



AI, data, and society (syllabus)

Course Information

Please note: This syllabus is from Spring 2020. Spring 2021 will be online/synchronous and there will be modifications to the syllabus.

Instructor

- Professor Prasanna (Sonny) Tambe
- tambe@wharton.upenn.edu, JMHH 558

Course Objectives

The progression of AI-based technologies promises to transform many aspects of business, labor, and even society. The goal of this course is to provide students with an understanding of the capabilities of modern AI technologies, with an emphasis on being able to critically assess where they can provide societal value, and where they may create new societal challenges. The course is not intended to provide a deep-dive into the workings of these technologies in the same way as a computer science course might. Rather, business and policy decision-makers will be confronted with a number of important issues as AI becomes integrated into the social decision-making fabric. This course is intended to provide a framework for people who may have to confront these legal, ethical, and economic challenges. In doing so, a course objective is to ensure that students who complete the course are comfortable enough in the inner-workings of these technologies to think critically across many AI contexts as well as different

domains ranging from public policy, to criminal justice, to health inspections, HR, and marketing.

The **0.5 CU** course is oriented around hands-on critical written assessments, labs, exams, and a presentation. Broadly, data rich firms in finance, tech, management, marketing, and other industries are increasingly adopting AI as a tool to accelerate and improve decision-making. It is important for modern managers to understand the opportunities and challenges introduced by data and AI so that they can credibly communicate about these issues with others in the firm. We will cover many of these issues, so that you will be able to think about the opportunities and challenges that arise when firms try to use AI to solve business problems.

Labs will reinforce your learning of how AI works, and how it is being used to solve business problems. We will focus in the labs on gaining experience with introductory machine learning concepts. Students will spend time inside and outside of the classroom combining data and code to provide a foundation for understanding the deep challenges that this will bring to organizations.

Course Overview

Within the last decade, there has been a dramatic rise in interest in the use of AI technologies applied to many domains, including finance, HR, policy, marketing, and strategy. As a result, the modern “digital leader” increasingly requires the use of technology, statistics, and data analysis skills to facilitate business analysis as well as a grasp of where AI technologies can perform well, and where they may fail. This includes knowing how to a) effectively frame data-driven questions, b) use AI-driven algorithms, and c) understand the capabilities of this new generation of tools that are becoming available to automate decision-making from data.

The class is a seminar style class, which includes readings and critical assessments due before class and in-class discussions of topics related to the application of AI technologies, including ethics, bias, and the potential for AI to fuel gains in productivity. The class also requires the completion of two labs to ensure that participants have a deep understanding of how algorithms are applied for decision-making and what the constraints are of these approach. A

background in coding is *not* required in this class. Labs will be based in "no-code" tools such as MS Azure ML Studio.

Throughout the semester, we will cover the applications of AI to a number of industries (e.g. medicine, journalism, criminal justice), including inviting guest speakers to hear about applications in specialized domains. A learning goal of this course is exposure to how AI is changing decision-making in different industry contexts, and how organizations are reacting to these changes.

Class time in most weeks will be dedicated to lecture and discussion of some of the most important topics facing the AI community. Class time in some weeks will be devoted to supervised work on projects that are meant to underscore AI-based challenges. Through these exercises and discussions, students are expected to become proficient at applying data to business decisions and at effectively analyzing big data sets to inform decisions about business problems using data analysis tools.

Course web site

We will be using Canvas to submit all assignments and receive grades. All course information will be posted on the Canvas website.

Required textbooks and software

There is no textbook. Readings will consist of selected online content which will be posted on Canvas. As part of your homework, you may be expected to view online videos or complete some online courses that introduce tools such Machine Learning Studio.

Deliverables and grading

During this course, you will be assigned a number of hands on data projects which you will spend time on both in class and out of class. You are expected to participate in classroom discussions (there is more information about participation below). The breakdown of points is as follows:

Copy of Grading rubric

<u>Aa</u> Name	<u>#</u> Percentage points
<u>AI Labs</u>	20
<u>Writeups</u>	20
<u>Exam</u>	30
<u>Professionalism + Participation</u>	15
<u>Capstone</u>	15

As mentioned above, a background in coding is *not* required in this class. Labs will be based in "no-code" tools such as MS Azure ML Studio. In corporate America, you will be expected to present your analytic findings and make a recommendation. Therefore deliverables may include short, informal analyses and an accompanying recommendation.

Group projects will be completed in small groups (two to three students, no more than three). You may also be asked to evaluate the contribution of each of your team members after the group project.

The classroom presentation and discussion presents a unique opportunity for you to develop and enhance your confidence and skills in articulating a personal position, sharing your knowledge, and reacting to new ideas. All of you have personal experience that can enhance our understanding of this subject, and we want to encourage you to share that experience.

Participation and Professionalism

This course, like many other courses at Wharton, uses learning methods that require active involvement (e.g. attendance, participation in discussions, and in-class exercises). Not only is this the best way to learn, but it also develops your communication and presentation skills. Regular attendance, participation, presentations, and in general, presenting yourself professionally are all very important, and are an important part of your grade. Active participation requires good preparation—thoughtful completion of homework before class is essential. We recognize that expressing viewpoints in a group is difficult, but it is an important skill for you to develop. We will do what we can to make this as easy as possible. Remember though that only regular and insightful contributions will be rewarded.

The grade we assign for your class participation and attendance is a careful, subjective assessment of the value of your input to classroom learning. We keep careful track of attendance, your contributions towards each class session, and these contributions can include (but are not restricted to) raising questions that make your classmates think, providing imaginative yet relevant analysis of a situation, contributing background or a perspective on a classroom topic that enhances its discussion, providing thoughtful feedback on the presentations of other students, and simply answering questions raised in class. A lack of preparation, negative classroom comments, or improper behavior (such as talking to each other, sleeping in the classroom or walking in and out of the class while the lecture is in process) can lower this grade.

In particular, because this class emphasizes discussion, attendance is expected. Missing an excessive amount of class without justification will negatively impact your final grade.

Grading Guidelines

At Wharton, we strive to create courses that challenge students intellectually and that meet the Wharton standards of academic excellence. If you believe that an assignment or project grade you received was unjustified, you can appeal the grade. To appeal the grade you must write a one-page explanation as to the reason for your appeal and hand it in along with your graded assignment.

Please think twice before appealing a grade: the entire assignment will be regraded, which may increase your grade, but may also lower it.

Points will be deducted from late assignments, labs, or projects at the rate of a 20% penalty for each day the submission is late. There are no exceptions except for the case of documented medical emergencies.

Accounts for Non-Wharton students

If you are a non-Wharton student and need to access public computers, you will need to create a Wharton Class Account. (Once created a class account will then link to their PennKey account and allow you to log into public computers with their PennKey accounts.)

For class account creation, please see:
<https://apps.wharton.upenn.edu/iam/accountcreator/>.

Course Schedule

Session list

# Session	Aa Name	☰ Deliverables
1	<u>Introduction + A brief history of AI</u>	
2	<u>How AI works I (Algorithms)</u>	
3	<u>How AI works II (Interpreting output)</u>	
4	<u>LAB 1: machine learning on azure</u>	Writeup #1: Bias
5	<u>Bias, ethics, fairness, and the law (Guest)</u>	
6	<u>AI for social good (Guest)</u>	
7	<u>Regulation: AI, privacy, and market power</u>	Writeup #2: Regulation
8	<u>LAB 2: deep learning</u>	Lab 1
9	<u>AI and UAV's (Guest)</u>	
10	<u>The interpretability challenge</u>	Lab 2
11	<u>AI and work I: the data driven workforce</u>	
12	<u>AI and work II: The future of jobs</u>	Writeup #3: Jobs
13	<u>In-class exam</u>	
14	<u>Group presentations/capstone</u>	In-class presentations

Session 1: Introduction + A brief history of AI

Objective: This session introduces expectations for the course. It covers the evolution of the AI, including varying viewpoints, from the 1920's to the present. Includes a discussion of the Turing test, the AI winter, the AI spring and the role of Moore's law and data in driving the resurgence of AI. Most of the time is spent on discussion of AGI and AI technological progress over the last five years.

Required readings:

- Can machines think?

- [How far are we from AGI?](#)
 - [The Future of AI](#)
-

Session 2: How AI Works I (Algorithms)

Objective: Introduction of ML techniques, with comparisons with pattern recognition and regression. Introduction to the role of training and test data sets in the ML workflow and how prediction works in the ML context. This session covers how to think about different popular ML algorithms (logistic regression, trees, etc.), and when to use each of them. Discussion of when different models work best, and what the "requirements" are for each to be successful. Coverage of popular platforms used for machine learning (Pytorch, Tensorflow, caret + Azure and AWS tools).

Required readings:

- [A Visual Introduction to Machine Learning](#)
 - [About Train, Validation and Test Sets in Machine Learning](#)
 - [Machine Learning Algorithm Cheat Sheet](#)
-

Session 3: How AI Works II (Interpreting output)

Objective: Discussion of how to evaluate AI models and further discussion of the costs and benefits of different AI models. Explanation of ROC curves, Type I and type II errors, and confusion matrices. Discussion of how the need for well labeled output limits which tasks are amenable to machine learning/automation. Will also cover platforms such as Kaggle, and how ML competitions are "specified", including prediction quality measures.

Required readings:

- TBD
-

Session 4: Lab 1 - ML on Azure

Objective: Exercise using Azure ML studio with HR data. Uses a drag and drop interface to code variables and predict HR outcomes. Assignment underscores tradeoffs between model accuracy and features. There is also a discussion in the lab of tradeoffs between data features and privacy. Finally, a brief discussion of bias, post-hoc fixes, and legal challenges.

Required readings:

- [AI in Human Resources Management: Challenges and a Path Forward](#)
- [What is Azure Machine Learning Studio?](#)

Session 5: Bias, fairness, and the law

Objective: Covers bias, fairness, and the law in the application of ML algorithms. Included discussions of tradeoffs between accuracy and bias. Extensions to specific applications like medicine and public policy.

Required readings:

- [Goel et al. "The Cost of Fairness"](#)
- [Huq, A. "Discriminatory Intent"](#)
- [What do we do about bias in AI?](#)

Optional reading:

- [These startups are building tools to keep an eye on AI](#)
- Skim [some relevant startups in this space](#)

Session 6: AI for social good

Objective: Covers the use of AI for social good (e.g. possible use cases to cover include counterterrorism, economic empowerment, preventing wildlife poaching). Likely with a guest speaker from DataKind or similar organization. Includes discussion of "citizen" data science.

Required readings:

[Applying artificial intelligence for social good](#)

Session 7: Regulation: AI, privacy, and market power

Objective: Cover returns to data scale, discussion of tradeoffs with privacy. How GDPR may affect AI innovation. Discussion of Alibaba, Baidu, and Tencent and the "dual axes" of AI power. Current regulation initiatives such as the Michigan Law prohibiting the use of video for HR decisions.

Required readings:

- [Executive Order on AI](#)
 - [Should AI be regulated?](#)
 - [Predictive Modeling with Big Data: Is Bigger really better?](#)
-

Session 8: Lab 2, Data and Deep Learning

Objective: Introduces students to neural networks and deep learning. This lab contrasts the benefits of hand-generating ML features vs. using a deep learning engine. Evaluates the predictive gains from the latter and discusses the challenges that deep learning offers for interpretability.

Session 9: AI and healthcare (Guest)

Objective: Discusses the specific challenges AI faces in healthcare. Includes issues with HIPAA and the bias issues that emerge when integrating AI into healthcare recommendations.

Required readings:

- [Computer algorithm fuels racial bias in US healthcare](#)
-

Session 10: The interpretability challenge

Objective: Discusses problems of responsible AI. Explainable AI. What are the trade offs in organizational adoption? What new technologies are emerging to deal with the responsible AI problem.

Required readings:

- [Artificial intelligence must explain itself](#)
 - [Leading your organization to Responsible AI](#)
-

Session 11: AI and labor I: the data driven workforce

Objective:

Discusses how the workforce and organizations are changing to adapt to AI. How new tools such as AutoML change the skill profile of the workforce in an AI-driven organization. Also discusses what types of managers are needed to govern AI driven decisions, including the emergence of new governance vehicles, such as AI councils and Chief Data Officers.

Required readings:

- [The Risks of AutoML and how to avoid them](#)
 - [What do machine learning practitioners actually do?](#)
 - [An Opinionated Introduction to AutoML and Neural Architecture Search](#)
 - [Google's AutoML: Cutting through the hype](#)
 - [Algorithms need managers too](#)
-

Session 12: AI and labor II: The future of jobs

Objectives: Discussion of the impact of AI on the future of work. Includes job displacement, robots, universal basic income, and other related topics. Includes a scenario planning exercise to discuss possible long-term outcomes for the relationship between AI, capital, and labor. Also discusses emerging AI technologies and how they might be combined with the blockchain and AR and other technologies to further affect workforce needs.

Reading:

- [How will machine learning transform the labor market?](#)
 - Excerpt from Race against the Machine or related HBR article
 - [Acemoglu and Restrepo](#)
-

Session 13: In-class exam

Session 14: Group presentations
