UCLouvain

linfo2364

2025

Mining Patterns in Data

The version you're consulting is not final. This course description may change. The final version will be published on 1st June.

5.00 credits	30.0 h + 15.0 h	Q2

Language :	English > French-friendly
Place of the course	Louvain-la-Neuve
Main themes	An important task in data mining is the discovery of patterns in data. Patterns are recurring structures in data; they can provide interpretable explanations for observations in data, can help to gain a better understanding in the structure of data, can be used to build better models, and can be used to solve other computational tasks (such as the construction of database indexes or data compression). Patterns can be found in many different forms of data, including data from supermarkets, insurance companies, scientific experiments, social networks, software projects, and so on. This course will provide an in-depth introduction to pattern mining. After an introduction to the basics of pattern mining, it will provide an in-depth discussion of a number of advanced pattern mining techniques. Topics that will be discussed are: • Categories of pattern mining tasks, including pattern and pattern set mining, supervised and unsupervised pattern mining, dataset types,and pattern scoring functions; • Algorithms for solving different pattern mining tasks; • Data structures for making pattern mining more efficient; • The implementation of pattern mining algorithms; • Mathematical foundations for the different categories of pattern mining tasks; • Complexity classes relevant to pattern mining; • Applications of pattern mining, with a special focus on the application of pattern mining techniques in software engineering.
Learning outcomes	**At the end of this learning unit, the student is able to :** Given the learning outcomes of the "Master in Computer Science and Engineering" program, this course contributes to the development, acquisition and evaluation of the following learning outcomes: • INFO 1 • INFO 2.1-4 • INFO 4.2-4 • INFO 5.5 • INFO 6.4 Given the learning outcomes of the "Master [120] in Computer Science" program, this course contributes to the development, acquisition and evaluation of the following learning outcomes: • SINF 1.M4, 1.M3 • SINF 2.1-4 • SINF 4.2-4 • SINF 5.5 • SINF 6.4 Students completing this course successfully will be able to • Identify the most appropriate pattern mining task for a given data set ; • Explain the advantages and disadvantages of pattern mining algorithms in relation to the problem to be solved ; • Identify appropriate approaches for evaluating the quality of patterns and apply them in various situations ; • Determine the computational complexity of pattern mining problems; • Develop new pattern mining algorithms for new applications.

1

Evaluation methods	The final grade is determined by 3 projects and an exam that is organized at the end of the semester. The grade is calculated following a 75% / 25% rule (final written exam / participation and grade obtained for projects during the semester). Every project counts equally. Failure to comply with the methodological instructions communicated by the teacher, particularly with regard to the use of online resources or collaboration between students, will result in an overall mark of 0. The use of generative AI tools **without prior permission** is strictly prohibited.
Teaching methods	• Lectures • Exercise sessions, during which exercises will be done that prepare for the exam and projets • 3 projets Even though preference will be given to face-to-face lectures and exercise sessions, depending on the health situation and the number of students enrolled, other forms of teaching (online, co-modal or hybrid) may be considered.
Content	• Frequent itemset mining: algorithms, data structures; • Constraint-based itemset mining: algorithms, data structures; • Patterns in sequences, trees, graphs: algorithms, data structures, complexity classes; • Pattern mining in supervised data: scoring functions, algorithms; • Pattern set mining in supervised data: scoring functions, models (decision trees, boosting), algorithms • Pattern set mining in unsupervised data: scoring functions (minimum description length principle, maximum entropy), algorithms • Applications of pattern mining: software repositories, traces, log files, cheminformatics, bioinformatics, industrial applications
Inline resources	https://moodle.uclouvain.be/course/view.php?id=3069
Bibliography	Charu C. Aggarwal, Jiawei Han (Eds.), Frequent Pattern Mining, Springer 2014 (ISBN: 978-3-319-07820-5) Chapitres de Siegfried Nijssen, Albrecht Zimmermann and Luc De Raedt, Essentials of Pattern Mining.
Other infos	During this course students wil have to implement a number of projects in Python. This course is impossible to follow without prior knowledge of Python. Hence, students should have followed a prior course in Python, such LEPL1401, LINFO1101 or LSINC1101.
Faculty or entity in charge	INFO

Programmes containing this learning unit (UE)

Program title	Acronym	Credits	Prerequisite	Learning outcomes
Master [120] in Data Science : Statistic	DATS2M	5		
Master [120] in Computer Science and Engineering	INFO2M	5		
Master [120] in Computer Science	SINF2M	5		
Master [120] in Mathematical Engineering	MAP2M	5		
Master [120] in Data Science Engineering	DATE2M	5		
Master [120] in Data Science: Information Technology	DATI2M	5		