Syllabus in Year and Term for CS 417 Machine Learning General Information

Contact Your Instructor

Name: *instructor name*
Email: *instructor email*
Phone: *instructor office phone #*
Office: Root Hall, *instructor office*

Lecture, Exam, Office Hours

Lecture: *days, times, location*
Section and CRN: *section #'s and CRN #'s*
Credit Hours: *# of credit hours*

Exam: *day, time*. Also check the Office of the Registrar's exam schedule

Instructor Office Hours: *your official office hours. You can also mention when you are normally available.*

GA Tutoring: See http://cs.indstate.edu/info/labs.html

Website: *link to course website if there is one*

Prerequisites

C or better in CS 401

Recommended Text

Note that the same texts are used for CS 417 as are used for CS 401, as follows.

The class will primarily use free online sources, including the online version of *R for Data Science*. The following are also recommended as supplementary texts.

Introduction to Computation and Programming Using Python: With Application to Understanding Data, 2nd edition, by Guttag, 2016 MIT Press.
Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, by McKinny, 2013 O'Reilly Media.
R Programming for Data Science, by Peng, 2015 Lean Publishing.
R for Data Science, by Grolemund et al., 2017 O'Reilly Media.

There are many free online resources to supplement the texts, including the following.

- Harvard online - Data Analysis for the Life Sciences (statistics in R)
- EdX Microsoft - Introduction to Python for Data Science

Course Announcements
Announcements regarding the course will be made both during class and via email to your @sycamores.indstate.edu email address. You should regularly check this email account or have it forwarded to an account that you check regularly.

Classroom conduct
You may not use cell phones, iPods/music players, etc. during class. You should be civil and respectful to both the instructor and your classmates, and you should arrive to class a few minutes before the scheduled lecture so you are ready for lecture to begin on time. You may use your computer during class if you are using it to follow along with the examples that are being discussed. You may not check email, facebook, work on other courses, etc. during class.

Course Description
The official description of this course from the catalog is

"A continuation of CS 401. A thorough study of data mining and machine learning algorithms, the overall process and the main techniques used in data mining and machine learning, including exploratory data analysis, predictive modeling, descriptive modeling, and model evaluation."

Course Outline
1. Data Mining
    - Problem types
        i. Clustering
        ii. Classification
        iii. Pattern detection
    - Algorithms
        i. k-means
        ii. k-nearest neighbor
        iii. Association rules – Apriori
        iv. Page rank
        v. Expectation maximization
2. Machine Learning
    - Naive Bayes classifiers
    - Maximum entropy classifiers
    - Classification trees
    - Artificial Neural Networks
        i. Feed forward networks
        ii. Convolutional networks
        iii. Recurrent networks
    - Linear regression
    - Logistic regression
    - Regression trees
    - Random forests - bagging and boosting

Learning Outcomes

The following learning outcomes from the CS BS outcomes library should be achieved during this course. Note that outcomes 3E.i and 3E.ii are met in CS 401 and are reinforced in CS 417. Outcomes 3E.iii, 3E.iv, and 3E.v are introduced in CS 401 and mastered in CS 417.

3E.i Can write programs to use data science packages to solve problems involved large datasets.

3E.ii - Can write programs to read in data from a variety of data formats, including cleaning and combining datasets.

Can explain the most important data science concepts, including tradeoffs and typical use cases
3E.iii - machine learning algorithms
3E.iv – regression
3E.v – modeling

Expected Amount of Work

If you take this class seriously and get what you should out of it, some weeks you will likely be spending around ***6 hours/week*** or more on the class. The students who get A's in their CS courses and have an easy time finding jobs do spend this much time on this course. Not everyone would need to spend this much time and not all weeks will be the same, but you should plan on putting in whatever time it takes. Note that the federal government definition of 1 credit hour as requiring 2 hours worth of time on the course for each credit hour of lecture, so you should think of this as the default for all of your courses.

Note - your classes should be more important than your part-time job.

Grading and Assignments

The students of this course have the following responsibilities: read assigned readings before lecture, attend lecture, complete homework assignments, take in-class quizzes, take exams, and complete a project.

Your "total" or "overall" grade will be computed as a weighted sum as below.

There may also some assignments labeled as "checkpoints" - which will mean that if you do not complete them you either cannot pass the class, or cannot get a grade of C (the assignment will specify whether it is a checkpoint for passing or a checkpoint for earning a C). Most assignments will *not* be labeled as checkpoints.

- **Project (15%):** graded based on correctness, proper coding style, completed sub-tasks I asked you to complete, good explanation of it during your presentation.
- **Homeworks and Quizzes (30%):** average of all homework and quiz grades, computed as a weighted average so that assignments worth more points count more. Quizzes cannot be "made up"; homeworks can be completed for 50% late credit.

- **Exams (45%):** 3 exams total, with the exam grade calculated as
  max(.2*exam1 + .3*exam2 + .5*exam3,
  .4*exam2 + .6*exam3,
  exam3)
- **Class Attendance (10%):** computed as .5*lecture + .5*lab attendance. Note that you are required to spend 5 hours per week in the unix lab and logged in and working. For unix lab attendance, one week will be dropped.

CS Course Policies

Note that this course follows all standard CS course policies. In particular check the CS course policies related to - cheating/plagiarism, attendance, missing exams. See http://cs.indstate.edu/info/policies.html for details.

Late Homeworks

For each assignment, when the assignment is handed out it will be indicated if late work will be accepted or not.

Start Homeworks Early

We suggest attempting a homework assignment the day it is given, or the day after, so that if you have a problem you can ask early. If you continue to have problems in trying to complete the assignment, you will have time to ask again. Many of the homework assignments require thought and problem solving, which takes "time on the calendar" not just "time on the clock". By that we mean that spending two hours on 3 consecutive days may be more productive than trying to spend 6 hours at once on the assignment.

Grade Cutoffs

We try to design homework assignments and exams so that a standard cutoff for grades will be close to what you deserve. After the first exam a grade will be created in Blackboard called "Letter Grade" that is what your letter grade would be if the semester ended today. Initially, I will likely assign the following grades: 93-100 A, 90-93 A-, 87-90 B+, 83-87 B, 80-83 B-, 77-80 C+, 73-77 C, 70-73 C-, 67-70 D+, 63-67 D, 60-63 D-, 0-60 F

Our goal is that the different grades have the following rough meaning.

A+/A
You can do *all* the assignments *on your own*.

B+/A-
You understand nearly everything, and should be all set to use this knowledge in other courses or in a job.

B-/B

Most things you understand very well and a few you might not (more towards the former for a B and more towards the latter for a C).

C/C+
Learned enough and have the minimum skills to move on in the subject.

D+/C-
You did put some effort in, and understand many things at a high level, but you haven't mastered the details well enough to be able to use this knowledge in the future.

D-
Students will normally *not* get an F if - you attend 80% of the lectures, complete some of the assignments up through the end of the course, and get nearly half of the problems on the final exam correct.

F
Normally, students that get an F simply stopped doing the required work at some point.

Blackboard

The course has a blackboard site. Click here to go to blackboard. You should see this course listed under your courses for the current term. The blackboard site is only used for giving you your grades (go to the course in blackboard, then click "My Tools", and then "My Grades"). All course content, schedule, etc. is kept in this google doc (which you are currently viewing).

For online courses - if any parts of blackboard are used for the course, mention it here if appropriate.

Academic Integrity

Follow the standard CS course policies in terms of what is and is not allowed on assignments: http://cs.indstate.edu/info/policies.html

Please ask the instructor if you have doubts about what is considered cheating in this course.

Special Needs / Student Disabilities

Standard language included in the syllabi for ISU courses.

Indiana State University recognizes that students with disabilities may have special needs that must be met to give them equal access to college programs and facilities. If you need course adaptations or accommodations because of a disability, please contact us as soon as possible in a confidential setting either after class or in my office. All conversations regarding your disability will be kept in strict confidence. Indiana State University's Student Support Services (SSS) office coordinates services for students with disabilities: documentation of a disability needs to be on file in that office before any accommodations can be provided. Student Support Services is

located on the lower level of Normal Hall in the Center for Student Success and can be contacted at 812-237-2700, or you can visit the ISU website under A-Z, Disability Student Services and submit a Contact Form. Appointments to discuss accommodations with SSS staff members are encouraged.

Once a faculty member is notified by Student Support Services that a student is qualified to receive academic accommodations, a faculty member is obligated to provide or allow a reasonable classroom accommodation under ADA.

Disclosures Regarding Sexual Misconduct

Standard language included in the syllabi for ISU courses.

Indiana State University fosters a campus free of sexual misconduct including sexual harassment, sexual violence, intimate partner violence, and stalking and/or any form of sex or gender discrimination. If you disclose a potential violation of the sexual misconduct policy I will need to notify the Title IX Coordinator. Students who have experienced sexual misconduct are encouraged to contact confidential resources listed below. To make a report or the Title IX Coordinator, visit the Equal Opportunity and Title IX website: http://www.indstate.edu/equalopportunity-titleix/titleix.

The ISU Student Counseling Center – HMSU 7th Floor | 812-237-3939 | www.indstate.edu/cns
The ISU Victim Advocate – Trista Gibbons, trista.gibbons@indstate.edu
HMSU 7th Floor | 812-237-3939 (office) | 812-230-3803 (cell)
Campus Ministries - United Campus Ministries | 812-232-0186
http://www2.indstate.edu/sao/campusinistries.htm
www.unitedcampusministries.org | ucmminister2@gmail.com
321 N 7th St., Terre Haute, IN 47807

For more information on your rights and available resources
http://www.indstate.edu/equalopportunity-titleix/titleix