SYLLABUS: INFO 4871/5871

REVISION 2, 2/19/2019

RECOMMENDER SYSTEMS

SPRING 2019, 3 CREDITS, 15 JANUARY – 2 MAY

INSTRUCTOR INFORMATION

Name: Dr. Robin Burke
Office Location: TLC 219
Office hours: Tuesday and Thursday, 2:00 pm – 3:00 pm and by appointment.

Contact: (Slack) info4871-s19.slack.com, (Email) robin.burke@colorado.edu
Note: Office currently has no phone. Phone number will be posted on Canvas when available.

COURSE COORDINATES

Lecture: Tuesdays and Thursdays 3:30 pm – 4:45 pm, Ketchum 1B71

COURSE INFORMATION

This is a research seminar that will explore the space of personalized information access applications known as recommender systems. This class will introduce students to a range of approaches for building recommender systems including collaborative, content-based, knowledge-based, and hybrid methods. Students will also explore a variety of applications for recommendation including consumer products, music, social media, and online advertising. The course will also examine controversies surrounding recommendation, including Pariser's "filter bubble", the deployment of personalization as a tool for electoral manipulation, and questions of algorithmic bias.

PREREQUISITES

INFO 4871 restricted to students with 57-180 credits (Junior or Senior). My expectation is that students in this class will be able to develop programs and manage data with the Python programming language and its associated data analysis tools including the Pandas and matplotlib libraries. Experience with the Jupyter notebook platform would be helpful. Students should be familiar with basic concepts of

statistics, including random variables, tests of statistical significance, and least-squares regression. Knowledge of regularization, linear algebra and multi-variate calculus is helpful but not required.

LEARNING GOALS

The learning goals of this course are aligned with the Baccalaureate Learning Goals for the University. In particular, those goals that state Baccalaureate Graduates of the University of Colorado Boulder will be able to:

- Formulate and investigate research, creative work and open-ended questions,
- Locate, evaluate and apply relevant evidence and technologies to solve problems in their disciplinary areas of study,
- Work collaboratively and individually

After taking this course, all students will be able to:

- Define key recommendation approaches, including collaborative, content-based, and hybrid recommendation, and associated data sources.
- Define evaluation paradigms for recommender systems, including k-fold cross-validation, A/B testing, and temporal evaluation.
- Define different learning approaches used in collaborative recommendation including instance-based, matrix factorization, learning-to-rank, and random walk.
- Interpret the results of recommender systems experiments using a variety of metrics including RMSE, nDCG, and AoC.
- Implement basic recommendation algorithms.
- Perform comparative experiments of recommendation algorithms.
- Transform raw data into formats appropriate for recommendation algorithms.
- Apply recommender systems ideas in an independent group project.

In addition to the goals above, graduate students in this class will be able to:

- Summarize recent research articles in recommender systems.
- Identify differences between recommendation domains and their consequences for the choice of recommendation approach.
- Develop and execute an independent research group project.

TEXTBOOKS AND MATERIALS

Textbooks: *Recommender Systems: The Textbook*, Charu C. Aggarwal, Springer International Publishing, 2016, 978-3-319-29657-9.

Note: The book can be downloaded in PDF and eBook format from Springer through the CU Boulder Library. You can also purchase a copy through Springer's "MyCopy" system for $25.

Other Readings: Other readings will be available through Canvas.

Programming platform: Programming assignments will be completed in Python 3 using the LKPy recommender systems framework using Jupyter notebook. Installation of the latest version of Anaconda 3 is recommended (www.anaconda.com/distribution/).

iClicker: This class will make use of iClicker technology. Each student will be expected to acquire an iClicker device and bring it to each class session. (For more information on clickers, see oit.colorado.edu/services/learning-spaces-technology/cuclickers/help/students)

ASSIGNMENTS

There are four different graded items in this course: course participation, programming assignments, quiz questions and answers, and projects. Graduate students have different project requirements.

PARTICIPATION

The participation grade has two components. The first is based on participation in the iClicker interactive response questions used in (most) class sessions after Week 1. Students get 1 point if they participate in at least 75% of the questions asked during a given class session. The lowest three session grades are dropped, for a total of 25 points.

The second participation component consists of 3 reaction papers worth 10 points each. There are three points in the semester when your fellow students will be presenting: Domain Days (2/19 and 2/21), Project Presentations (4/23, 4/25, 4/30), and Project Posters (5/2). For each presentation type, students will pick one presentation and write a 2-page reaction paper. A suggested outline will be provided.

PROGRAMMING ASSIGNMENTS

There are five programming / data analysis assignments to be completed during the quarter. All assignments will be completed in Python 3 using the Jupyter notebook platform. Data and scaffolding code will be supplied.

QUIZZES

This class will use the "social quiz" site SCuiz for testing students' knowledge of course content: scuiz-server.appspot.com. Each student must create questions related to the course material in order to get access to questions to answer. Each correctly-answered question is worth one point. To get full credit for this course element, you will need to accumulate 260 points – about 20 answers per week, starting week 2. An additional 20 points will be accepted as extra credit. You get access to 10 questions when

you create one, so you will need to create, on average, 2 questions per week. Note that the system automatically filters out questions that are too easy, so if you wait too long to participate, you will be left with only difficult questions and may have a hard time getting the needed points. You can also gain points by challenging questions that are incorrect or confusing. INFO 4871 and 5871 have different SCuizzes.

GROUP PROJECTS

The projects in this course will be group projects. Groups will be assigned by the instructor. However, students will be able to nominate classmates with whom they would like to work.

INFO 4871

Students enrolled in INFO 4871 will complete a group project using a topic chosen in collaboration with the instructor. Full details on the project requirements will be available later in the semester. The project can take one of several forms: an implementation augmentation project, in which the students add some new functionality to an existing recommender system or platform, a data+evaluation project in which the students develop a new recommendation data set and use it to perform a comparative evaluation of different recommendation algorithms or systems, or an implementation project in which the students create a prototype recommender systems implementation.

The project has five milestones: a project proposal, a midterm progress report, a project poster, an in-class presentation of the project poster, and a project report.

INFO 5871

Students enrolled in INFO 5871 will complete two group projects over the course of the quarter.

"Domain Days" presentation: The first project is an in-class presentation on a recommendation application domain. The domain will be assigned by the instructor.

Research project: The second project will be a research project on a topic chosen by the group in collaboration with the instructor. The project can take one of several forms: a research proposal accompanied by a detailed literature review, a replication of a published study of an existing recommendation algorithm, or a small-scale research study.

The project has five milestones: a project proposal, a midterm progress report, a project presentation, an in-class delivery of the presentation, and a project report.

High-quality projects may be nominated by the instructor for revision and submission to a suitable peer-reviewed publication venue (e.g. as a paper, poster or demo to the ACM International Conference on Recommender Systems, or similar conference.)

GRADING

4871 STUDENTS

Participation: (35 points) 10%
Assignments: (50 points) 30%
SCuiz points: (260 points + 20 points extra credit) 20%
Group project: (200 points) 40%

5871 STUDENTS

Participation: (35 points) 10%
Assignments: (50 points) 30%
SCuiz points: (260 points + 20 points extra credit) 20%
Domain presentation: (50 points) 10%
Group project: (150 points) 30%

GRADING SCALE

93.0% and above:	A
90.0%--92.9%:	A-
87.0%--89.9%:	B+
83.0%--86.9%:	B
80.0%--82.9%:	B-
77.0%--79.9%:	C+
73.0%--76.9%:	C
70.0%--72.9%:	C-
67.0%--69.9%:	D+
63.0%--66.9%:	D
60.0%--62.9%:	D-
Below 60.0%:	F

LATE POLICY

Students in this class have four late days that can be used for an assignment or combination of assignments with no penalty. After the four days have been used, no further late assignment will be accepted, unless a documented need for accommodation is provided. The following individual assignments cannot be submitted late: social quiz answers, reaction papers due on 5/5.

COURSE CALENDAR

Note: Calendar is tentative and may change during the semester. Any changes will be discussed in class

and announced via email and Canvas. Check Canvas for the most accurate information.

Class Format

The format of each class will vary and will include lecture, discussion, activities, and student presentations. Readings are expected to be completed before class.

Week 1 (1/15, 1/17), Introduction

Syllabus and course expectations. Tools and resources. Overview and history of recommender systems.
1/17: Reading, Aggarwal, Ch. 1
1/17: Social quiz I start

Week 2 (1/22, 1/24), Collaborative Recommendation

User profiles and ratings. Neighborhood-based methods.
Due: Acquire clicker and bring to class
1/22: Reading, Aggarwal, Ch. 2.1-2.4
1/22: Homework 1 assigned

Week 3 (1/31), Regression Models

I will be out of town this week. Class is canceled on 1/29. There will be a guest instructor on 1/31.
1/29: No class
1/31: Reading, Aggarwal, Ch. 2.6

Week 4 (2/5, 2/7), Matrix Factorization

Mathematical foundations. Singular value decomposition. Matrix factorization. Introduction to the LKPy package.
2/5: Reading, Aggarwal, Ch. 3.6-3.8
2/5: Homework 1 due
2/5: Group agreement due
2/5: Domain assigned

Week 5 (2/12, 2/14), Learning Factorizations / Content-Based Recommendation

Factorization via gradient descent and alternating least squares.
2/12: Reading,
2/12: Homework 2 assigned

Week 6 (2/19, 2/21), Domain Days

2/18: Social quiz I due
2/19, 2/21: (INFO 5871) Recommendation application domain presentation due

Week 7 (2/26, 2/28), Project Brainstorming

2/26: Domain days reaction paper due
2/26: Attendance for INFO 4871 students only
2/26: Homework 2 due
2/28: Attendance for INFO 5871 students only

Week 8 (3/5, 3/7), Evaluation

Recommender systems evaluation. On-line and off-line methods. Accuracy and ranking metrics. User reception of recommendations.
3/5: Reading, Aggarwal, Ch. 7
3/5: Project proposal due
3/5: Homework 3 assigned

Week 9 (3/12, 3/14), Factorization (conclusion) / Content-based Recommendation

SVD++, Non-negative matrix factorization. Wrap-up. Content-based recommendation. Naïve Bayes.
3/14: Reading, Aggarwal, Ch .4

Week 10 (3/19, 3/21), Hybrid Recommendation and Context-Aware Recommendation

Hybrid recommender systems designs. Defining and representing context. Contextual modeling and filtering.
3/19: Homework 3 due
3/19: Reading, Aggarwal, Ch. 6
3/19: Homework 4 assigned
3/21: Reading, Aggarwal, Ch. 8

Week 11, Spring Break

Week 12 (4/2, 4/4), Recommendation in Networks

Recommendation tasks in networks. Ranking algorithms. Neighborhood methods. Random walks. Heterogeneous network methods.
4/1: Social quiz II due
4/2: Homework 4 due
4/2: Reading, Aggarwal, Ch. 2.7 and Ch. 10
4/2: Social quiz III start
4/2: Project progress report
4/2: Homework 5 assigned

Week 13 (4/9, 4/11), Learning to Rank

Optimizing ranking metrics. Smooth approximations. Pair-wise and list-wise methods.

4/9: Reading, Aggarwal, Ch. 13.2, Rendel, et al. “BPR: Bayesian Personalized Ranking from Implicit Feedback”, UAI 2009; and Shi, et al. “CLiMF: Learning to Maximize Reciprocal Rank with Collaborative Less-is-More Filtering”, RecSys 2012.

Week 14 (4/16, 4/18), Fairness-Aware Recommendation

Criticisms of personalization. The “filter bubble” and algorithmic bias. Multistakeholder and multi-objective recommendation. Fairness-aware recommendation models.
4/16: Reading, Dwork, et al. Fairness Through Awareness, ITCS 2012; Kamishima, et al. Model-Based Approaches for Independence-Enhanced Recommendation, FATML 2016; and Burke, Multi-sided Fairness for Recommendation, FATML 2018.
4/16: Homework 5 due

Week 15 (4/23, 4/25), Presentations 1 & 2

Student presentations of final projects.
4/23, 4/25: (INFO 5871) Project presentation due

Week 16 (4/30, 5/2), Presentations 3 / Posters

Student presentations of final projects.
4/30: (INFO 5871) Project presentation due
5/2: (INFO 4871) Poster presentation due

Finals Week (5/5), No In-Person Meeting

5/5: Project report due
5/5: Project presentation reaction paper due
5/5: Project poster reaction paper due
5/5: Social quiz III due

Course Technology

Canvas

Course materials, including lecture slides, homework handouts, data sets, and readings will be posted to the Canvas site for this course. Grade information will also be uploaded to Canvas.

SCuiz

Students will be assigned accounts on the SCuiz social quiz site. No enrollment emails are sent. You will need to go to scuiz.crowdgrader.com and create an account using the email address in course roster. If you do not see a widget indicating that the quiz for this course is available, please let me know as soon as possible.

SLACK

This course will use Slack for course discussion and help. This collaboration tool is accessible via the web at info4871-s19.slack.com and also via dedicated desktop and mobile apps. When project groups are formed, each group will have its own Slack group. You are encouraged to use these groups as they are accessible to me and I can assist with project questions.

ACCOMMODATION STATEMENT

I am committed to providing everyone the support and services needed to participate in this course. If you qualify for accommodations because of a disability, please submit to your professor a letter from Disability Services in a timely manner so that your needs can be addressed. Disability Services determines accommodations based on documented disabilities. Contact Disability Services at 303-492-8671 or by e-mail at dsinfo@colorado.edu. If you have a temporary medical condition or injury, see Temporary Medical Conditions: Injuries, Surgeries, and Illnesses guidelines under Quick Links at Disability Services website and discuss your needs with me.

RELIGIOUS OBSERVANCES

Campus policy regarding religious observances requires that faculty make every effort to deal reasonably and fairly with all students who, because of religious obligations, have conflicts with scheduled exams, assignments or required assignments/attendance. If this applies to you, please speak with me directly as soon as possible at the beginning of the term.

CLASSROOM BEHAVIOR

Students and faculty each have responsibility for maintaining an appropriate learning environment. Those who fail to adhere to such behavioral standards may be subject to discipline. Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, color, culture, religion, creed, politics, veteran's status, sexual orientation, gender, gender identity and gender expression, age, ability, and nationality. Class rosters are provided to the instructor with the student's legal name. I will gladly honor your request to address you by an alternate name or gender pronoun. Please convey this information on the student information sheet distributed in class. For more information, see the policies on class behavior and the student code.

DISCRIMINATION AND HARASSMENT

The University of Colorado Boulder (CU-Boulder) is committed to maintaining a positive learning, working, and living environment. CU-Boulder will not tolerate acts of discrimination or harassment

based upon Protected Classes or related retaliation against or by any employee or student. For purposes of this CU-Boulder policy, "Protected Classes" refers to race, color, national origin, sex, pregnancy, age, disability, creed, religion, sexual orientation, gender identity, gender expression, veteran status, political affiliation or political philosophy. Individuals who believe they have been discriminated against should contact the Office of Discrimination and Harassment (ODH) at 303-492-2127 or the Office of Student Conduct (OSC) at 303-492-5550. The full policy on discrimination and harassment has more information.

HONOR CODE

All students of the University of Colorado at Boulder are responsible for knowing and adhering to the academic integrity policy of this institution. Violations of this policy may include: cheating, plagiarism, aid of academic dishonesty, fabrication, lying, bribery, and threatening behavior. All incidents of academic misconduct shall be reported to the Honor Code Council (honor@colorado.edu; 303-735-2273). Students who are found to be in violation of the academic integrity policy will be subject to both academic sanctions from the faculty member and non-academic sanctions (including but not limited to university probation, suspension, or expulsion). The Honor Code Office has more information.