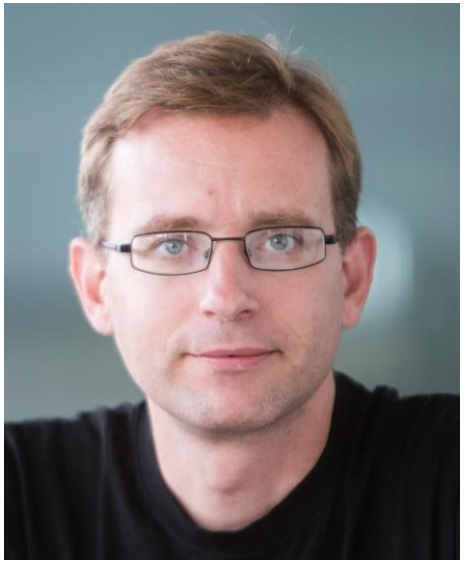
MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Machine Learning

Lecture 1: Introduction to the module

Introduction



- Dr Christian Beder
  - Lecturer at the Department of Computing
  - Email: Christian.Beder@mtu.ie
  - Office Room: B180A

- Qualification & Experience:
  - Lecturer in the CS department teaching AI subjects such as Machine Learning, Decision Analytics and Computer Vision
  - Researcher at CIT's Nimbus centre 2010-2019 working on large-scale collaborative European Research Projects
  - PhD in Photogrammetry 2006
  - MSc in Computer Science 2002

What is Machine Learning?

- Machine learning is the use of statistical analysis of existing data to infer a relation between input and output with the goal of computing this relationship for new data

- Depending on the output data we distinguish
  - Classification
  - Regression
  - Clustering

- Depending on the input data we distinguish
  - Supervised learning algorithms
  - Unsupervised learning algorithms
  - Reinforcement learning algorithms

What is Classification?

- Decide what category an object belongs to based on a set of measured features

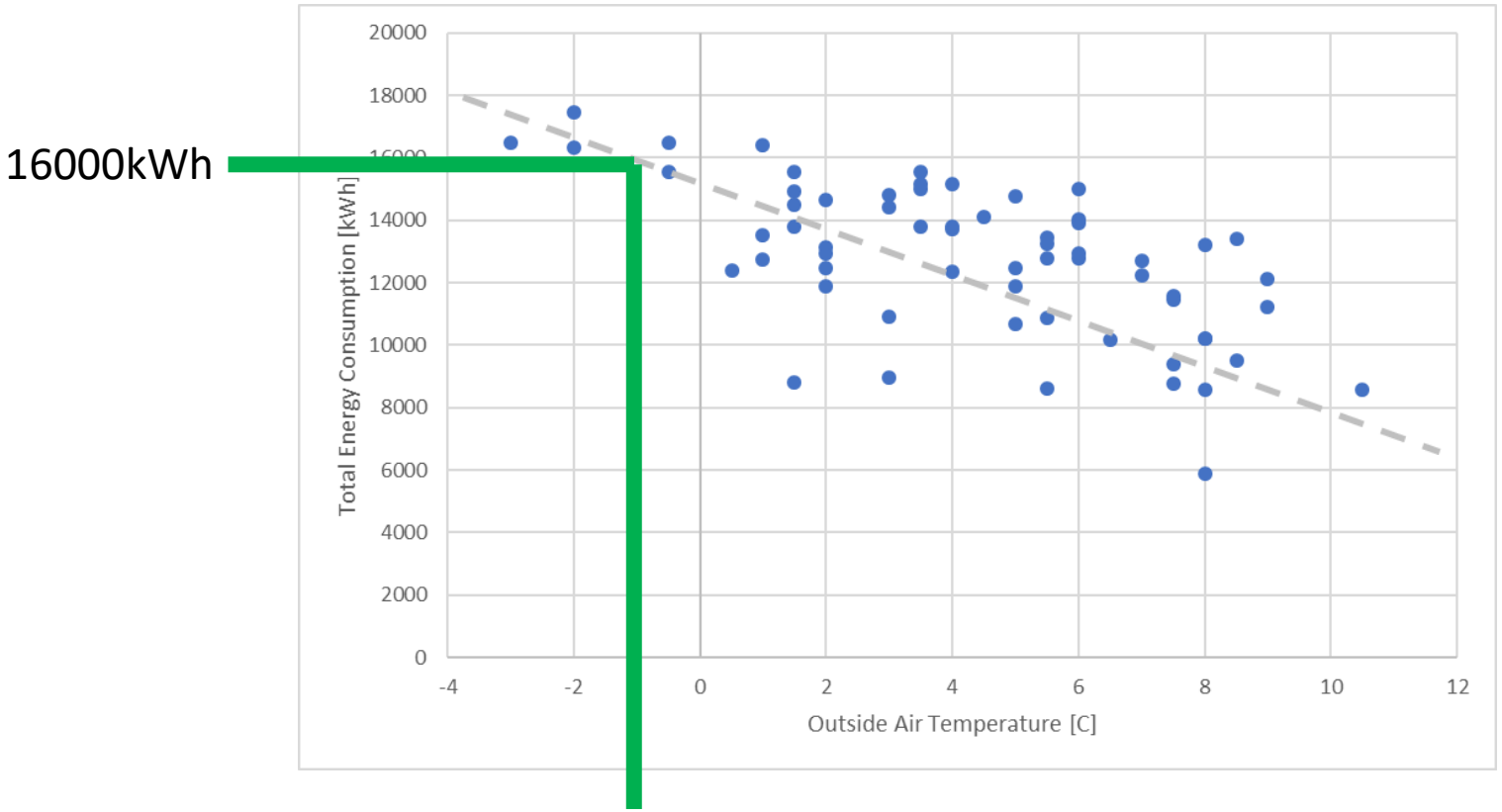
Is it an apple or an orange?

→ “Apple”

→ “Orange”

What is Regression?

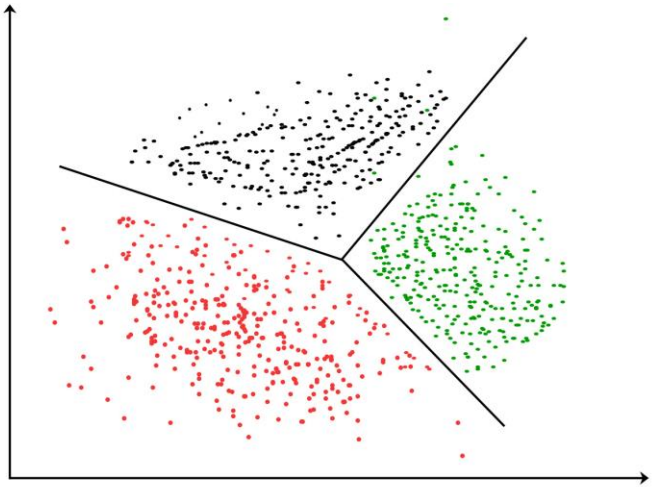
- Predict a numerical quantity or a relation between quantities based on a set of measurements



If it is -1 deg outside, what will my heating bill look like?

What is Clustering?

- Determine which measurements in a dataset form a cluster, i.e. can be considered similar for some subsequent task

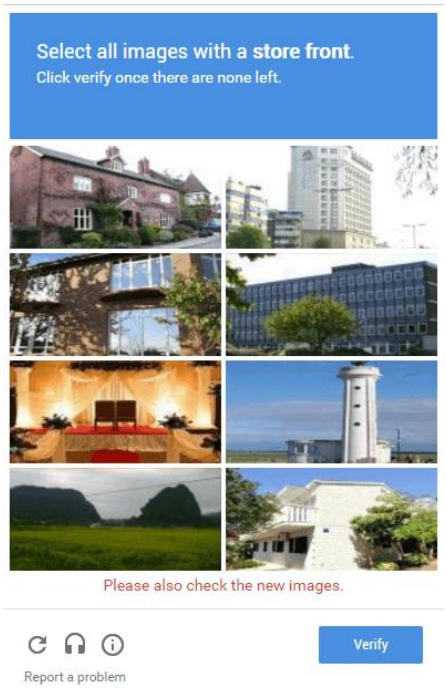


What looks similar?
What does not?

What is Supervised Learning?

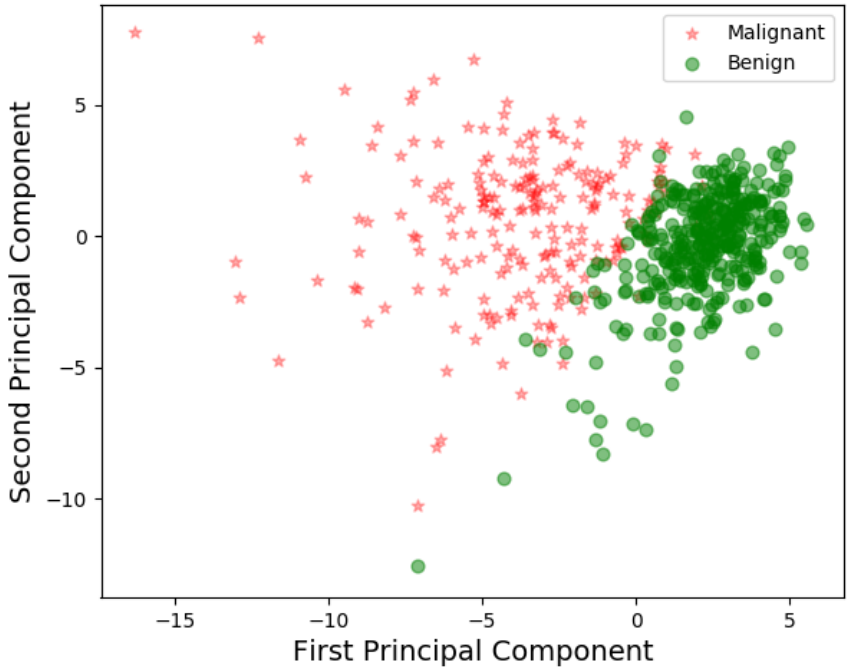
- Supervised learning algorithms use labelled training data to determine a functional relationship between input and output
- Generating/obtaining this training data is not easy as somebody need to assign labels for a large dataset
- Can sometimes be crowd-sourced

Workers at the headquarters of Ruijin Technology Company in Jiaxian, in central China's Henan Province. They identify objects in images to help artificial intelligence make sense of the world.Credit...Yan Cong for The New York Times



What is Unsupervised Learning?

- Discover patterns in the data without labels based solely on similarity between data-points
- Example are dimensionality reduction and clustering



What is Reinforcement Learning?

- Learn a strategy based on observed rewards

- Requires experimentation with the environment
- Very well-suited for synthetic environments, such as computer games, that can be tried and repeated without cost

Learning outcomes

Module descriptor
https://courses.cit.ie/index.cfm/page/module/moduleId/13881

- LO1: Apply machine learning methodologies to facilitate pre-processing, dimensionality reduction and model selection.
- LO2: Select and apply appropriate machine learning algorithms to datasets from a specific application domain.
- LO3: Analyse and evaluate the performance of machine learning algorithms.
- LO4: Develop a machine learning algorithm for solving a real-world problem.
- LO5: Implement and apply optimization algorithms for solving complex problems with a high dimensional search space.

More informal learning outcomes

- Understand what type of problems can be solved using machine learning.
- Be able to select the right approach to solve a given machine learning problem.
- Understand how to evaluate and visualise the performance of a machine learning algorithm to be able to assess and compare.
- Solve real-world problems using a machine learning method.
- Solve real-world problems using appropriate optimisation algorithms.

Course breakdown & Assessment

Module descriptor
https://courses.cit.ie/index.cfm/page/module/moduleId/13881

- 2x1h lecture / week (**Tue 4pm-5pm** & **Fri 9am-10am**)
  - Presentation of theory and concepts

- 1x2h lab session / week split into two groups
  (A: **Wed 2pm-4pm**, B: **Thu 11am-1pm**)
  - Apply concepts seen in class
  - Work on assignments

- 100% Continuous Assessment
  - 30% assignment due in week 6 (**21/10/2022**)
  - 35% assignment due in week 9 (**18/11/2022**)
  - 35% assignment due in week 12 (**09/12/2022**)

Module content

- L1-L2: Introduction
  - Examples of ML applications, categories of ML algorithms, typical design-cycle and data flow in a ML problem, issues with dimensionality and overfitting
- L3-L5: Tools
  - NumPy, Pandas, scikit-learn
- L6-L7: Data processing & methodology
  - feature extraction, noise/outliers, results visualisation and analysis techniques, cross-validation
- L8-L17: Supervised learning
  - Bayesian classification, Bayesian networks, decision trees, k-nearest neighbours, linear separation, support vector machines, multi-layer neural networks

Module content

- L18-L20: Regression & optimisation
  - Parameter estimation, gradient descent, simulated annealing
- L21-L24: Unsupervised learning
  - K-means clustering, principal component analysis, generative adversarial networks

Assignments

- A1: Bayesian classification (due: **21/10/2022**)
- A2: Supervised learning (due: **18/11/2022**)
- A3: Optimisation (due: **09/12/2022**)

All assignments comprise of

- Implementation of the algorithm, including data input/output and visualisation
- Technical description and basic evaluation
- Report on research and detailed evaluation

Development environment

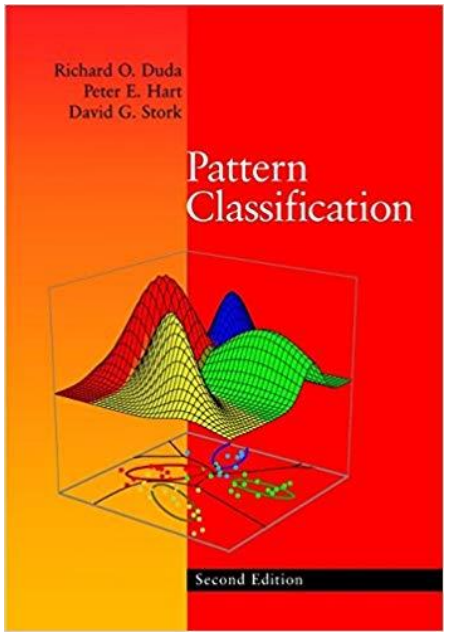
- All lab exercises and assignments will be developed in **Python3**
- The recommended IDE is Spyder and is part of the Anaconda package: https://www.anaconda.com/
- The Anaconda distribution also includes
  - NumPy https://numpy.org/
  - Pandas https://pandas.pydata.org/
  - and scikit-learn https://scikit-learn.org/stable/

Recommended resources

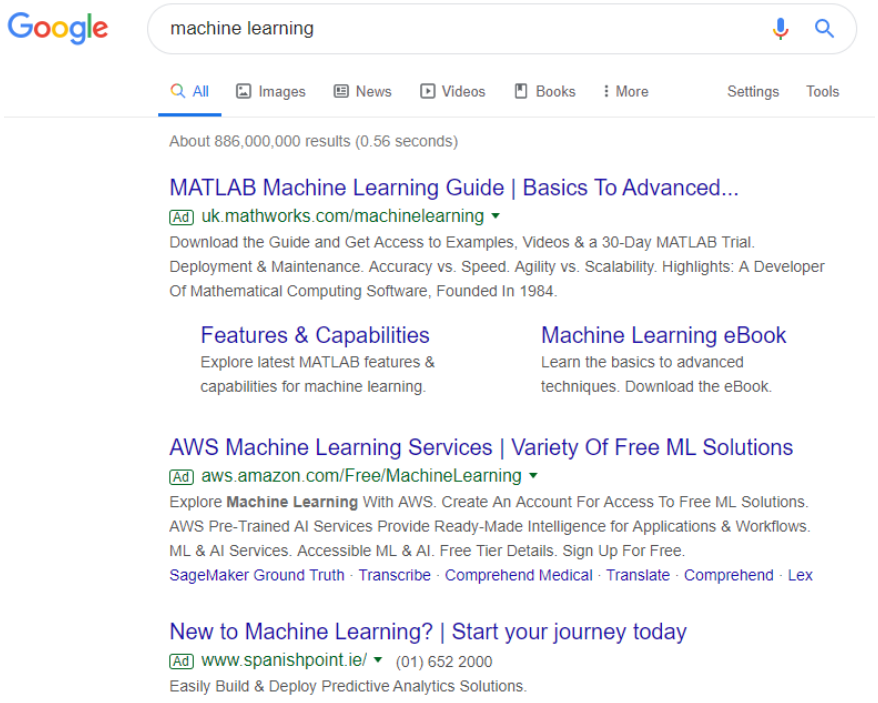
- Machine Learning Lecture by Professor Andrew Ng at Stanford

https://www.youtube.com/watch?v=UzxYlbK2c7E&list=PLA89DCFA6ADACE599

- Richard O. Duda, David G. Stork, Peter E.Hart
  Pattern Classification, Second Edition
  Wiley, 2001



Example: Targeted advertising



Example: Spam filters

CORK INSTITUTE OF TECHNOLOGY
INSTITIÚID TEICNEOLAÍOCHTA CHORCAÍ

Quarantine Digest

Christian Beder
christian.beder@cit.ie

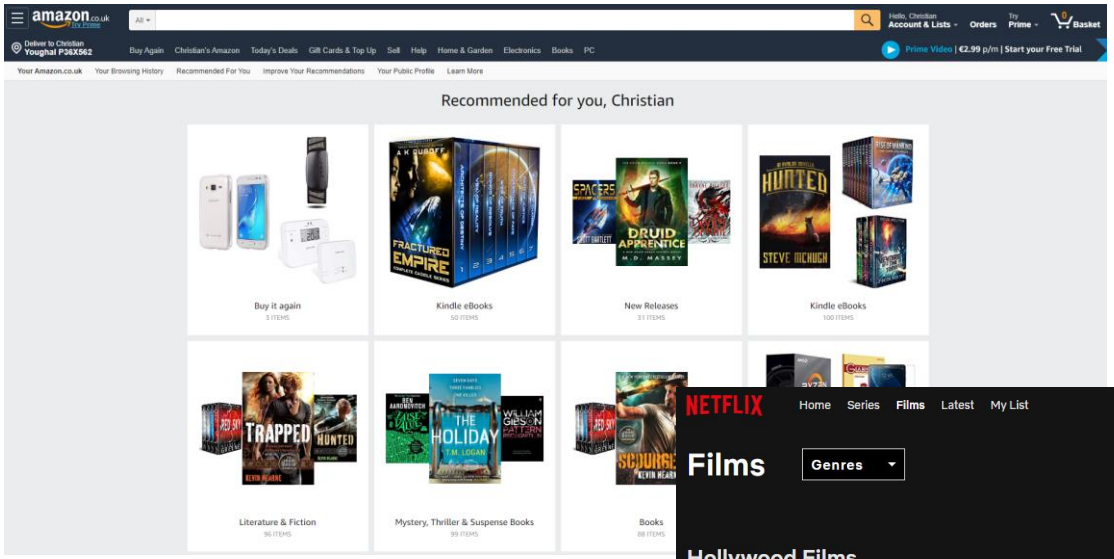
2019-09-05
19:01 IST

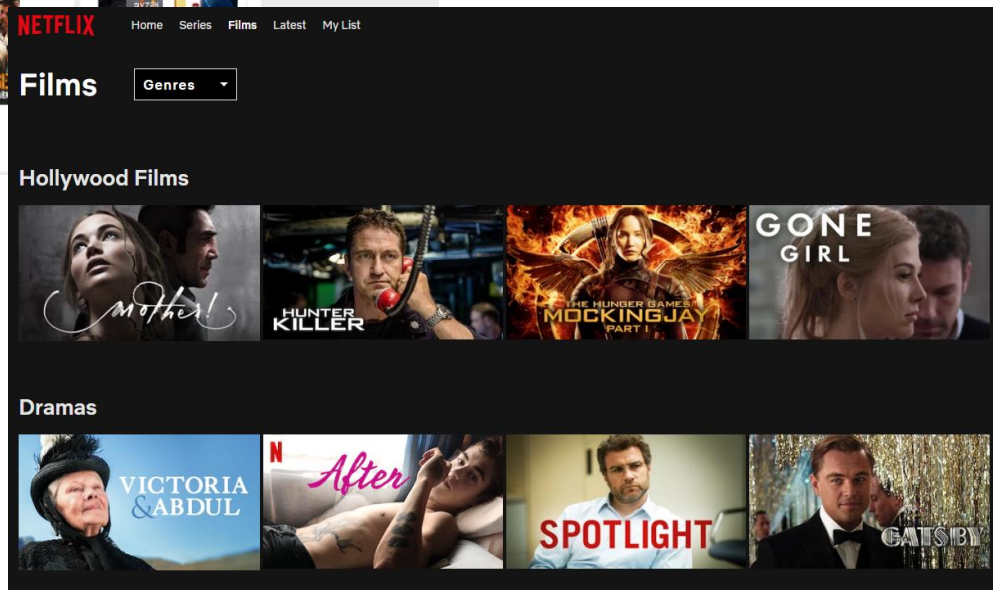
Quarantined (by Score)

Address	Subject	Delivery Date/Time	Action
RocÃo Gea - Matrikon Denmark and Ireland <info@matrikone...< b=""> 0100016d00c2cbcd-3725434d-da58-4fc5-9424-ede1b4a3694e-000000@amazonses.com	Transform Excel into a Reporting Tool for Process and Equipment Data	05 Sep 19 - 10:30 AM	Preview Release Release & Approve Block
"Shirley" aciabdlllmq@adk.biz	Re: Fiber Distribution Cabinet	05 Sep 19 - 07:27 AM	Preview Release Release & Approve Block

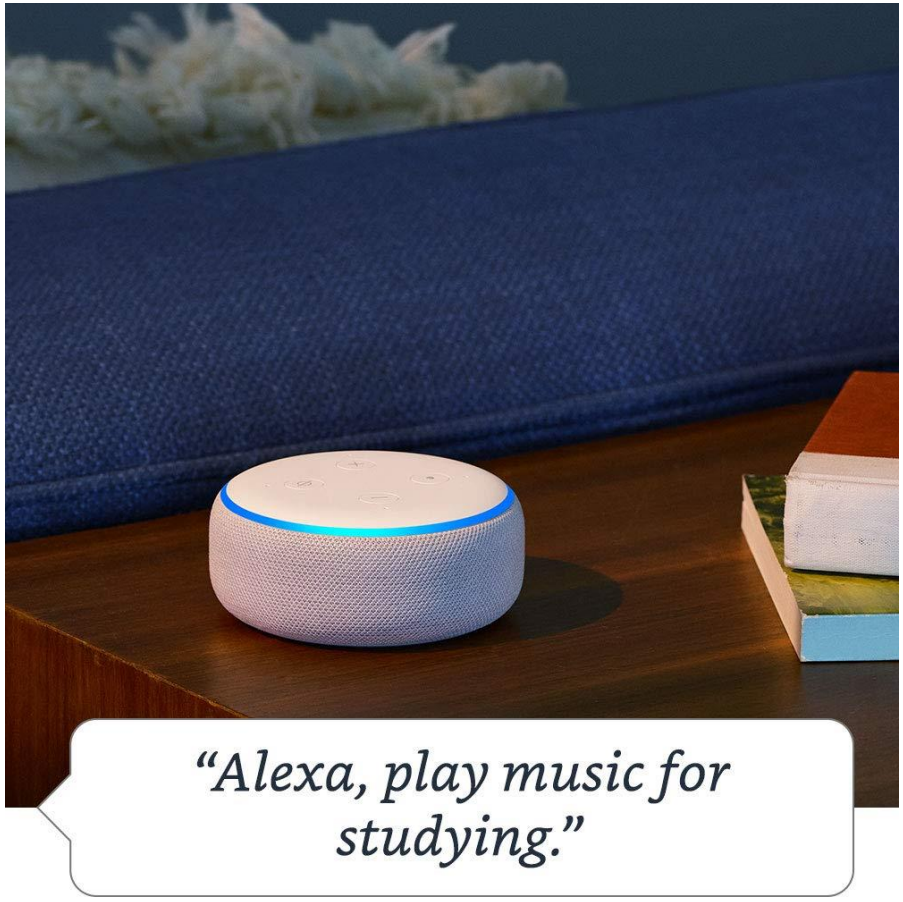
Request new digest

Example: Recommender systems

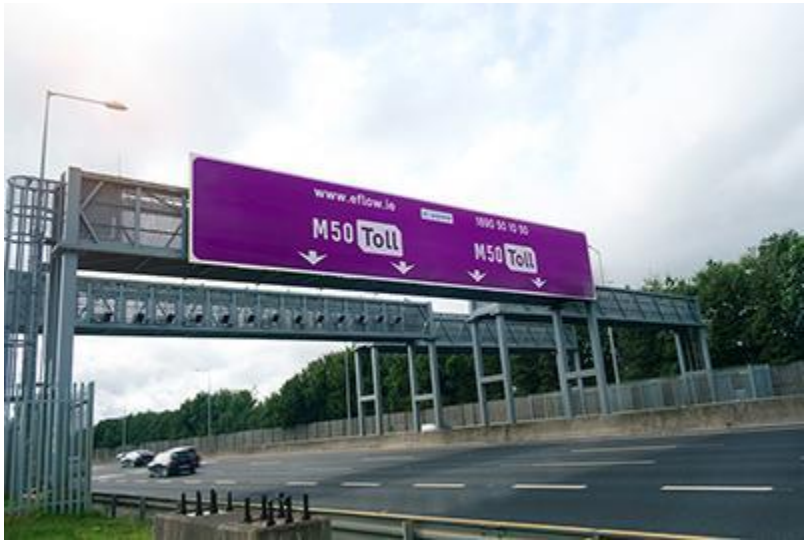




Example:
Speech recognition & natural language processing

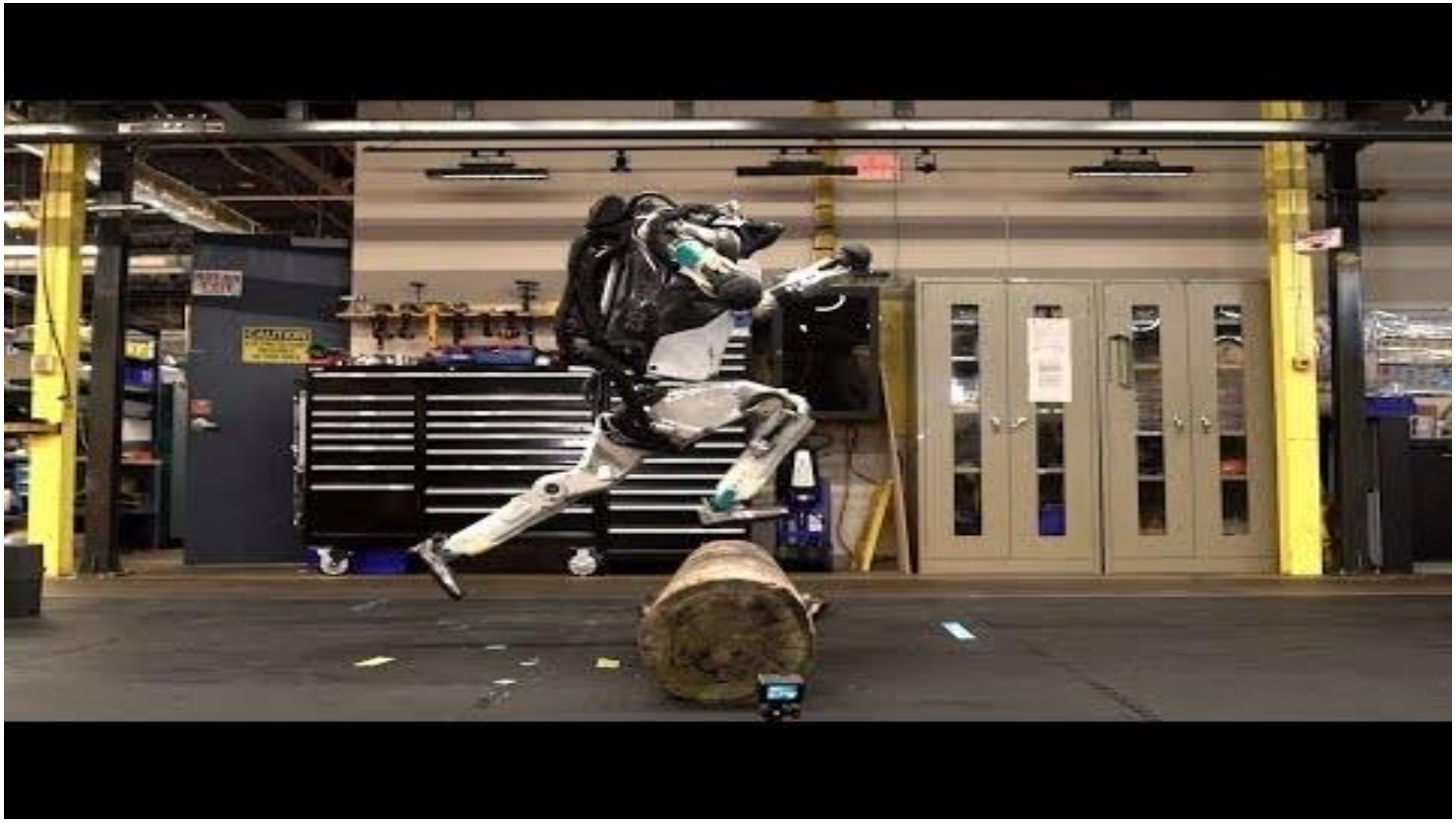


Example: Image understanding



Example: Games

Example: Robotics



https://www.youtube.com/watch?v=LikxFZZO2sk

Example: Deep fakes

https://www.washington.edu/news/2017/07/11/lip-syncing-obama-new-tools-turn-audio-clips-into-realistic-video/

Thank you for your attention