COMS30035, Machine learning: Introduction

Edwin Simpson (based on slides by Rui Ponte Costa)

Department of Computer Science, SCEEM
University of Bristol

September 7, 2023

What is machine learning?

- Machine Learning (ML) is the study of computer **algorithms** that **learn** to perform a task from **data** or experience.

What is machine learning?

- Machine Learning (ML) is the study of computer **algorithms** that **learn** to perform a task from **data** or experience.
- The algorithm learns a **model** of the data for making **predictions**, **decisions**, or to help **understand** the data.

What is machine learning?

- Machine Learning (ML) is the study of computer **algorithms** that **learn** to perform a task from **data** or experience.
- The algorithm learns a **model** of the data for making **predictions**, **decisions**, or to help **understand** the data.
- It is typically grounded in Statistics and seen as a subfield of Artificial Intelligence.

Machine learning interest [Google trends]

Examples:

What kind of tasks can we learn from data?

Examples: Linear regression

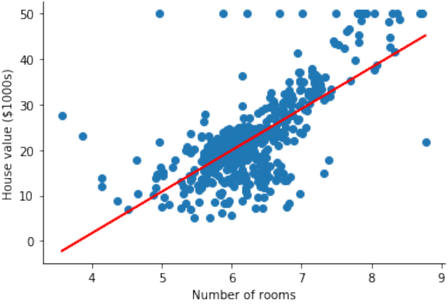
Observe one numerical variable and use it to predict another using a linear model.
Pros: Simple model **Cons**: Simple model

Example: Can we predict property prices in Boston?

Examples: Linear regression

Observe one numerical variable and use it to predict another using a linear model.
Pros: Simple model **Cons**: Simple model

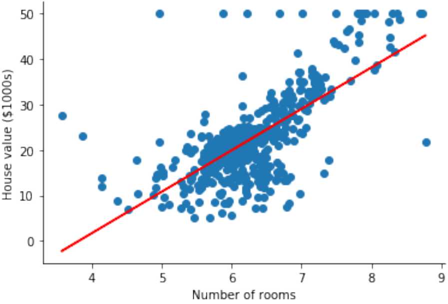
Example: Can we predict property prices in Boston?

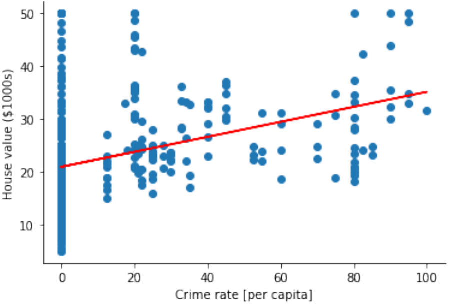


Examples: Linear regression

Observe one numerical variable and use it to predict another using a linear model.
Pros: Simple model **Cons**: Simple model

Example: Can we predict property prices in Boston?





Examples: Classification

When we want to automatically separate data into **discrete classes**.

Examples: Classification

When we want to automatically separate data into **discrete classes**.

Example: Is it a dog or a bagel?

Examples: Dimensionality reduction

Example application: compress data with many dimensions to 2D so that we can visualise it.

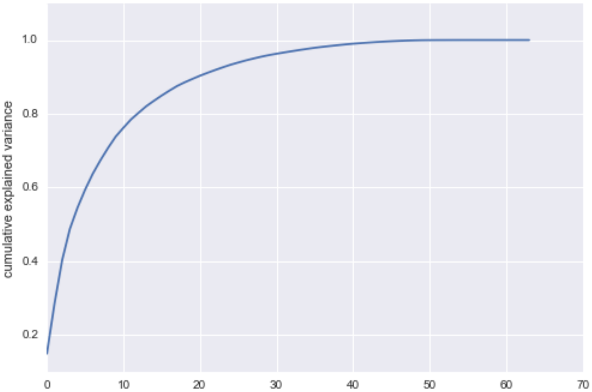
Pros: Easy to compute (e.g., using PCA). **Cons**: Loss of feature semantics.

Examples: Dimensionality reduction

Example application: compress data with many dimensions to 2D so that we can visualise it.

Pros: Easy to compute (e.g., using PCA). **Cons**: Loss of feature semantics.

Example: How many features/components are enough to explain 80% of the variance in a given dataset? More here.



Examples: Generating text in natural language

Pros: Can give impressive results (e.g. using deep learning).
Cons: Models can be hard to interpret and control.

Examples: Generating text in natural language

Pros: Can give impressive results (e.g. using deep learning).
Cons: Models can be hard to interpret and control.

Example using a Transformer (a type of neural network):

Input: "Machine Learning is the study of computer algorithms that learn a model of data."

Examples: Generating text in natural language

Pros: Can give impressive results (e.g. using deep learning).
Cons: Models can be hard to interpret and control.

Example using a Transformer (a type of neural network):

Input: "Machine Learning is the study of computer algorithms that learn a model of data."

Output/completion: "Traditional AI, such as natural language processing, machine translation, image recognition, and machine translation relies heavily on well-defined data, and can perform well without much further manipulation of the data. Typically, a machine learning system recognizes patterns of data and then uses those patterns to predict outcomes of a given scenario. (see prediction)"

Try it out **app.inferkit.com/demo**!

Examples:

- What is a **model**?

Examples:

- What is a **model**?
- A way of understanding how **input** data relates to the desired **outputs**

Examples:

- What is a **model**?
- A way of understanding how **input** data relates to the desired **outputs**
- E.g., a function that takes images as input and outputs a label "dog" or "bagel"

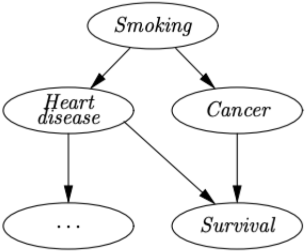
Examples: Graphical models

When we use a graphical model to represent an explicit probabilistic model.

Examples: Graphical models

When we use a graphical model to represent an explicit probabilistic model.

Example: What are the changes of survival for a smoker?



$\Pr(cancer \mid smoking)$	$\Pr(cancer \mid \neg smoking)$
$\Pr(\neg cancer \mid smoking)$	$\Pr(\neg cancer \mid \neg smoking)$

How does it relate to other units

- This unit builds directly on previous units

How does it relate to other units

- This unit builds directly on previous units
  - Data Science for CS [introduction to data science in **Python**]

How does it relate to other units

- This unit builds directly on previous units
  - Data Science for CS [introduction to data science in **Python**]
  - Algorithms I and II [design/analyse algorithms, and data structures]

How does it relate to other units

- This unit builds directly on previous units
  - Data Science for CS [introduction to data science in **Python**]
  - Algorithms I and II [design/analyse algorithms, and data structures]
  - Math for CS I and II [algebra and statistics]

How does it relate to other units

- This unit builds directly on previous units
  - Data Science for CS [introduction to data science in **Python**]
  - Algorithms I and II [design/analyse algorithms, and data structures]
  - Math for CS I and II [algebra and statistics]
- And is a building block for more advanced units (4th year):
  - Applied Deep Learning (neural networks)
  - Information Processing & the Brain (ML and neuroscience)
  - Cloud Computing and Big Data (ML on the cloud)
  - Advanced Computer Vision (ML for vision)
  - Applied Data Science (ML and data management)

How does it relate to other units

- This unit builds directly on previous units
  - Data Science for CS [introduction to data science in **Python**]
  - Algorithms I and II [design/analyse algorithms, and data structures]
  - Math for CS I and II [algebra and statistics]
- And is a building block for more advanced units (4th year):
  - Applied Deep Learning (neural networks)
  - Information Processing & the Brain (ML and neuroscience)
  - Cloud Computing and Big Data (ML on the cloud)
  - Advanced Computer Vision (ML for vision)
  - Applied Data Science (ML and data management)
  - Individual Project (e.g., on applications or ML or advanced ML)

What we will cover

We will focus on the following textbook

- **Bishop, C. M., Pattern recognition and machine learning (2006). Available for free here.**

With some sections from:

What we will cover

We will focus on the following textbook

- **Bishop, C. M., Pattern recognition and machine learning (2006). Available for free here.**

With some sections from:

- Murphy, K., Machine learning a probabilistic perspective (2012). The book is also freely available here.

What we will cover

What we will cover

1. Part 1: Overview, ML concepts, revisiting regression, classification and neural networks

What we will cover

1. Part 1: Overview, ML concepts, revisiting regression, classification and neural networks
2. Part 2: Kernel methods (SVM), Probabilistic graphical models, (mixture models, EM), Dimensionality reduction (PCA, ICA)

What we will cover

1. Part 1: Overview, ML concepts, revisiting regression, classification and neural networks
2. Part 2: Kernel methods (SVM), Probabilistic graphical models, (mixture models, EM), Dimensionality reduction (PCA, ICA)

3. Part 3: Sequential data (HMM, LDS), Selecting and Combining Models (Ensembles)

What we will cover

1. Part 1: Overview, ML concepts, revisiting regression, classification and neural networks [» Edwin Simpson]
2. Part 2: Kernel methods (SVM), Probabilistic graphical models, (mixture models, EM), Dimensionality reduction (PCA, ICA)

3. Part 3: Sequential data (HMM, LDS), Selecting and Combining Models (Ensembles)

What we will cover

1. Part 1: Overview, ML concepts, revisiting regression, classification and neural networks [» Edwin Simpson]
2. Part 2: Kernel methods (SVM), Probabilistic graphical models, (mixture models, EM), Dimensionality reduction (PCA, ICA) [» James Cussens, unit director]
3. Part 3: Sequential data (HMM, LDS), Selecting and Combining Models (Ensembles)

What we will cover

1. Part 1: Overview, ML concepts, revisiting regression, classification and neural networks [» Edwin Simpson]
2. Part 2: Kernel methods (SVM), Probabilistic graphical models, (mixture models, EM), Dimensionality reduction (PCA, ICA) [» James Cussens, unit director]
3. Part 3: Sequential data (HMM, LDS), Selecting and Combining Models (Ensembles) [» Edwin Simpson]

Unit Outline

On the unit BB page; and uob-coms30035.github.io

Weeks	*First lecture*	*Second lecture*	*Labs [Thu 9am-12pm]*	*Live lecture [Q&A class; Tues 9-10am]*
1	Introduction [stream]	Machine learning concepts [stream]	L1: Revision of Jupyter Notebook and ML libraries [answers]	General questions about the unit
2	Revisiting regression [stream], Bayesian regression [stream], Classification and neural networks [stream]	Kernel machines [Kernels 1, stream] [Kernels 2, stream] [Kernels 3, stream]	L2: Regression, nnets and SVMs	ML concepts, regression, classification and nnets
3	Introduction to graphical models [Prob revision, stream, mp4] [PGM 1, stream, mp4] [Netica demo, stream, mp4] [PGM 2, stream, mp4] [PGM 3, stream, mp4]	Bayesian ML using graphical models	L3: Probabilistic graphical models [answers]	Kernel machines and probabilistic graphical models
4	k-means and mixtures of Gaussians	The EM algorithm	L4: k-means and EM [answers]	The k-means and EM
5	PCA	kernel PCA and ICA	L5: PCA and ICA [answers]	PCA and ICA
6	**Reading week**			
7	Sequential data	Sequential data	L6: Hidden Markov Models [answers]	Modelling sequential data
8	Selection and Combination	Trees, Mixtures and Crowds	L7: Trees and Ensemble methods [answers]	Combining models using ensembles and probabilistic methods
9-11	**Coursework weeks**			
12	**Review week**			

Scheduled Sessions

- Lectures: Tuesday, 3pm, Queens 1.15 and Friday, 9am, Queens 1.15

[1] For those doing the CS undergrad, we will follow a similar setup to Data-driven CS.

Scheduled Sessions

- Lectures: Tuesday, 3pm, Queens 1.15 and Friday, 9am, Queens 1.15
- Labs: Thursday 9am - Noon in MVB 2.11

[1] For those doing the CS undergrad, we will follow a similar setup to Data-driven CS.

Scheduled Sessions

- Lectures: Tuesday, 3pm, Queens 1.15 and Friday, 9am, Queens 1.15
- Labs: Thursday 9am - Noon in MVB 2.11
    - Lab Environment [Jupyter + Python] [1]

[1] For those doing the CS undergrad, we will follow a similar setup to Data-driven CS.

Scheduled Sessions

- Lectures: Tuesday, 3pm, Queens 1.15 and Friday, 9am, Queens 1.15
- Labs: Thursday 9am - Noon in MVB 2.11
  - Lab Environment [Jupyter + Python] [1]
  - Lab relates to the previous Friday + Tuesday lecture

[1] For those doing the CS undergrad, we will follow a similar setup to Data-driven CS.

Scheduled Sessions

- Lectures: Tuesday, 3pm, Queens 1.15 and Friday, 9am, Queens 1.15
- Labs: Thursday 9am - Noon in MVB 2.11
  - Lab Environment [Jupyter + Python] [1]
  - Lab relates to the previous Friday + Tuesday lecture
  - This week we introduce Scikit-learn and revisit linear regression

[1] For those doing the CS undergrad, we will follow a similar setup to Data-driven CS.

Scheduled Sessions

- Lectures: Tuesday, 3pm, Queens 1.15 and Friday, 9am, Queens 1.15
- Labs: Thursday 9am - Noon in MVB 2.11
  - Lab Environment [Jupyter + Python] [1]
  - Lab relates to the previous Friday + Tuesday lecture
  - This week we introduce Scikit-learn and revisit linear regression
- Teams: Please post any questions here! We will try to reply or answer in the lectures

[1] For those doing the CS undergrad, we will follow a similar setup to Data-driven CS.

Assessments

- **Option 1, 100% Exam**:
  - COMS30033, 10 CP

Assessments

- **Option 1, 100% Exam**:
  - COMS30033, 10 CP
  - Mathematical questions, explanations of ML methods and concepts, working through toy examples

Assessments

- **Option 1, 100% Exam**:
  - COMS30033, 10 CP
  - Mathematical questions, explanations of ML methods and concepts, working through toy examples
  - Completing labs is extremely beneficial to your understanding of ML – make sure you attend labs!

Assessments

- **Option 1, 100% Exam**:
  - COMS30033, 10 CP
  - Mathematical questions, explanations of ML methods and concepts, working through toy examples
  - Completing labs is extremely beneficial to your understanding of ML – make sure you attend labs!
- **Option 2, 50% Exam, 50% Lab**:
  - Same exam as above
  - Coursework will be a small project involving an ML challenge and experiments
  - Assessed based on a written report project
  - The methods taught in labs will be needed for the coursework
  - Coursework released at beginning of week 9 (coursework period is full-time during weeks 9-11)

Assessments

- **Option 1, 100% Exam**:
  - COMS30033, 10 CP
  - Mathematical questions, explanations of ML methods and concepts, working through toy examples
  - Completing labs is extremely beneficial to your understanding of ML – make sure you attend labs!
- **Option 2, 50% Exam, 50% Lab**:
  - Same exam as above
  - Coursework will be a small project involving an ML challenge and experiments
  - Assessed based on a written report project
  - The methods taught in labs will be needed for the coursework
  - Coursework released at beginning of week 9 (coursework period is full-time during weeks 9-11)
- Discussion with others is encouraged, but submissions need to be unique [**plagiarism is taken seriously!**]