ITMD 561 SYLLABUS

ITMD 561 Web Intelligence

Hours: 3 credit hours / 45 contact hours

Instructor: Yong Zheng, Ph.D.

Textbook, title, author, and year:

Recommender Systems Handbook 2nd edition, Francesco Ricci and Lior Rokach, 2015
Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit, Steven Bird, Ewan Klein, and Edward Loper, 2009

Specific course information

a. **Catalog description:** Embrace the dynamic landscape of the digital realm with our groundbreaking course in Web Intelligence. This course is designed to empower students with a profound understanding of Web Mining, Natural Language Processing (NLP), Information Retrieval (IR), and Recommender Systems (RecSys), pivotal components in shaping the intelligent future of the web. Web Mining can uncover the hidden gems within the vast expanse of the web through the exploration of web mining techniques. NLP is a popular topic in AI and job market. From sentiment analysis to language generation, students will gain hands-on experience in developing systems that can understand, interpret, and generate human language effectively. We will also navigate the intricacies of IR and learn the art of efficiently accessing and presenting relevant information from the vast web ecosystem. Topics in RecSys help students explore the art and science of personalized content delivery. This course will feature both knowledge (e.g., concepts, algorithms) and practical skills (e.g., tools and libraries by Python) in these related topics.
b. **Prerequisites:** ITMD 514

Specific goals for the course

a. **Course Outcome:** Upon completion of the Web Intelligence course, students will emerge with a comprehensive skill set and understanding of web mining, Natural Language Processing (NLP), information retrieval, and recommender systems, as well as skills of using Python to solve problems in these topics. They will demonstrate mastery in applying web mining techniques to extract valuable insights from diverse data sources, effectively implementing NLP algorithms for tasks like sentiment analysis and text summarization. Additionally, students will showcase the ability to design and optimize information retrieval systems, utilizing search algorithms and relevance ranking, while also constructing and evaluating personalized recommender systems. Through interdisciplinary integration, they will tackle complex problems, fostering a holistic approach to web intelligence. Ethical considerations will be ingrained, enabling responsible decision-making in the development and implementation of intelligent systems. Practical applications and hands-on projects will highlight students' capacity to address real-world challenges, fostering innovation and creativity in the evolving landscape of web technologies. Effective communication of complex technical concepts will be emphasized, preparing students for diverse roles in data science, machine learning, and web development.
b. **Course Student Outcomes:**
Upon successful completion of the course the student should be able to do the following:
   - Demonstrate proficiency in web mining techniques to extract patterns and insights from diverse web data sources.
   - Apply Natural Language Processing (NLP) algorithms for tasks such as sentiment analysis, named entity recognition, and text summarization.
   - Design and implement efficient information retrieval systems, employing search algorithms, indexing strategies, and relevance ranking.
   - Construct collaborative filtering, content-based filtering, and hybrid recommender systems for delivering personalized content recommendations.
   - Integrate knowledge from web mining, NLP, information retrieval, and recommender systems to solve interdisciplinary problems.
   - Execute hands-on projects and practical applications, applying acquired knowledge to address contemporary challenges in these topics.
   - Deliver optimal technical and policy technology solutions for the problems of business, industry, government, non-profit organizations, and individuals using web intelligence technologies and methods.

Topics to be covered

a. Overview of Data Mining & Machine Learning
b. Web Mining: Web Usage Mining
c. Web Mining: Web Structure Mining
d. NLP: Basic Techniques in Text Mining
e. NLP: Semantic Technologies
f. NLP: Distributed Representations and LLMs
g. NLP Applications: Sentimental Analysis, Web Content Mining
h. Introduction to Information Retrieval
i. IR: Data Preprocessing, Basic IR by using NLP
j. Exam and IR: Ranking Methodologies
k. Introduction to RecSys: Categories, Tasks, Evaluation Methods
l. Recommendation Algorithms: Collaborative Filtering
m. Recommendation Algorithms: Content-Based Approaches and Hybrid Models
n. Different Types of Recommender Systems
o. Final Project Presentations