



**190F**  
Fall 2018

# Foundations of Data Science

## Lecture 1

---

Introduction

# Welcome to 190F: Foundations of Data Science

---



Prof. Ben Marlin  
Instructor  
LGRC A355



Prof. Pat Flaherty  
Instructor  
LGRT 1444



Mahmood Jasim  
Head TA

---

# **Data Science**

# What is Data Science?

---

Drawing useful conclusions from data using computation

- **Exploration**

- Identifying patterns in information
- Uses visualizations

- **Inference**

- Quantifying whether those patterns are reliable
- Uses randomization

- **Prediction**

- Making informed guesses
  - Uses machine learning
-

# Applications

---

- Data science is driven by applications
  - Data analysis is playing an increasingly important role in many fields including Biology, Chemistry, Economics, Earth Systems, Education, Environmental Science, Finance, Geography, Geology, Kinesiology, Linguistics, Management, Political Science, Public Health, Psychology, Sociology, ...
  - Every data-driven subject brings new challenges
-

# Examples

## In fight against fake news, technology outsmarts humans at detecting misinformation

Researchers have demonstrated an algorithm solution that is comparable to and sometimes correctly identifying fake news stories.

By: IANS | New York | Published: August 22, 2018 11:23 AM

0  
SHARES

f SHARE



Zach Gibson / Getty

## Inside Facebook's efforts to protect the U.S. election

POLITICS

Is it enough?

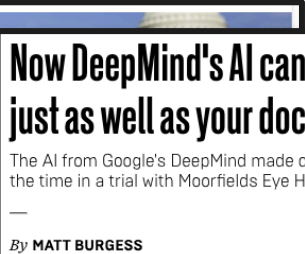
By Kurt Wagner | August 22, 2018



SHARE



Two weeks ago, on a conference call with executives announcing



## Now DeepMind's AI can spot eye diseases just as well as your doctor

The AI from Google's DeepMind made correct diagnoses in a trial with Moorfields Eye Hospital.

By MATT BURGESS

13 Aug 2018



PHYSICS

## LHC Physicists Embrace Brute-Force Approach to Particle Hunt

The world's most powerful particle collider has yet to turn up new physics—now some physicists are turning to a different strategy

By Davide Castelvecchi, Nature magazine on August 15, 2018

# **Course Structure**

# What does the course cover?

---

- An introduction to programming in Python with a focus on manipulating, visualizing, and analyzing data.
  - An introduction to statistics that is grounded in computer simulations.
  - An introduction to predictive modeling and machine learning.
-



# Course Components and Grading

---

- Lectures M/W/F
- Weekly labs
- Weekly homework assignments
- Evening Midterm exam & final exam

Homework	35%
Labs	15%
Midterm Exam	25%
Final Exam	25%

---

# Course Technology

---

- **Moodle:** Online gradebook
  - **Piazza:** Online discussion forums, course Q&A, announcements, and instructor DM.
  - **Github.io:** Course website (lecture slides, demos, assignments, labs, etc.)
  - **DataHub:** Web-based Python compute environment for completing labs and homework assignments.
  - **Links to all resource can be found on Moodle.**
-

# Course Policies

---

- Late Homework
  - Re-grades
  - Academic Honesty
  - <https://umass-data-science.github.io/190fwebsite/policies/>
-

# Collaboration Policy

---

## Asking questions is highly encouraged

- You can discuss homework and lab questions with each other
- Do not take notes or pictures out of discussions
- The work you turn in must be your own

## The Limits of collaboration

- Don't share solution material of any type with each other
  - Copying solutions from any source will be dealt with under UMass' Academic Honesty procedures: <https://www.umass.edu/honesty/>
-

# Getting Help

---

- The course staff are here to help you be successful in the course!
  - When you need help come to office hours or post on the Piazza discussion forums.
  - The lab sessions are also a good time to ask questions and get help.
-

**Let's Dive In!**

# In fight against fake news, technology outsmarts humans at detecting misinformation

Researchers have demonstrated an algorithm-based automated solution that is comparable to and sometimes better than humans at correctly identifying fake news stories.

By: IANS | New York | Published: August 22, 2018 11:23 AM

0  
SHARES

