CS 414-Introduction to Machine Learning

Syllabus Spring-2020

Instructor: Prof. Christian Lopez, 569 Rockwell Integrated Science Center, lopezbec@lafayette.edu
Class: MWF 10:00 AM- 10:50 AM
Office Hours: Mondays & Wednesdays 3:20 PM -4:20PM (or by appointment)
Couse Website: https://sites.lafayette.edu/lopezbec/courses/spring-2020/
Prerequisite: CS202, **co-requisite** MATH 272

Course Description

This course is an introduction to both the theoretical and practical aspects of the design and implementation of algorithms that enable machines to "learn" from examples (i.e., Machine Learning). Instead of programing machines by specifying a set of instructions that dictate exactly how they should perform a task, a new paradigm is developed whereby machines are presented with examples from which they can automatically identify ("learns") suitable rules to perform a task. This allows computer programs to improve their performance on a given task through experience (i.e., more data).

This course will provide students an in-depth introduction to the areas of Supervised and Unsupervised Machine Learning. The course will cover core Machine Learning algorithms for classification, regression, clustering, and dimensionality reduction.

On the theory side, the course will focus on understanding algorithms and the relationships between them. On the applied side, the course will focus on effectively using machine learning methods to solve real-world problems with an emphasis on model selection, regularization, hyperparameter optimization, and presentation and interpretation of results. Specific topics will include linear and logistic regression, support vector machines, artificial neural networks and deep learning, principal component analysis, k-means clustering, decision trees, and random forests.

Student Learning Outcomes

Upon completion of this course, students will be able to:

- Understand the paradigms of supervised and unsupervised machine learning.
- Explain the fundamental issues and challenges of machine learning.
- Identify the strengths and weaknesses of multiple machine learning approaches.
- Formalize a task as a machine learning problem.
- Identify suitable algorithms to tackle different machine learning problems.
- Develop models and apply machine learning frameworks to solve practical problems.

Expectations

Our classroom is a studio class, where sessions will involve a series of activities interwoven with lectures. In some instances, we will adopt a flipped-classroom approach where prior to the face-to-face session, you will gain exposure to new concepts and material by watching lecture videos, doing guided readings, and/or completing the required activities online; and during class you will do active work that calls for the analysis and/or the application of the concepts and material learned prior to class.

There will be classroom discussions, dialogues, and exercises led by the faculty. You are expected to actively participate in these discussions and work on the particular exercises that allow you and your student colleagues to learn by doing, learn by observing the results of others, and to learn from one another while trying out new ideas. You will learn by doing computer programming, interacting with your professor, and other students; and engaging with your professor's instruction as well as external media. You must bring your computer to every class.

What you learn will depend directly on your willingness to participate, be involved, and complete assignments and exercises. Therefore, try things even if you think they might "fail," ask questions—to faculty and to each other, and come to class ready to participate.

Textbook

For this course will use the book:

- Géron, A. (2017). *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, Inc.(ISBN-10: 1491962291).

There is an electronic version of this book (e-book) available through the library: https://ebookcentral-proquest-com.ezproxy.lafayette.edu/lib/lafayettecol-ebooks/detail.action?docID=4822582

There are many good books and articles for Machine Learning. We will cover material from different textbooks that are freely available online, such as:

- Trevor Hastie, Robert Tibshirani and Jerome Friedman (2009). *Elements of Statistical Learning* [https://web.stanford.edu/~hastie/ElemStatLearn/]
- Shai Shalev-Shwartz, Shai Ben-David (2014). *Understanding Machine Learning: From Theory to Algorithms* [https://www.cse.huji.ac.il/~shais/UnderstandingMachineLearning/understanding-machine-learning-theory-algorithms.pdf]
- Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar(2012). *Foundations of Machine Learning* [https://mitpress.ublish.com/ereader/7093/?preview=#page/Cover]
- Hal Daume III (2017). *A Course in Machine Learning* [http://ciml.info/]
- Ian Goodfellow, Yoshua Bengio and Aaron Courville (2016) *Deep Learning* [http://www.deeplearningbook.org/]

Grading Scheme

The grading breakdown of this class is shown below.

Attendance	6 %
Participation	10 %
Quizzes	10 %
Homework	30 %
Midterm	10 %

Final Exam	15 %
Final Project	19 %

Grading Scale
Typically, grades are assigned as follows from your final numerical grade:

A: 93-100	B+: 87-89	C+: 77-79	D+: 67-69	F: 0-59
A-: 90-92	B: 83-86	C: 73-76	D: 63-66	
	B-: 80-82	C-: 70-72	D-: 60-62	

Assignments Types

Attendance: The success of this course is highly dependent on your active participation. As such, only four unexcused absences will be allowed. You will be considered absent if you arrive to class more than 10 minutes late or not at all. The first unexcused absence will not be penalized. Each additional unexcused absence will reduce your *Attendance* grade by 2%. Disciplinary action for students with greater than four absences will be handled on a case by case basis. If you have a dean's excuse or a coach's letter for a particular day, you will be excused from class but not assignments.

Participation: To get a perfect participation grade come to each class prepared and on time, contribute to class discussions at least once a week or more, allow other students to contribute, and maintain a high level of respect and professionalism with your peers. In this course, we will use a Q&A Webservice call Piazza (see Moodle). Any questions about assignments or course material should be posted first in Piazza. This can be done anonymously or not. Class participation includes asking or answering questions (in class or Piazza), expressing an opinion about a topic of discussion, and involvement in group projects or in-class group activities. All sincere efforts to participate are admired, so don't worry, just speak up. You are even welcome to express an opinion different than mine. All types of participation count except participation that shows you failed to prepare for class, did not read the instructions of an assignment, or are disrespectful to others. In addition, each day, a student will give a brief presentation (e.g., 2-3mins) regarding a Machine Learning Tech news. Finally, in order to continuously improve the course and ensure students are making the most out of it, there will be anonymous feedback repositories where students are encouraged to provide constructive feedback regarding the things they liked or would like to improve from the class, activities, lectures, etc.

***Quizzes*:** Some of the material in this class is cumulative, so it is essential to stay current in the class. To help you stay current, there will be quizzes- mostly unannounced. Some of the quizzes will test your knowledge of the previous day or week. The intention here is to encourage you to actively review the material each day.

***Homework*:** As in any other learning endeavor, practice is important while learning computer programming. The homework, as well as the final project, are meant to help you practice and gain proficiency. Homework includes completing the assigned problems. The answers will be provided after the deadline for submitting the homework has passed. Homework also includes assignments that will be distributed occasionally through Moodle. You will receive the points for a problem that you complete in an acceptable way. If you do obtain outside assistance in completing the problem, be sure that you **do NOT copy** the answer. Rather, learn from the outside assistance, and then do the problem on your own. Directly copying an answer would constitute cheating. Similarly, avoid asking/responding to specific questions about an assignment in Piazza (e.g., what is the code for completing X). Instead, ask/respond to questions that help you understand the underlying principles

of the assignment.

Exams: The midterm and final exams will be composed of a written in-class portion and a take-home portion. Exam absences will receive a score of zero for the exam (unless a dean's excuse is given). Before the written portion of the exams begins, you are required to close your course materials and put them and your phone in front of the class to avoid the temptation to look at them during the exam. Each exam will have a self-assessment component to help you evaluate your own learning progress and encourage self-reflection of how to continuously improve.

Final Project: There is a final team project in this class to build an application that integrated a Machine Learning model to solve a practical problem. Teams will be comprised of up to four students. This project will require your team to propose the project to be written, train and test the model, and then deploy it. You will be required to submit one or more intermediate solutions as you develop your full applications. During the last class, you will be asked to demonstrate your game. 6% of the Final Project grade will be based on anonymous peer-evaluation. That is, your team members will evaluate your performance and contributions to the project.

Class Participation Rubric

	Strong Work	**Needs Development**	**Unsatisfactory**
Listening	Actively and respectfully listens to peers and instructor	Sometimes displays lack of interest in comments of others	Projects lack of interest or disrespect for others
Preparation	Arrives fully prepared with all assignments completed, and notes on reading, observations, questions	Sometimes arrives unprepared or with only superficial preparation	Exhibits little evidence of having read or thought about the assigned material
Quality of Contributions	Comments are relevant and reflect an understanding of assigned text(s) or assignments; previous remarks of other students; and insights about assigned materials	Comments sometimes irrelevant, betray lack of preparation, or indicate lack of attention to previous remarks of other students	Comments reflect little understanding of either the assignment or previous remarks in a seminar
Impact on Class	Comments frequently help move the class conversation forward	Comments sometimes advance the conversation, but sometimes do little to move it forward	Comments do not advance the conversation or are actively harmful to it
Frequency of Participation	Actively participates at appropriate times	Sometimes participates but at other times is "tuned out."	Seldom participates and is generally not engaged

Class participation deserving 100% participation grade will be strong in most categories; participation that is strong in some categories but needs development in others will receive an 80%; a grade of 60% reflects a need for development in most categories; 40% of work is typically unsatisfactory in several categories; and 0% of work is unsatisfactory in nearly all categories.

Technology

In this class, students will be introduced to several programs, applications, and frameworks to help their learning journey (e.g., Google Coolab, JupyterHub, Anaconda, TensorFlow). In this course, we will extensively use Python3. Technology in the classroom should enhance the learning environment for all students. Use of technology for purposes defined by the College as academic dishonesty is prohibited. In the event that students receive permission in advance to digitally record a class (audio or video), the material should not be posted to the internet for public access, unless a prior agreement has been made with me. The use of technology in my classes should reflect two key values:

- **That we are here for a common purpose – education**. Use of technology in the classroom by the faculty member and the students should always support student learning. If you are using your phone, tablet, or computer in class, be prepared to show me how you are using the technology to support your learning.

- **That the classroom should be a place of mutual respect**. Students need to respect my efforts to create a classroom environment and to organize the course in ways that support the learning of all students. Students also need to respect their fellow classmates and their classmates' rights not to be distracted from participating fully in the classroom.

Tentative Schedule

This is a tentative schedule, subject to change. Check Moodle for the most up to date information:

Week	Start Date	Topics
1	27,Jan	Introduction to Machine Learning
2	03,Feb	Linear Regression
3	10,Feb	Logistic Regression
4	17,Feb	Regularization
5	24,Feb	Artificial Neural Networks
6	02,Mar	Support Vector Machines
7	09,Mar	Dimensionality Reduction
8	16,Mar	**SPRING BREAK**
9	23,Mar	Training and Testing Machine Learning Models
10	30,Mar	Decision Tress
11	06,Apr	Deep Learning
12	13,Apr	Deep Learning
13	20,Apr	Unsupervised Machine Learning Models
14	27,Apr	Anomaly Detection
15	04,May	Final Project Implementation
		FINAL EXAM TBD

- **Check E-Mail and Moodle Daily.** Information about the class, including assignment updates and schedule changes, will be posted to Moodle and/or sent by e-mail. Not reading your e-mail or checking Moodle will not be accepted as a reason for me to accept a late assignment or your absence in a class activity.

Communication
My preference is for you to address me as either Professor Lopez or Dr. Lopez. If you have a preference regarding how you would like to be addressed, please let me know.

If you need to schedule a meeting or have a request of me that will require time outside of class, please be sure to follow up any conversation we might have about the request immediately before, during, or after class with an e-mail to confirm that I have placed the request on my calendar. Because class time can be busy, by the time I return to my office, there is a chance I will have been distracted and forget our conversation.

Students often worry about how to e-mail a professor. I recommend reading some guidelines/advice at http://web.wellesley.edu/SocialComputing/Netiquette/netiquetteprofessor.html.

Religious Observances:
If you plan to be absent from class due to the observance of a religious holiday, please communicate this to me before the end of the second week of class. You will need to get a Dean's Excuse for religious purposes. This dean's excuse needs to be approved before the "drop/add" deadline (see academic calendar).

Disability:
In accordance with Lafayette College policy, reasonable academic accommodation and support services will be available to students who have a documented disability. Students should review the information available at http://attic.lafayette.edu/disability-services/

Diversity and Inclusiveness Statement
Lafayette College is committed to creating a diverse community: one that is inclusive, responsive, and is supportive of each and all of its faculty, students, and staff. The College seeks to promote diversity in its many manifestations. These include but are not limited to race, ethnicity, socioeconomic status, gender identity and expression, sexual orientation, religion, disability, and place of origin.

The College recognizes that we live in an increasingly interconnected, globalized world, and that students benefit from learning in educational and social contexts, in which there are participants from all manner of backgrounds. The goal is to encourage students to consider diverse experiences and perspectives throughout their lives. All members of the College community share a responsibility for creating, maintaining, and developing a learning environment in which **difference is valued, equity is sought, and inclusiveness is practiced.**

Federal Credit Hour Statement
The student work in this course is in full compliance with the federal definition of a four-credit hour course. Please see Registrar's Office web site (https://registrar.lafayette.edu/wp-content/uploads/sites/193/2013/04/Federal-Credit-Hour-Policy-Web- Statement.doc) for the full policy and practice statement.