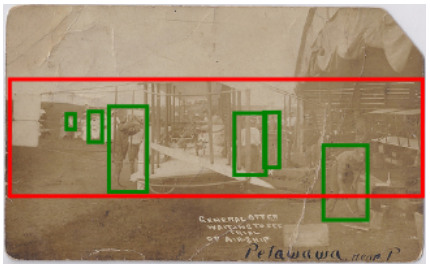
# Computational Learning Theory - MT 2018
# Introduction and Course Details

Varun Kanade

University of Oxford
October 8, 2018

# What is Machine Learning?






- 200-basic level categories
- Here: Six pictures containing airplanes and people
- Dataset contains hundres of thousands of images
- Imagenet competition (2010-)
- All recent successes through very deep neural networks!

# What is Machine Learning?

| Movie / User | Alice | Bob | Charlie | Dean | Eve |
|---|---|---|---|---|---|
| The Shawshank Redemption | 7 | 9 | 9 | 5 | 2 |
| The Godfather | 3 | ? | 10 | 4 | 3 |
| The Dark Knight | 5 | 9 | ? | 6 | ? |
| Pulp Fiction | ? | 5 | ? | ? | 10 |
| Schindler's List | ? | 6 | ? | 9 | ? |

Netflix competition to predict user-ratings (2008-09)

Any individual user will not have used most products

Most products will have been used by some individual

# Supervised Learning

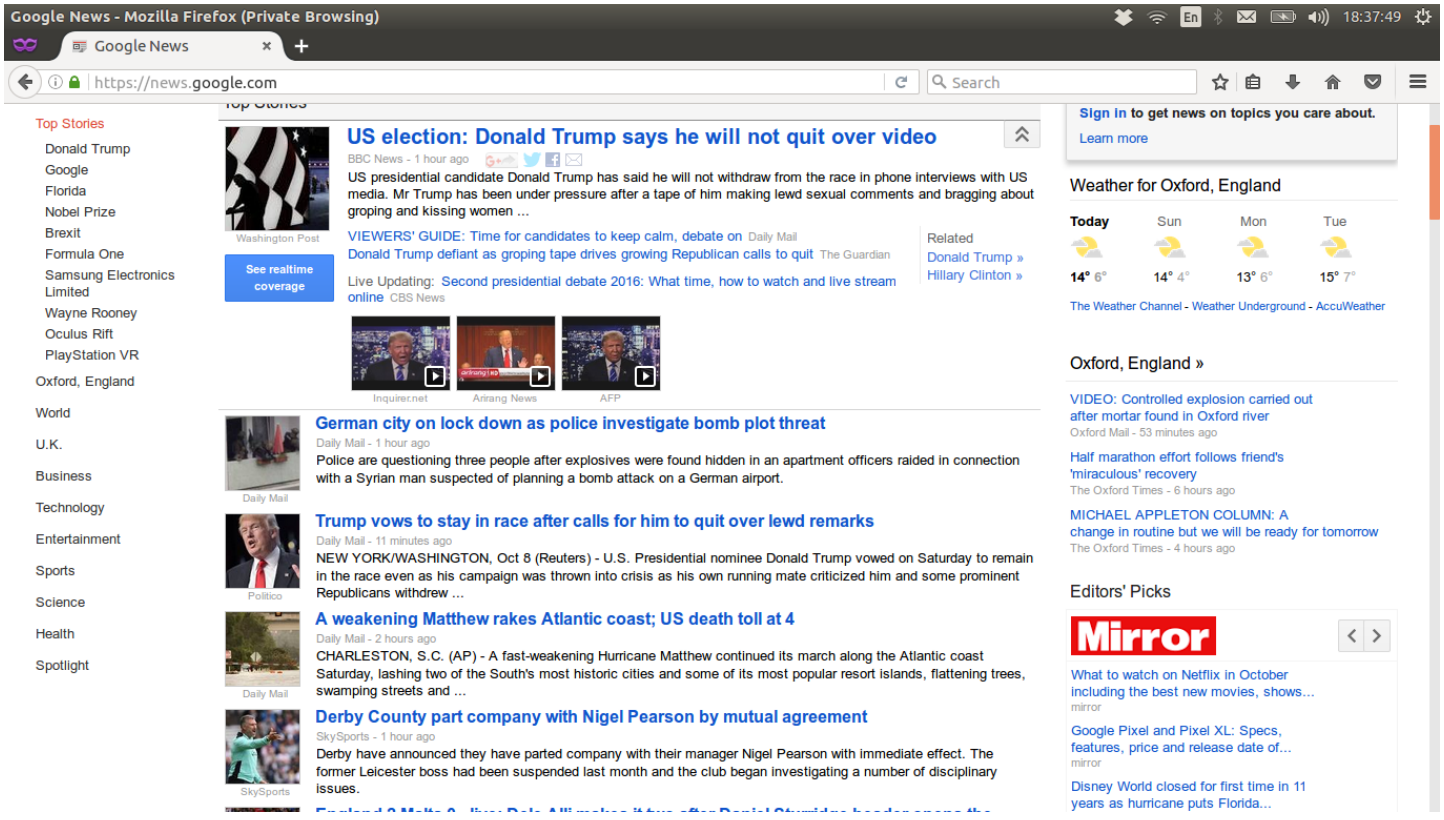
Training data has inputs $\mathbf{x}$ (numerical, categorical) as well as outputs $y$ (target)

Regression: When the output is real-valued, *e.g.,* housing price

Classification: Output is a category

- Binary classification: only two classes *e.g.,* spam
- Multi-class classification: several classes *e.g.,* object detection

# Unsupervised Learning : Group Similar News Articles




Group similar articles into categories such as politics, music, sport, *etc.*

In the dataset, there are no labels for the articles

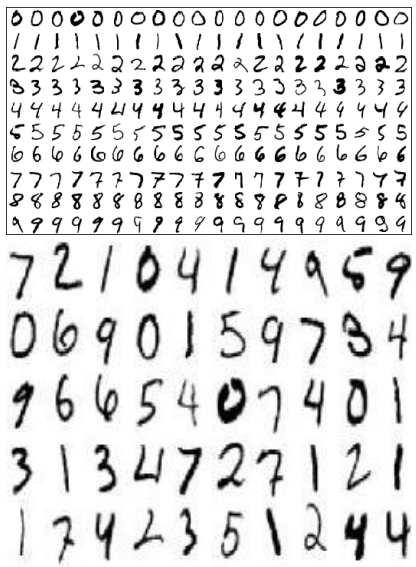
# Active and Semi-Supervised Learning

## Active Learning

- Initially all data is unlabelled
- Learning algorithm can ask a human to label some data

## Semi-supervised Learning

- Limited labelled data, lots of unlabelled data
- How to use the two together to improve learning?



# Outline

What is Machine Learning?

What is Learning Theory?

Course Logistics

# What is Learning Theory?

The goal of (computational) learning theory is to develop formal models to analyse questions arising in machine learning

- How much data do we need to learn?
- What amount of computational resources are necessary for learning?
- Are there hard learning problems?

# What is Learning Theory?

In this course we'll cover several models that aim to capture questions that are of interest in modern machine learning

- (How) can we learn in the presence of noisy data?
- What can we learn when data is obtained in an online manner?
- (How) can we do useful machine learning while preserving privacy?
- Can we learn when data and computational power is distributed?

# What is Learning Theory?

Towards the end of the course we'll cover some of the latest topics in the area

- Can we develop a theoretical understanding of neural networks?
- Connections to information theory, game theory, etc.
- Conference on Learning Theory (COLT)

# Outline

What is Machine Learning?

What is Learning Theory?

Course Logistics

# Course Information

## Website

`www.cs.ox.ac.uk/people/varun.kanade/teaching/CLT-MT2018/`

## Lectures

Mon 15h-16h, Thu 16h-18h (about 20 contact hours)

## Classes

Weeks 3-7

Instructors: Alexandros Hollender, Philip Lazos, Francisco Marmolejo, David Martínez

## Office Hours

After Monday lecture (16-17h)

# Course Information

## Textbooks

- Kearns and Vazirani - An Introduction to Computational Learning Theory
- Several additional texts for suggested reading on website
- Papers and (rough) lecture notes will be posted

## Assessment

- Take Home Exam

## Piazza

- Use for course-related queries
- Sign-up at `piazza.com/ox.ac.uk/other/cltmt2018`

# Who should take this course?

In this course, we will cover

- Mathematical formulations for different learning paradigms
- Definitions, theorems, proofs
- Design and analysis of learning algorithms
- Provable guarantees on run-time and sample complexity

# Who should take this course?

In this course, we will cover

- Mathematical formulations for different learning paradigms
- Definitions, theorems, proofs
- Design and analysis of learning algorithms
- Provable guarantees on run-time and sample complexity

In this course, we will not cover

- Practical applications of learning algorithms - although understanding the theory will likely make you a better practitioner

# Who should take this course?

It is expected that you will be familiar with most of the following

- The notion polynomial time, space, etc.
- Big O notation
- Basic probability theory - expectation, independence, etc.

# Who should take this course?

It is expected that you will be familiar with most of the following

- The notion polynomial time, space, etc.
- Big O notation
- Basic probability theory - expectation, independence, etc.

It'd be helpful if (though not necessary that) you've seen at least some of the following

- Basic complexity theory such as NP-completeness
- Applied Machine Learning
- Optimisation algorithms - Linear Programming

# Who should take this course?

This is an advanced theoretical course. If you are taking this course, you should

- Be keen to understand the theory behind machine learning algorithms
- Be able to fill in details of algorithms and proofs omitted in the lectures
- Develop an ability to read research papers