Web Mining

La fouille du Web

Domain : Mathematics and computer science
Field : IT
Master's Program : RTIC

Presented by : Dr. Benazi Makhlouf
Year of study : second year
Academic year : 2024/2025

Practical Information

- **Course Title:** Web Mining

- **Teaching unit** : Fundamental (UEF 3.2)

- **Semester** : 03

- **Master's title** : RTIC

- **Coefficient** : 2

- **Credits** : 4

- **VH** 1.5h lecture, 1.5h tutorials 4 Personal work

- **evaluation method :** tutorials : 50%, Final Exam: 50%.

- **Course Instructor**: Benazi Makhlouf

- **Email** : makhlouf.benazi@univ-msila.dz

- **Useful sites**: https://elearning.univ-msila.dz/moodle/

Course objectives

The goal of this module is to provide comprehensive insights into web mining, focusing on web content, structure, and usage. It also covers various social network analysis techniques.

Prerequisite knowledge:

Students should have a solid understanding of basic data mining concepts, including approaches and algorithms such as association rules, classification, and clustering.

References:

Yanchun Zhang, editor "Web Mining and Social Networking", 2011, Springer.

Anthony Scime , "Web Mining," 2005, Idea Group Publishing .

Content of the subject:

1. **Introduction**:

   ○ Overview of data mining and web mining.

   ○ Web communities and social networks.

2. **Basic Concepts Recap**:

   ○ Web data models.

   ○ Similarity functions.

   ○ Information retrieval and performance evaluation.

   ○ Social network fundamentals.

3. **Web Content Mining**:

   ○ Vector space model.

   ○ Web search techniques.

   ○ Latent Semantic Indexing (LSI).

   ○ Automatic topic extraction.

Content of the subject:

4. **Web Structure Mining**:

   o PageRank and HITS algorithms.

   o Web community discovery.

   o Graph-based modeling.

   o Link information classification.

5. **Web Usage Mining**:

   o User interest modeling via clustering.

   o Latent semantic analysis.

   o User access pattern discovery.

   o Web log (weblog) exploitation.

6. **Extraction and Analysis of Social Networks on the Web**:

   o Evolution of web communities.

   o Social behavior analysis.