Class Starting at 14:10

This session will be recorded

Introductory Applied Machine Learning

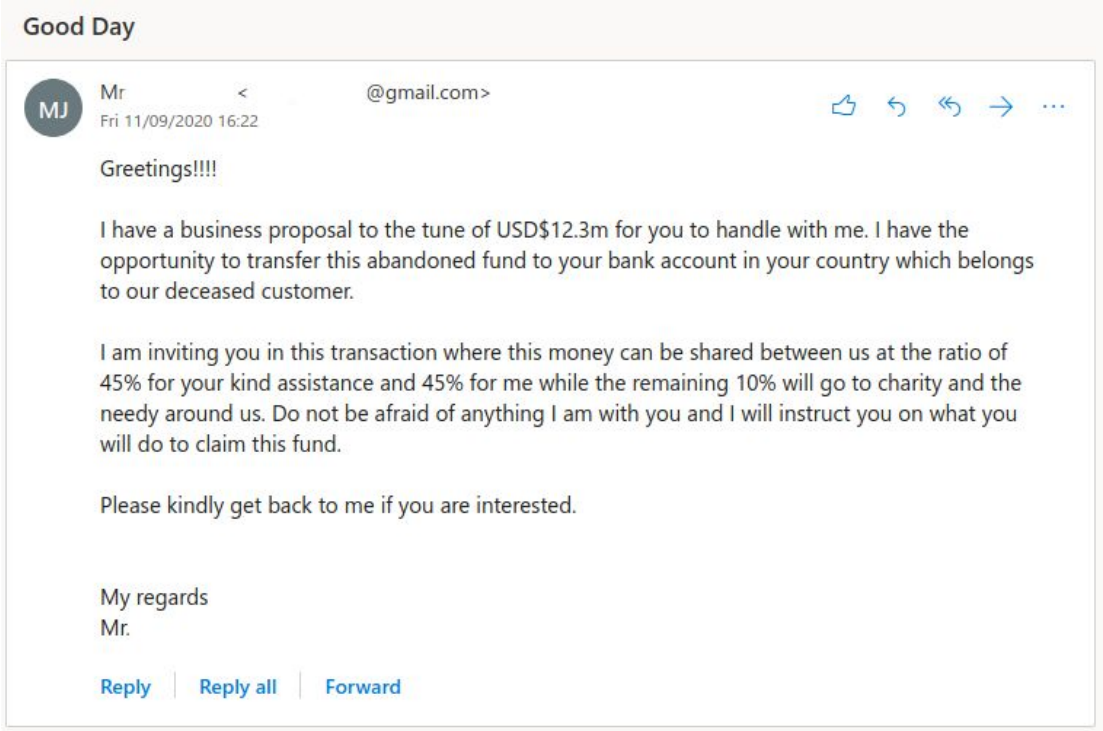
Week 1 Class 1: Introduction

Oisin Mac Aodha

Course Instructors

+ a big team of TAs, demonstrators, and tutors helping out

How to Make a Spam Filter?



What is Machine Learning?

Machine Learning (ML) is the study and development of algorithms that learn from data in order to make predictions about new data.

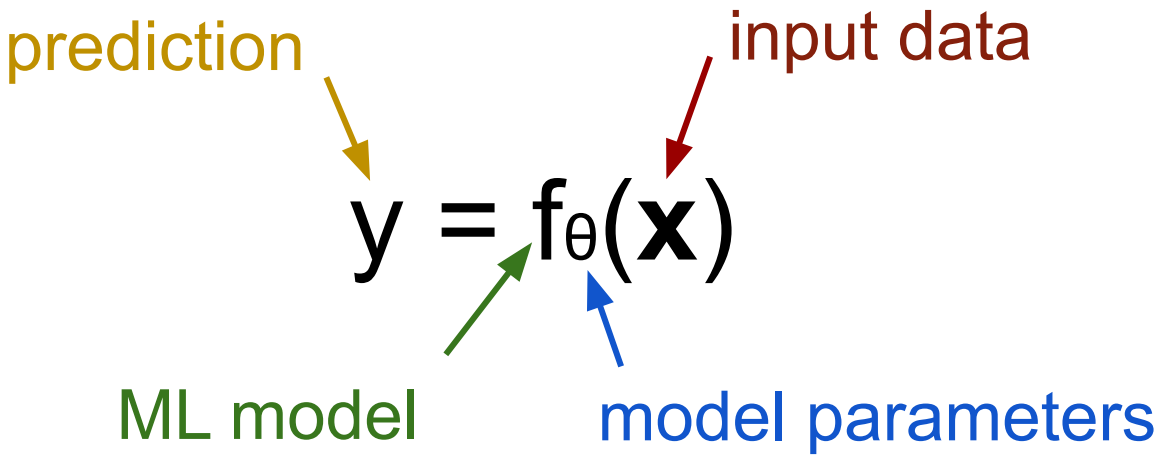
What is Machine Learning?

Machine Learning (ML) is the study and development of algorithms that learn from data in order to make predictions about new data.

$$y = f_\theta(\mathbf{x})$$

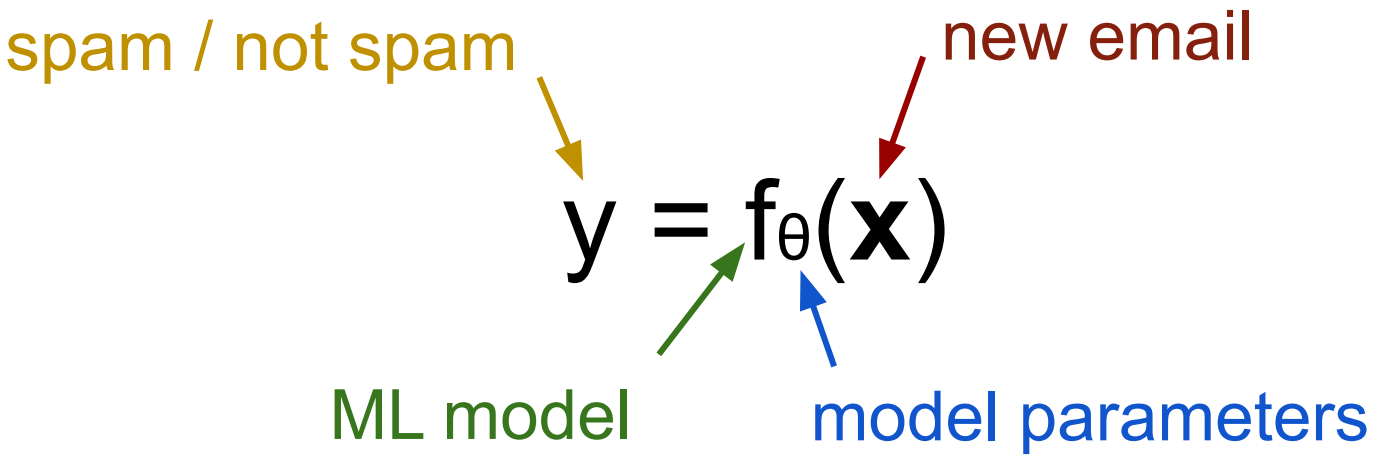
What is Machine Learning?

Machine Learning (ML) is the study and development of algorithms that learn from data in order to make predictions about new data.



What is Machine Learning?

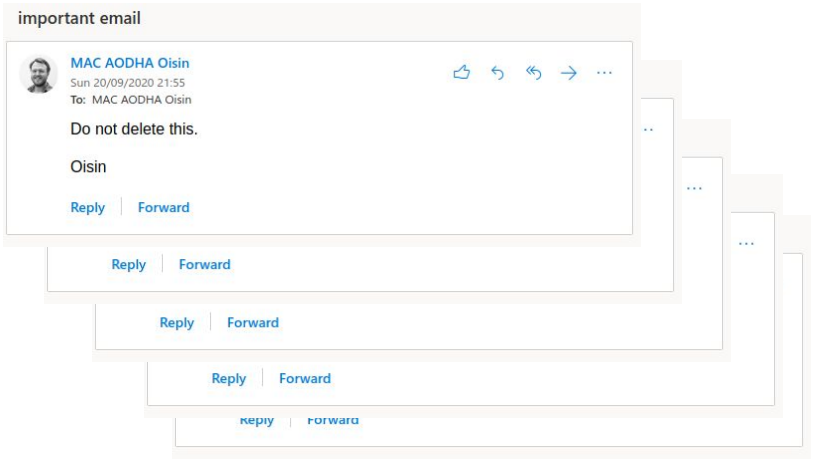
Machine Learning (ML) is the study and development of algorithms that learn from data in order to make predictions about new data.



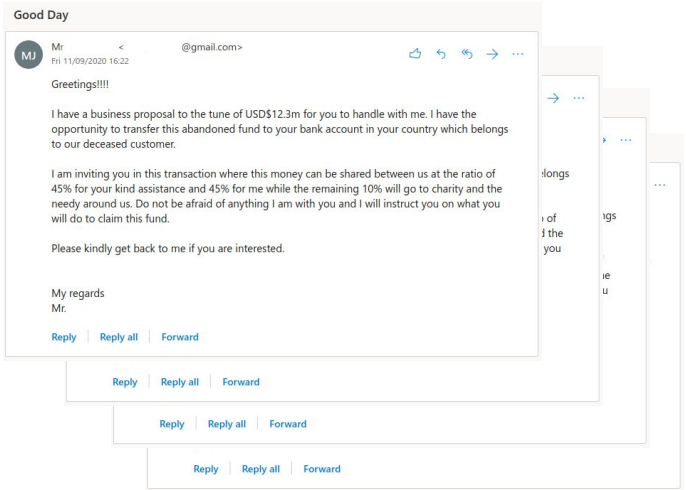
How to Estimate the Model Parameters?

We will use data to optimise for the model parameters that make the fewest mistakes on previously seen data

Not spam emails



Spam emails



Core questions

What is the task I am trying to solve?

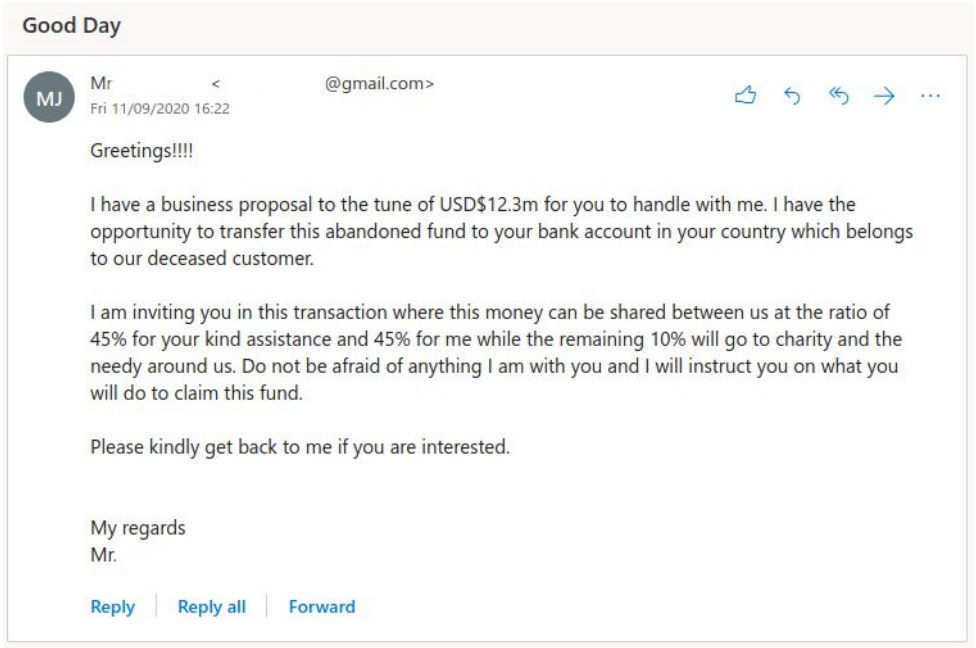
How should I measure performance?

What model, $\mathbf{f}$, should I use?

How can I estimate the parameters, $\boldsymbol{\theta}$, of my model, $\mathbf{f}$?

How should I represent my data to create features, $\mathbf{x}$?

Example ML Tasks - Classification



Is this email spam?

x contains “features” encoding the email
y is a discrete class label - integer

e.g. 0 = not spam, 1 = spam

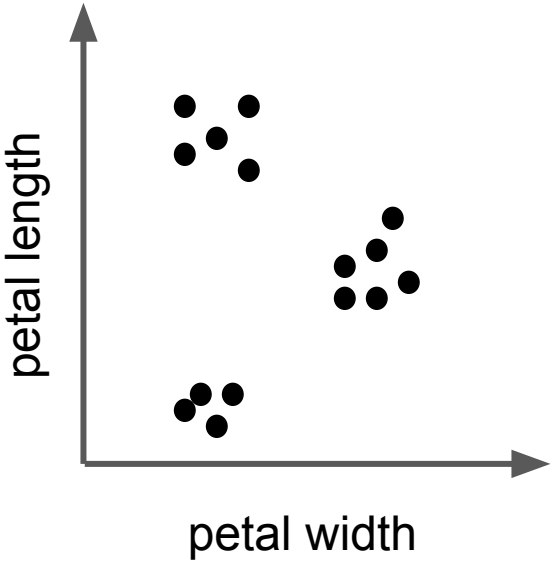
Example ML Tasks - Classification

What species is this?

x contains “features” encoding the image
y is a discrete class label - integer

e.g. 0 = crow, 1 = starling, 2 = sparrow, ...

Example ML Tasks - Clustering

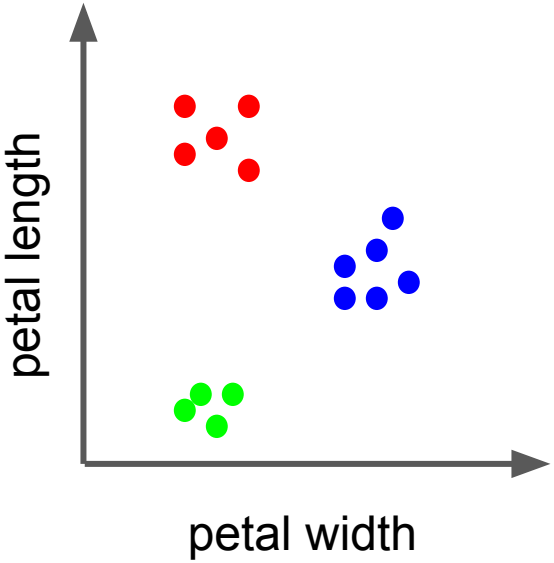


Grouping items, where we do not know the groups in advance

Find structure in the data

x contains the real world features/measurements
y is the id of the group - integer

Example ML Tasks - Clustering



Grouping items, where we do not know the groups in advance

Find structure in the data

x contains the real world features/measurements
y is the id of the group - integer

Example ML Tasks - Regression

How much will this house sell for?

x contains features encoding house size, location, ...
y is the price - a continuous value

Image from https://en.wikipedia.org/wiki/Holyrood_Palace

Question

You have been tasked with creating a model that can predict a student's final grade, between 0 and 100, in a given subject. As input you have information about their attendance, coursework performance, etc.

You also have access to data from last year's students to estimate the parameters of your model.

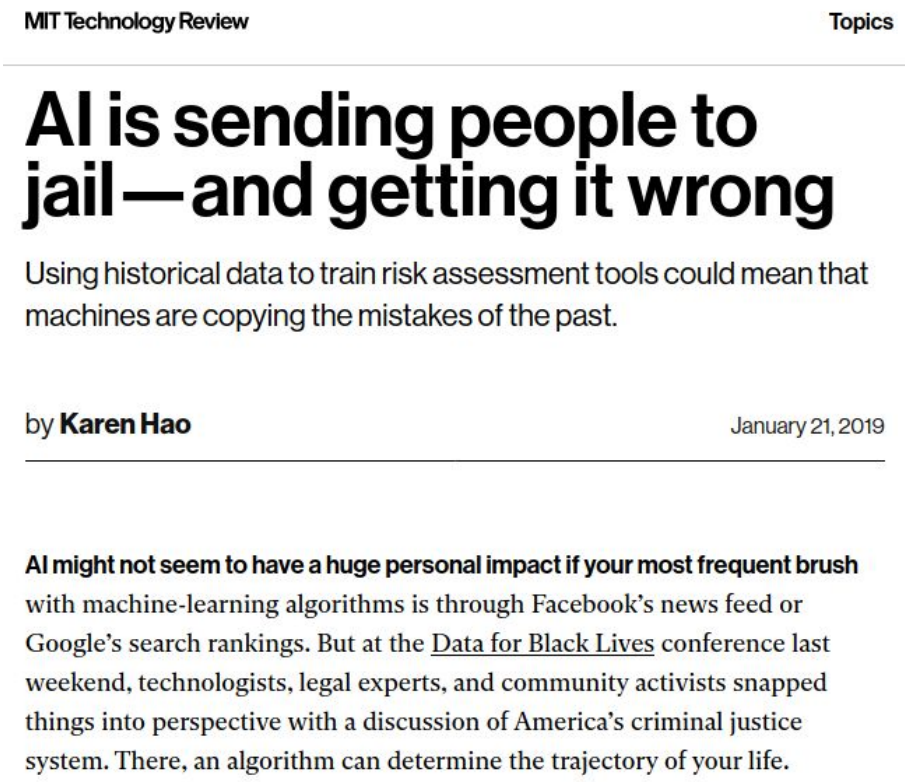
Q: Should you pose this as a classification, clustering, or regression task?

With Great Power ...

Machine learning is a powerful tool that can be used for good or bad

Some of the bad stems from

- incompetence / lack of understanding
- not caring
- ignoring biases in data
- blindly trusting automated systems
- ...



MIT Technology Review Topics

AI is sending people to jail—and getting it wrong

Using historical data to train risk assessment tools could mean that machines are copying the mistakes of the past.

by **Karen Hao** January 21, 2019

AI might not seem to have a huge personal impact if your most frequent brush with machine-learning algorithms is through Facebook's news feed or Google's search rankings. But at the Data for Black Lives conference last weekend, technologists, legal experts, and community activists snapped things into perspective with a discussion of America's criminal justice system. There, an algorithm can determine the trajectory of your life.

https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/

What is IAML?

An introduction to a number of machine learning methods and concepts

Teaches you how these algorithms work

Gives you the tools to know which algorithm to apply for a given task and how to evaluate it

Practical experience of applying machine learning to real problems

Gives you a foundation for learning about more advanced machine learning algorithms

What is IAML not?

IAML is not an advanced Machine Learning course
Instead try **Machine Learning and Pattern Recognition (MLPR)**
Semester 1

IAML is not a Deep Learning course
Instead try **Machine Learning Practical (MLP)**
Semesters 1 and 2

INFR10069 vs INFR11182

INFR10069

Level 10

Typically Undergraduate students

INFR11182

Level 11

Typically MSc students

Note:
Lectures, classes, tutorials, and labs are the same for L10 and L11

Make sure you are signed up for the correct course

Course Structure

Week	Topics	Lab	Tutorial
1	Introduction (no pre-class video) Mathematical Preliminaries	Lab 0: Introduction to Python and ML packages	
2	Thinking About Data Naive Bayes	Lab 1: Data Analysis and Visualisation and Multinomial Naive Bayes	
3	Decision Trees Generalisation and Evaluation		Tutorial 1: Naive Bayes and Feature Engineering
4	Linear Regression Logistic Regression	Lab 2: Decision Trees and Linear Regression	
5	Optimisation and Regularisation Support Vector Machines I		Tutorial 2: Decision Trees and Gaussian Naive Bayes
6	Support Vector Machines II Nearest Neighbours	Lab 3: SVMs, Evaluation	
7	K-Means Gaussian Mixture Models		Tutorial 3: Logistic Regression
8	Principal Components Analysis Hierarchical Clustering	Lab 4: Clustering, PCA and Evaluation	
9	Perceptrons, Neural Networks		Tutorial 4: PCA, Clustering and Evaluation

Course Structure

Introduction

Supervised

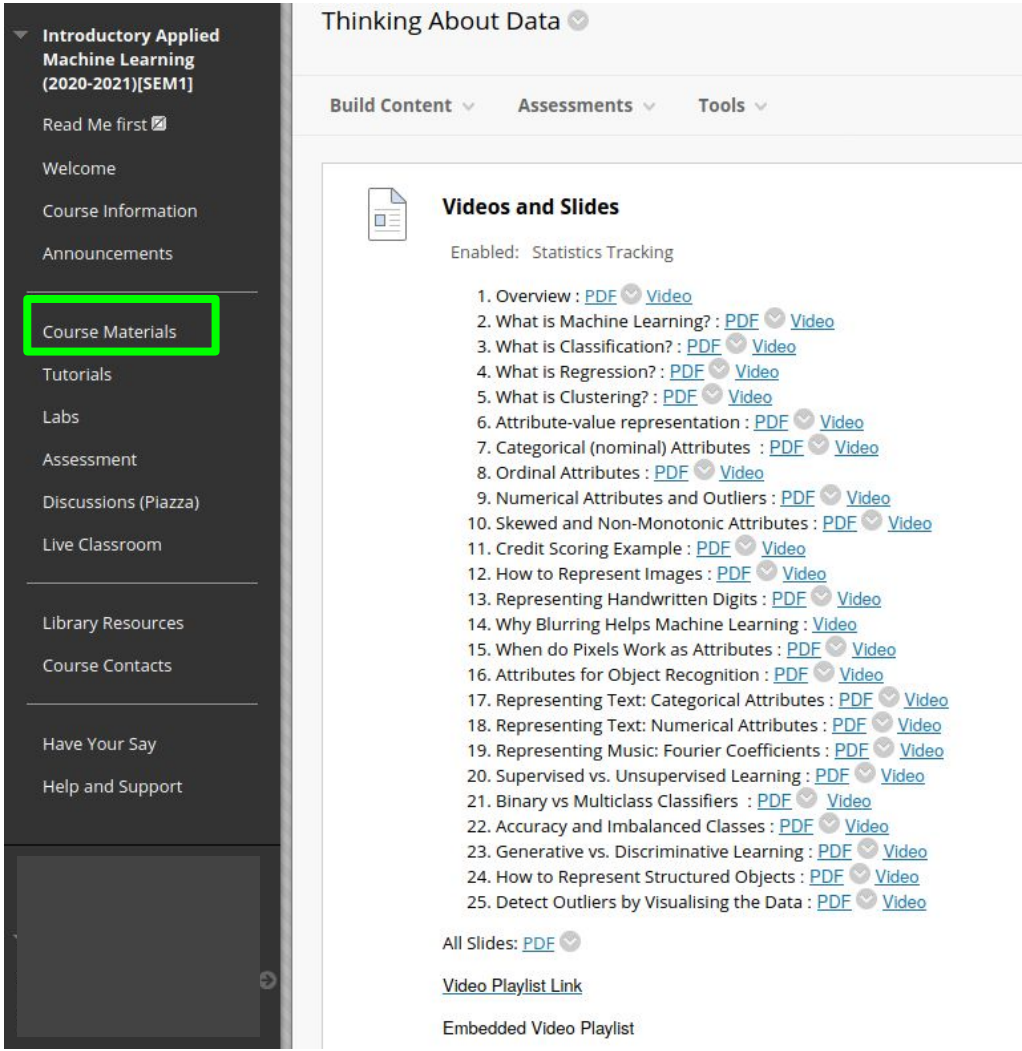
Unsupervised

Neural Networks

Week	Topics	Lab	Tutorial
1	Introduction (no pre-class video) Mathematical Preliminaries	Lab 0: Introduction to Python and ML packages	
2	Thinking About Data Naive Bayes	Lab 1: Data Analysis and Visualisation and Multinomial Naive Bayes	
3	Decision Trees Generalisation and Evaluation		Tutorial 1: Naive Bayes and Feature Engineering
4	Linear Regression Logistic Regression	Lab 2: Decision Trees and Linear Regression	
5	Optimisation and Regularisation Support Vector Machines I		Tutorial 2: Decision Trees and Gaussian Naive Bayes
6	Support Vector Machines II Nearest Neighbours	Lab 3: SVMs, Evaluation	
7	K-Means Gaussian Mixture Models		Tutorial 3: Logistic Regression
8	Principal Components Analysis Hierarchical Clustering	Lab 4: Clustering, PCA and Evaluation	
9	Perceptrons, Neural Networks		Tutorial 4: PCA, Clustering and Evaluation

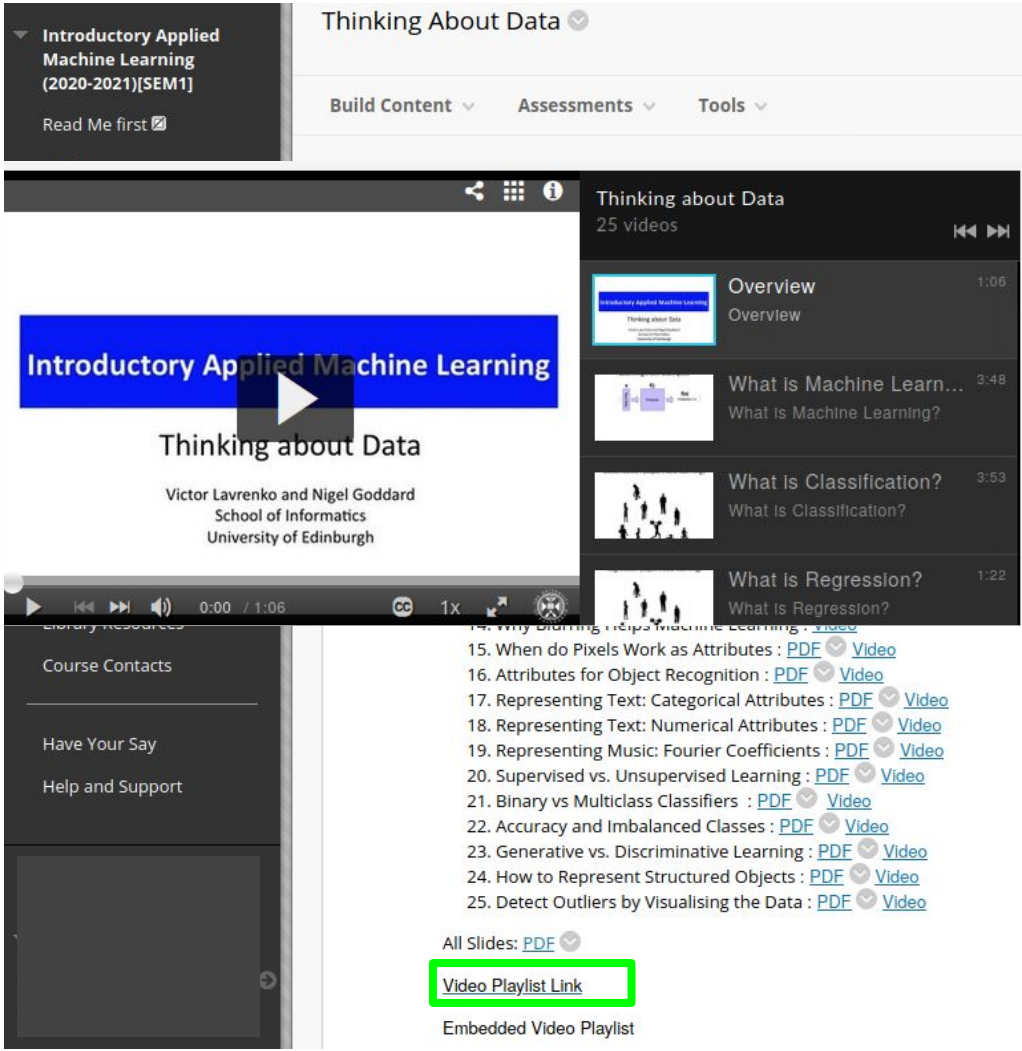
Lectures

Pre-recorded - you watch on your own in advance of the class meetings



Lectures

Pre-recorded - you watch on your own in advance of the class meetings

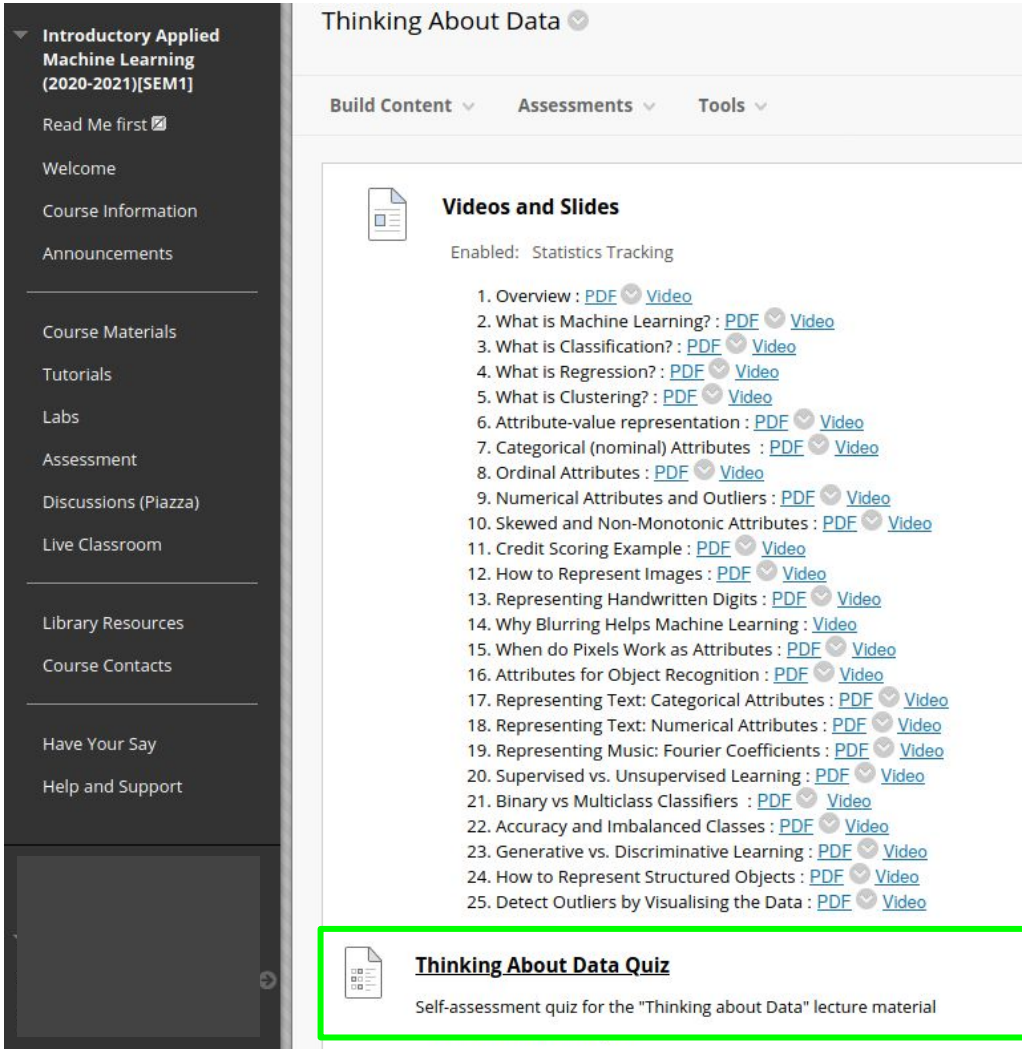


Lectures - Quiz

After you have watched the lecture - take the online quiz

You can take each quiz as many times as you like

Note: There are **no** quizzes for the lectures in week 1



Class Meetings

You are in/watching the class meeting right now!

Twice a week: **Mon @ 14:10** and **Thurs @ 14:10**

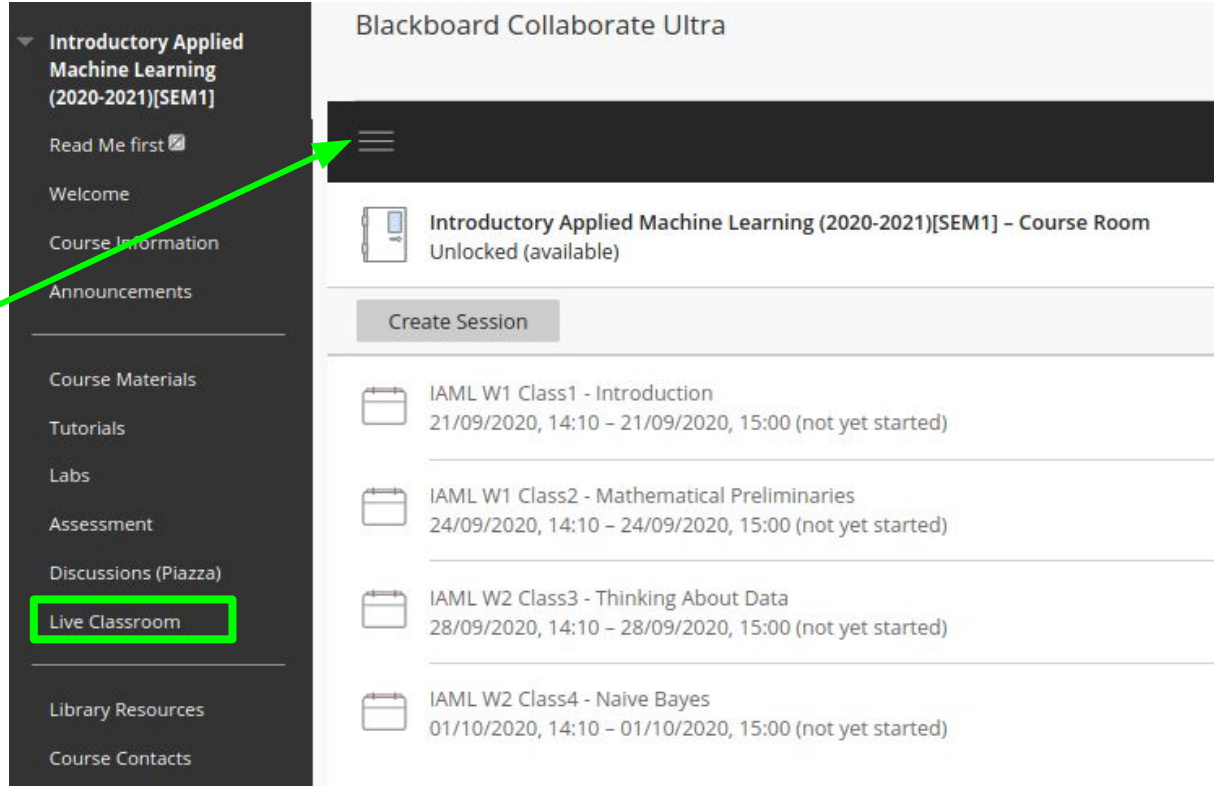
We will go over additional examples

Opportunity to ask questions (best to do this on Piazza in advance - more on this later)

Class Meetings

Class meetings will be recorded so you can watch them offline later

Click here to find them



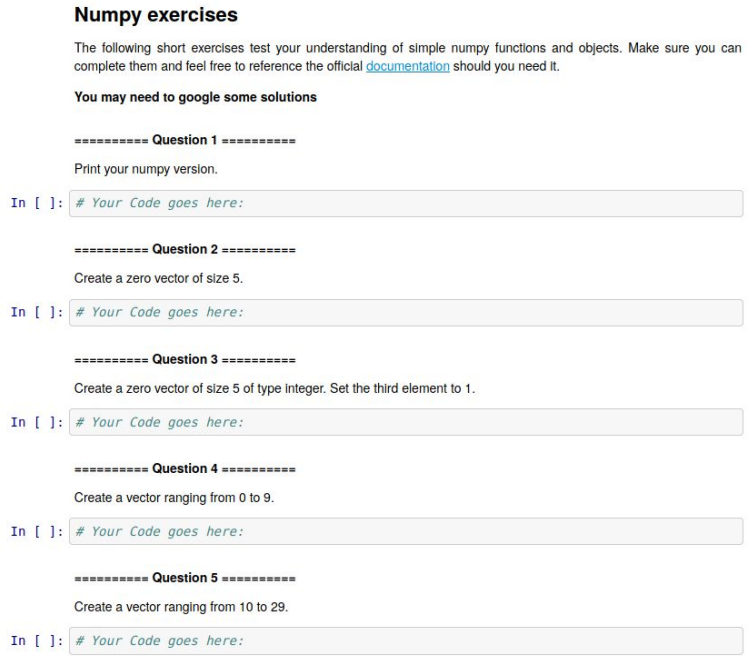
Labs

Weeks: 1, 2, 4, 6, 8, and 10

Python using Jupyter Notebooks

Lab sessions will be 1 hour long and conducted via Blackboard Collaborate in smaller groups

Lab groups for week 2 onwards will be provided soon - you need to register for the course



Numpy exercises

The following short exercises test your understanding of simple numpy functions and objects. Make sure you can complete them and feel free to reference the official documentation should you need it.

You may need to google some solutions

========== **Question 1** ==========

Print your numpy version.

```
In [ ]: # Your Code goes here:
```

========== **Question 2** ==========

Create a zero vector of size 5.

```
In [ ]: # Your Code goes here:
```

========== **Question 3** ==========

Create a zero vector of size 5 of type integer. Set the third element to 1.

```
In [ ]: # Your Code goes here:
```

========== **Question 4** ==========

Create a vector ranging from 0 to 9.

```
In [ ]: # Your Code goes here:
```

========== **Question 5** ==========

Create a vector ranging from 10 to 29.

```
In [ ]: # Your Code goes here:
```

Labs: Week 1

Week 1: Lab “00 - Introduction.ipynb”

Setting up your Python environment either locally (recommended) or by remotely logging into Informatics computers (DICE)

Introduction to Python and some of the core libraries we will use e.g. numpy

Important to get this set up now so you can do the labs and courseworks in future weeks

Labs: Week 1 Drop in Sessions

Special drop in lab sessions in Week 1 if you are stuck or have problems with the first lab

Thurs 24th Sept 9-7pm and **Fri 25th Sept 9-5pm**

See Learn for information on how to join

Try to spread out over the day, if busy, try again later

Tutorials

Weeks: 3, 5, 7, and 9

Smaller groups

Exam style questions

They will help you further understand the material so that you are better prepared for the exam

Will happen online - more information will follow soon

Courseworks

CW1 worth 20% - due 20th October

CW2 worth 30% - due 20th November

Both will be available on Learn ~2 weeks before the submission deadline

Actively engaging with the labs and tutorials will be very helpful preparation for the courseworks

Make sure you do the coursework corresponding to your course code

Courseworks: How to submit

You will submit a PDF document generated using Latex

A template will be given to you

More information will be provided closer to the submission date

Exam

Worth 50% - will take place in December

See Learn (Assessments page) for links to previous exam papers

Again, make note of the different course codes

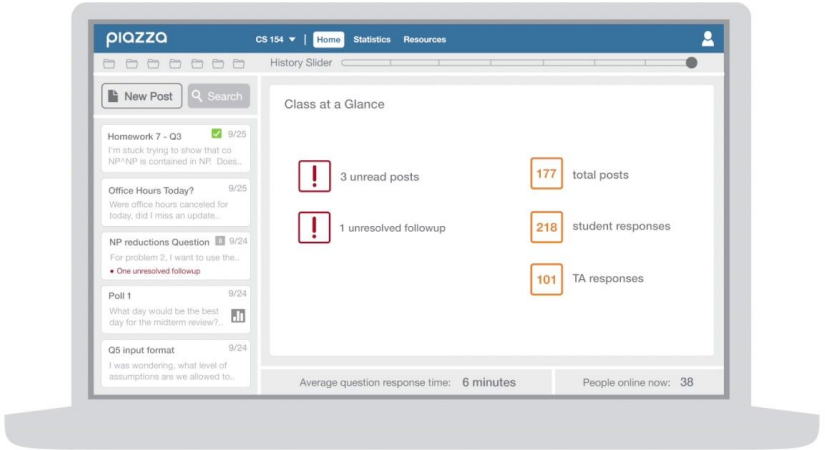
IAML Piazza

piazza.com/ed.ac.uk/fall2020/infr10069andinfr11182

Make sure you are signed up if taking IAML

Your opportunity to ask questions in advance of class sessions

Use the relevant topics



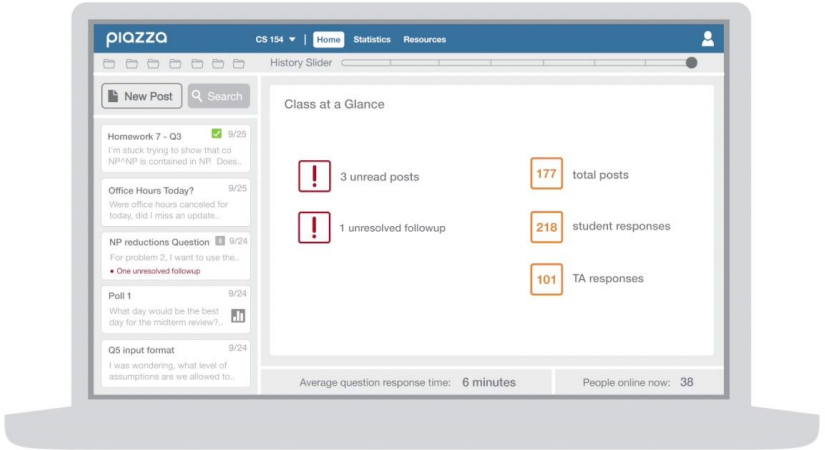
IAML Piazza

piazza.com/ed.ac.uk/fall2020/infr10069andinfr11182

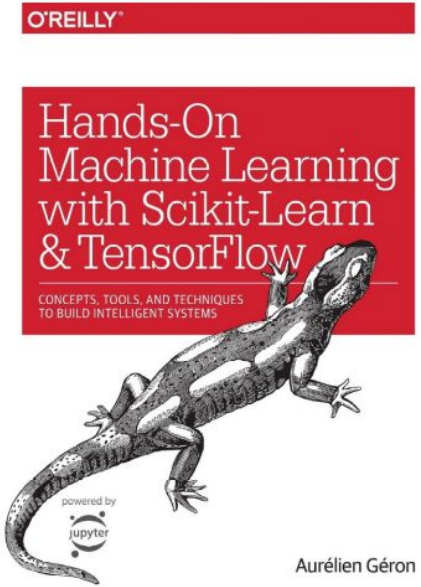
You can also answer each other questions - peer learning

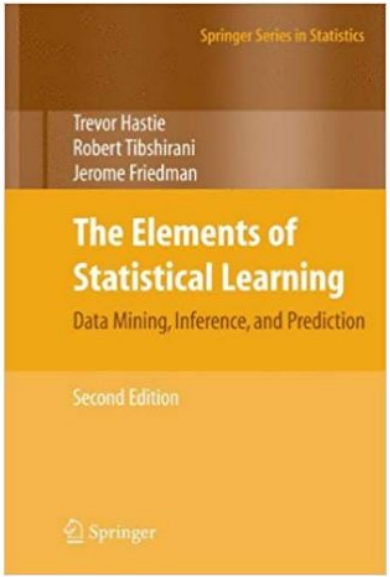
We will have a TA monitoring Piazza and they will also respond to questions

No discussion of coursework



Recommended Texts





See Resource List on Learn for how to access these books and others online

Best of Luck!

What do you want to get out of IAML? - use this to guide your learning

Advice

Monitor your email and Piazza for announcements

Watch the lectures before the class sessions

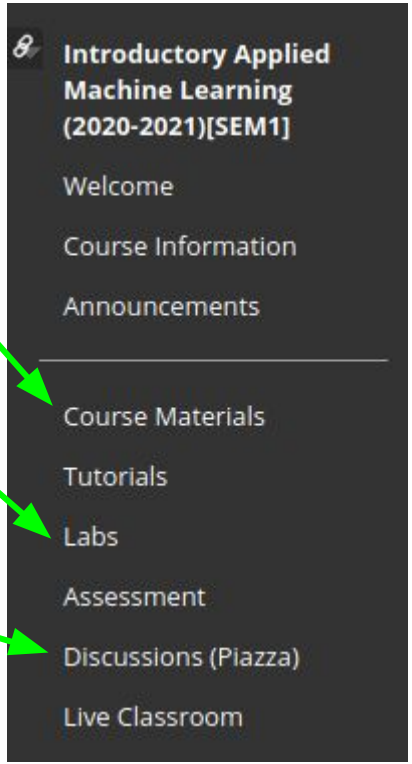
Don’t leave the courseworks until the last minute

Start on the labs and tutorials on your own before the online sessions

Engage in the labs and tutorials

Week 1: Your tasks for this week

1) Study the “Maths and Probability” lectures on Learn before Thursday’s class meeting
2) Start “Lab 0 - Intro to Python”. Ask questions in the lab drop in sessions on Thurs and Fri this week if stuck (see “Labs” on Learn)
3) Ask questions on Piazza about the lectures - earlier the better



https://course.inf.ed.ac.uk/iaml

Questions?

Write in the chat box