Recommender Systems

Iordanis Koutsopoulos, Associate Professor, AUEB, jordan@aueb.gr, 210 82 03 933

Overview

The course will cover fundamental and practical aspects of Recommender systems, focusing on theory as well as on the practical use and applications of Recommender systems. Recommender systems are around us and are encountered on multiple domains such as e-commerce, content and media distribution, social media and so on. The course aims to explain both basics and advanced topics and concepts for recommender systems.

Key Outcomes

By completing the course the students will be able to:

- Understand the basic concepts of recommender systems
- Solve mathematical optimization problems pertaing to recommender systems
- Carry out performance evaluation of recommender systems based on various metrics
- Implement machine-learning and data-mining algorithms in recommender systems data sets.
- Design and implement a simple recommender system.
- Learn about advanced topics and current applications of recommender systems in other realms such as mobile computing.

Requirements and Prerequisites

The students should have a basic calculus and algorithms background. Basic knowledge in machine learning and data mining will be very helpful, although the various machine-learning and data-mining concepts will be explained when encountered in the course. A good command in programming will be useful as well.

Required Course Materials

There is no required textbook. All course materials will be provided in class and will be available for downloading. The students may need to bring their laptops in class in some lectures in order to try out interactively the material being presented.

Books

There are many books on the subject, and a lot of free resources on the Internet. We will use part of the material from the following books:

- C.C. Aggarwal, Recommender Systems: The Textbook, Springer, 2016.
- F. Ricci, L Rokach, B. Shapira and P.B. Kantor, Recommender systems handbook, Springer 2010.
- J. Leskovec, A. Rajaraman and J. Ullman, Mining of massive datasets, 2nd Ed., Cambridge, 2012. (Chapter 9).
- M. Chiang, Networking Life, Cambridge, 2010. (Chapter 4).

Software/Computing requirements

The Python programming language and its libraries and tools that are relevant to mathematical optimization and data processing will be useful for the course. For the project assignments, we will leverage public datasets from different domains such as movies, products or services recommender systems. These datasets will be announced in the class.

Grading

The final grade will be computed as follows:

- Final exam: 50% of the final grade.
- 2 projects, each counting towards 25% of the grade.

Note that the projects count towards the final grade only if a 5/10 performance in the final is exam is achieved.

Participation

In-class contribution is an important part of our shared learning experience. You are expected to actively participate in the discussions in the class during the interactive part of the class. Please arrive to class on time and stay to the end of the class period. Chronically arriving late or leaving class early is unprofessional and disruptive to the entire class. Repeated tardiness will have an impact on your grade. Turn off all electronic devices prior to the start of class. Cell phones tablets and other electronic devices are a distraction to everyone.

Assignments

There will be two project assignments. Indicative subjects can be:

- Processing and analysis of public recommender systems datasets, and performance evaluation and comparison.
- Buidling a simple recommender system, focusing more on the functionality rather than the user interface.

Late assignments will either not be accepted or will incur a grade penalty unless they are due to documented serious illness or family emergency. Exceptions to this policy for reasons of civic obligations will only be made available when the assignment cannot reasonably be completed prior to the due date, you make suitable arrangements, and give notice for late submission in advance.

Attendance Requirements

Class attendance is essential to success in this course and is part of your grade. An excused absence can only be granted in cases of serious illness or grave family emergencies and must be documented. Job interviews and incompatible travel plans are considered unexcused absences. Where possible, please notify the instructor in advance of an excused absence.

Students are responsible for keeping up with the course material, including lectures, from the first day of this class, forward. It is the student's obligation to bring oneself up to date on any missed coursework.

Code of Ethics

Students may not work together on graded assignments unless the instructor gives express permission.

Exercise integrity in all aspects of one's academic work including, but not limited to, the preparation and completion of all other course requirements by not engaging in any method or means that provides an unfair advantage. In any case of doubt, students must be able to prove that they are the sole authors of their work by demonstrating their knowledge to the instructor.

Clearly acknowledge the work and efforts of others when submitting written work as one's own. Ideas, data, direct quotations (which should be designated with quotation marks), paraphrasing, creative expression, or any other incorporation of the work of others should be fully referenced. No plagiarism of any sort will be tolerated. This includes any material found on the internet. Reuse of material found in question and answer forums, code repositories, other lecture sites, etc., is unacceptable. You may use online material to deepen your understanding of a concept, not for finding answers.

Please report observed violations of this policy. Any violations will incur a fail grade at the course and reporting to the senate for further disciplinary action.

Course Syllabus

The course comprises ten units of three hours each. The following topics will be covered:

Working with recommender systems datasets. Advanced topics: context-aware recommender systems; user interaction with Recommender systems and impact on content caching.

Unit 1: Introduction

Introduction and basic taxonomy of recommender systems (RSs). Traditional and non-personalized RSs. Overview of data mining methods for recommender systems (similarity measures, classification, Bayes classifiers, censembles of classifiers, lustering, SVMs, dimensionality reduction). Overview of convex and linear optimization principles.

Unit 2: Content-based recommender systems

The long-tail principle. Domain-specific challenges in recommender systems. Content-based recommender systems. Advantages and drawbacks. Basic componenets of content-based RSs. Feature selection. Item representation Methods for learning user profiles.

Unit 3: Collaborative Filtering (CF)-based RSs: Mathematical foundations

Mathematical optimization in CF RSs. Optimization objective. Baseline predictor through least squares. Regularization and overfitting. Temporal models. Step-by-step solution of the RS problem.

Unit 4: Collaborative Filtering (CF)-based RSs: systematic approach

Nearest-neighbor collaborative filtering (CF). User-based and item-based CF, comparison. Components of neighborhood methods (rating normalization, similarity weight computation, neighborhood selection). Hybrid recommender systems.

Unit 5: Advanced CF methods

Matrix factorization models and dimensionality reduction. Matrix Decomposition. Latent factor models. Solution via alternative projections method. Examples. The Netflix data challenge. Constraint-based RSs. Introduction to tensors and their applications.

Unit 6: Performance evaluation of RSs

Experimental setttings. Working with RSs data sets. Examples. The cold-start problem. Evaluation metrics. Rating prediction and accuracy. Other metrics (fairness, coverage, diversity, novelty, serendipity).

Unit 7: Context awareness and Learning principles in RSs

Context-aware recommender systems. Contextual information models for RSs. Incorporating context in Rs. Learning to rank. Active learning in RSs. Multi-armed bandits and Reinforcement learning in RSs. Dynamic RSs.

Unit 8: User behavior understanding in RSs

Foundations of behavioral science. User choice and decisions models. Choice models in RSs. Digital nudging and user choice engineering principles. Applications and examples for recommender systems.

Unit 9: Applications of RSs for content media, social media and communities

Music and video RSs. Datasets. Group recommender systems. Social recommendations. Recommending friends: link prediction models. Similarities and differences of RSs with task assignment in mobile crowdsensing. Social network diffusion awareness in RSs.

Unit 10: Advanced topics: Network aspects of content RSs

Recommender systems for video content distribution. Implications of recommender systems in 5G wireless networks. RSs for optimizing wireless network performance. Case studies (i) Joint content recommendations and content caching in small cells wireless networks (ii) The interplay of RSs and User access point association.