operations in the relational model.	- lectures, - discussion, - computer exercises, - independent work, - seminar In the teaching process, the following methods will be used: - Oral Presentation Method - Discussion Method - Practical Work Method - Writing Method - Demonstration Method	Written exam, Oral exam
Q2, Q3, Q4	Introduction to data mining. Data warehouses. Data analysis and processing. Discovery and presentation of knowledge in mining.		Written exam, Oral exam
Q5, Q6, Q7, Q8	Algorithms in data mining: associative rule, classification, prediction. Evaluation of knowledge. Implementation of mining in real databases. Clustering. Advanced methods in data mining.		Written exam, Oral exam, seminar

2. GRADING POLICY

2.1. Grading of students' work during the semester and on the final exam

A student's performance in the course will be evaluated and graded during the semester as well as on the final exam. The total number of points a student can earn during the semester is 70 (based on described student activities). Through all forms of continuous monitoring and assessment during the semester, a student must accumulate at least 50% of grading points in order to be eligible to take the final exam. It is possible to earn a maximum of 30 points on the final exam. The passing threshold on the final exam cannot be lower than 50% of successfully completed exam tasks.

Students who earn between 0% and 49.9% of the grading points that could have been obtained through forms of continuous monitoring and assessment during the semester will receive a grade of F (Insufficient) and cannot earn ECTS credits. They must retake the course. The same applies to students who fail the final exam in three offered exam sessions.

2.2. Minimal requirements for access to the final exam / passing grade

ACTIVITY	
Midterm Exam	20
Seminar	10
TOTAL:	35
OTHER REQUIREMENTS:	**/**

SVEUČILIŠTE U RIJECI
FAKULTET ZA MATEMATIKU

Sveučilište u Rijeci •Fakultet za matematiku
Radmile Matejčić 2 • 51 000 Rijeka • Hrvatska
T: (051) 584-650 • F: (051) 584-699
http://www.math.uniri.hr• e-adresa: math@math.uniri.hr

2.3. Final grade – grading scale

Based on the total sum of grading points earned during the course and on the final exam, the final grade is determined according to the following distribution:

GRADE	POINTS
Excellent (5) , A	90% - 100%
Very good (4), B	75% - 89,9%
Good (3), C	60% - 74,9%
Sufficient (2), D	50% - 59,9%
Insufficient (1), F	0% - 49,9%

3. LITERATURE

3.1. Required literature

1. J. Leskovec, A. Rajaraman, J. D. Ullman, Mining of Massive Datasets, Cambridge University Press, 2014.
2. Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar, Introduction to Data Mining, 2nd ed., Pearson, 2019.

3.2. Recommended literature

1. B. Schölkopf, A. J. Smola, Learning with Kernels. Support Vector Machines, Regularization, Optimization, and Beyond, MIT Press, Massachusetts, 2002.
2. T. Hastie, R.Tibshirani, J. Friedman, Data Mining, Inference, and Prediction, Springer-Verlag New York, 2009.

4. ADDITIONAL INFORMATION

4.1. Class attendance

Any form of disruption during the class will not be tolerated as well as the usage of mobile phones.

4.2. Informing students

All relevant informations will be provided via the online course. It is the responsibility of a student to be regularly informed.

4.3. Other relevant information

4.4. Assessment of quality and performance for the course

At the end of the semester, an anonymous survey will be conducted in which students will evaluate the quality of the classes held. After the end of the semester, an analysis of the performance of the students in the exams held in that semester will be conducted.

4.5. *Examination terms*

Summer	30.06.2025. 14.07.2025.
Fall	01.09.2025.

SVEUČILIŠTE U RIJECI
FAKULTET ZA MATEMATIKU

Sveučilište u Rijeci •Fakultet za matematiku
Radmile Matejčić 2 • 51 000 Rijeka • Hrvatska
T: (051) 584-650 • F: (051) 584-699
http://www.math.uniri.hr• e-adresa: math@math.uniri.hr

5. COURSE OUTLINE

DATE	*TIME*	*MODE OF INSTRUCTION*	*TOPIC*	*GROUP*	*LECTURE HALL*
04.03.	12:15 - 13:45	L	Course Introduction On Data Mining and Applications	SVI	O - 360
06.03.	14:15 - 15:45	E	Course Introduction. Basics of the Python Programming Language. Working with Data, Fundamentals of Databases. Data Loading, Manipulation, Data Visualization.	SVI	O - 363
11.03.	12:15 - 13:45	L	MapReduce	ALL	O - 360
13.03.	14:15 - 15:45	E	MapReduce	ALL	O - 363
18.03.	12:15 - 13:45	L	Finding similar itmes	ALL	O - 360
20.03.	14:15 - 15:45	E	Finding similar itmes	ALL	O - 363
25.03.	12:15 - 13:45	L	Data mining	ALL	O - 360
27.03.	14:15 - 15:45	E	PageRank algorithm	ALL	O - 363
01.04.	12:15 - 13:45	L	Frequent itemsets, Market-Basket model, Association rules	ALL	O - 360
03.04.	14:15 - 15:45	E	Frequent itemsets, Market-Basket model, Association rules	ALL	O - 363
08.04.	12:15 - 13:45	L	Advanced Methods for Frequent Itemset Detection	ALL	O - 360
10.04.	14:15 - 15:45	E	**I. MIDTERM**	ALL	O - 363
15.04.	12:15 - 13:45	L	Clustering	ALL	O - 360
17.04.	14:15 - 15:45	E	Clustering	ALL	O - 363
22.04.	12:15 - 13:45	L	Recommendation systems	ALL	O - 360
24.04.	14:15 - 15:45	E	Recommendation systems	ALL	O - 363
29.04.	12:15 - 13:45	L	Mining Social-Networks graphs	ALL	O - 360
06.05.	12:15 - 13:45	L	Dimensionality reduction	ALL	O - 360
08.05.	14:15 - 15:45	E	Mining Social-Networks graphs	ALL	O - 363
13.05.	12:15 - 13:45	L	Dimensionality reduction	ALL	O - 360
15.05.	14:15 - 15:45	E	**II. MIDTERM**	ALL	O - 363
20.05.	12:15 - 13:45	L	Social-Network graphs	ALL	O - 360
22.05.	14:15 - 15:45	S	Student Presentations	ALL	O - 363
27.05.	12:15 - 13:45	L	Machine learning for databases	ALL	O - 360
29.05.	14:15 - 15:45	S	Student Presentations	ALL	O - 363
03.06.	12:15 - 13:45	L	Neural networks and deep learning	ALL	O - 360
05.06.	14:15 - 15:45	S	Student Presentations	ALL	O - 363
10.06.	12:15 - 13:45	L	Advertising on the Web	ALL	O - 360
12.06.	14:15 - 15:45	S	Student Presentations	ALL	O - 363
17.06.	14:15 - 15:45	S	Student Presentations	ALL	O - 363

Minor changes are possible.
Up to 40% of the teaching activities may be online.

L – lectures
E – exercises
S – seminars