Artificial Intelligence: Risks and Benefits

Professor Ahmed Bouridane
Head – Computer and Electronic Security Systems (CESS)
Northumbria University at Newcastle
United Kingdom
Email: ahmed.bouridane@northumbria.ac.uk


J.RANK L


northumbria
UNIVERSITY NEWCASTLE


alamy stock photo


alamy stock photo


alamy stock photo

My expertise

Security related applications

- Cryptography: Algorithms and architectural implementations
- Steganography: the art of hiding covert information without its mere detection
- Video Telesurveillance : Person (re)Identification in the crowd
- Cybersecurity: Algorithmic implementations

Biomedical (outside security applications)

- Quantitative Pathology: Machine learning to detect and classify cancer early
- Quantitative Retinopathy and diabetes detection

Ahmed Bouridane

Northumbria University
University
Machine Learning
Biometric Security
Information Security and Biomedic...

GET MY OWN PROFILE

	All	Since 2014
Citations	5074	2767
h-index	32	23
i10-index	141	89

TITLE	CITED BY	YEAR
Digital image watermarking using balanced multiwavelets L Ghouti, A Bouridane, MK Ibrahim, S Boussakta IEEE transactions on signal processing 54 (4), 1519-1536	252	2006
Simultaneous feature selection and feature weighting using Hybrid Tabu Search/K-nearest neighbor classifier MA Tahir, A Bouridane, F Kurugollu Pattern Recognition Letters 28 (4), 438-446	188	2007
FPGA implementations of fast Fourier transforms for real-time signal and image processing IS Uzun, A Amira, A Bouridane IEE Proceedings-Vision, Image and Signal Processing 152 (3), 283-296	181	2005
An effective and fast iris recognition system based on a combined multiscale feature extraction technique M Nabti, A Bouridane Pattern recognition 41 (3), 868-879	164	2008
Design and implementation of a high level programming environment for FPGA-based image processing D Crookes, K Benkrid, A Bouridane, K Alotaibi, A Benkrid IEE Proceedings-Vision, Image and Signal Processing 147 (4), 377-384	136	2000
Fusion of finger-knuckle-print and palmprint for an efficient multi-biometric system of person recognition A Meraoumia, S Chitroub, A Bouridane 2011 IEEE International Conference on Communications (ICC), 1-5	84	2011
A multispectral computer vision system for automatic grading of prostatic neoplasia M Roula, J Diamond, A Bouridane, P Miller, A Amira Proceedings IEEE International Symposium on Biomedical Imaging, 193-196	80	2002
Protecting fingerprint data using watermarking K Zebbiche, L Ghouti, F Khelifi, A Bouridane First NASA/ESA Conference on Adaptive Hardware and Systems (AHS'06), 451-456	73	2006
An FPGA based coprocessor for GLCM and Haralick texture features and their application in prostate cancer classification MA Tahir, A Bouridane, F Kurugollu Analog Integrated Circuits and Signal Processing 43 (2), 205-215	69	2005

My Transfer of Know-How to DZA

A quick look at my past and current academic ctivities with DZA

- 8 full scholarships awarded to young researchers so far
- 5 successful PhD supervision so far
- 3 currently ongoing
- Hosting researchers (on average 2 each year)
- Research linkages with: Batna, Guelma, Jijel, Tiaret, Ouargla
- EMP, CGN

Learning from Data

- Germany's climate research centre generates 10 petabytes per year
- Google processes 24 petabytes per day
- The Large Hadron Collider produces 60 gigabytes per minute (~12 DVDs)
- There are over 50m credit card transactions a day in the US alone.
- A large number of multimedia data (video, audio etc) are uploaded on social media platforms

- **What might we want to do with that data?**
  - **Prediction**
    - **what can we predict about this phenomenon?**
  - **Description**
    - **how can we describe/understand this phenomenon in a new way?**

To take any action first we need to understand the data (learn from the data) and automatically (machine learning)

Over 13 million pages

Over 800 million users

Over 6 billion photos

Over 24 hrs of video per minute

Learning from Data

northumbria UNIVERSITY

- How can we extract knowledge from data to help humans take decisions?
- How can we automate decisions from data?
- How can we adapt systems dynamically to enable better user experiences?

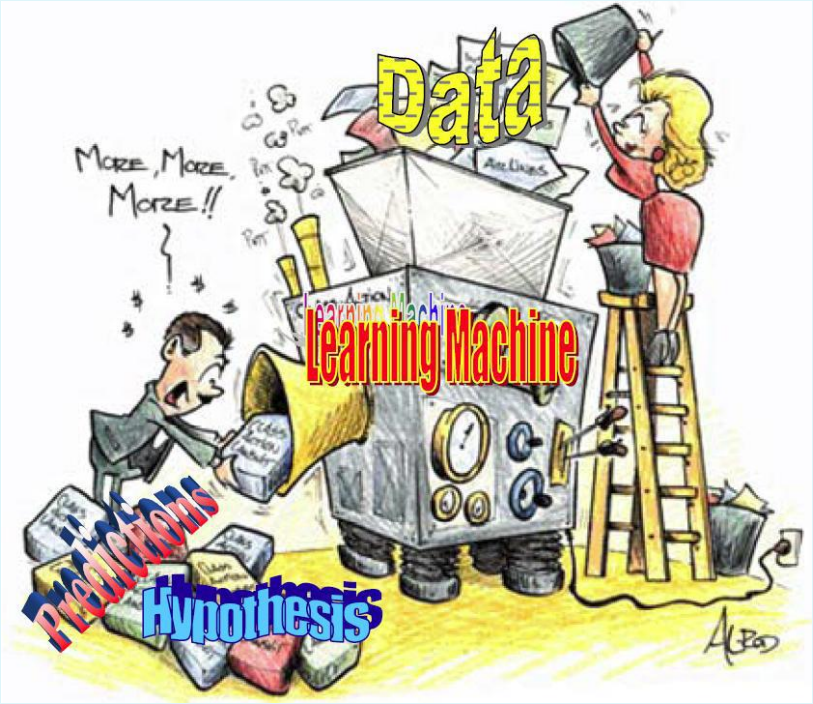
- Ultimately we want to:
  - Write code explicitly to do the above task?
  - Write code to make the computer **learn** how to do the task?

Artificial Intelligence

What is Artificial Intelligence?

Study of algorithms and techniques to mimic “human” intelligence.

Study of algorithms that improve their performance on a particular task as they learn from data (experience) using machines/computers: This is called Machine Learning



What is Artificial Intelligence?

northumbria UNIVERSITY

Computational models of human behavior?

- Computer programs that behave (externally) like humans

Computational models of human "thought"

- Computer programs that operate (internally) the way humans do

Computational systems that behave intelligently?

- What does it mean to behave intelligently?

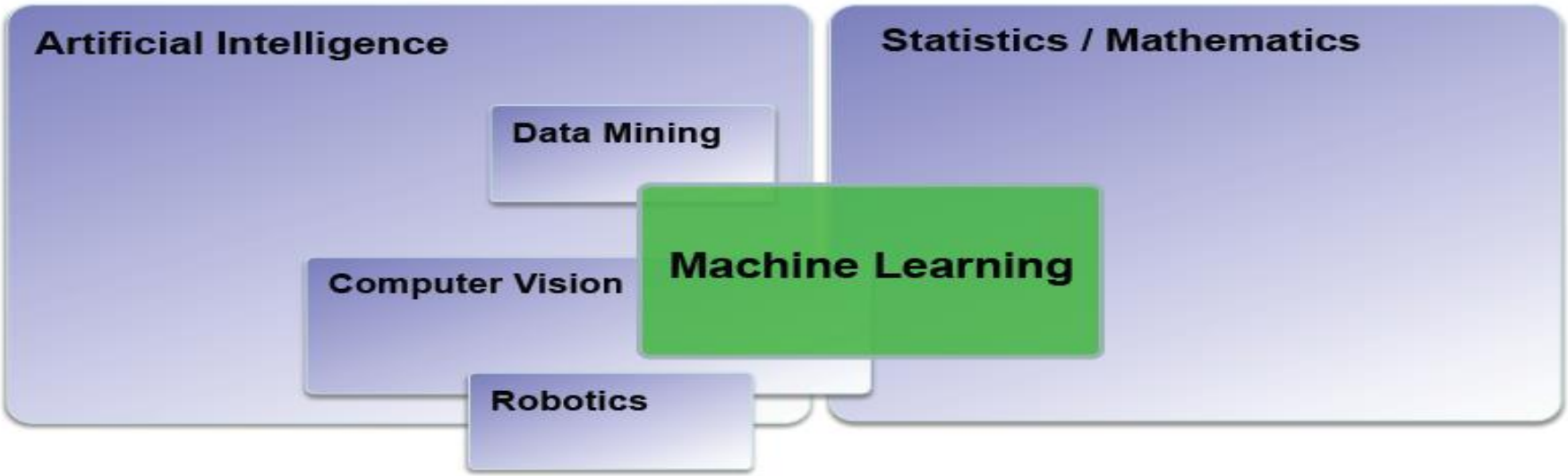
AI applications

- Monitor trades, detect fraud, schedule shuttle loading, etc.

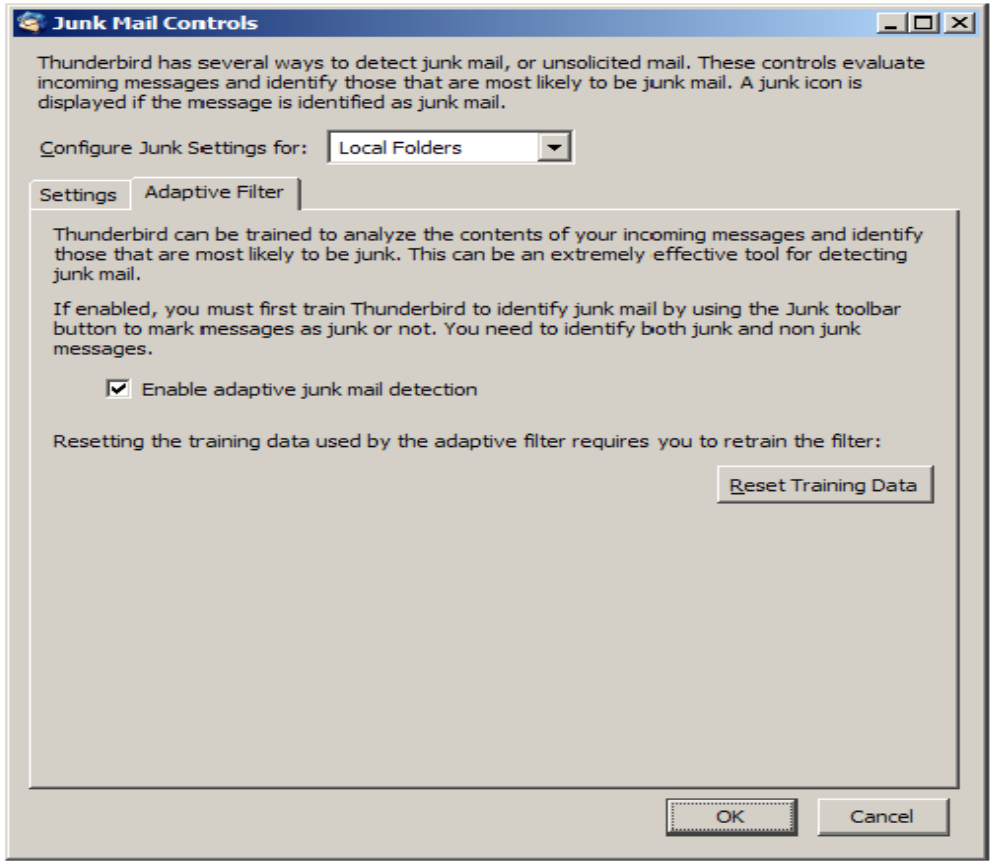
Machine Learning?

Is Machine Learning actually AI?

- What is it?
- Where does it fit?
- What is it **not**?



Some Examples – SPAM Filter



To: you@gmail.com
GET YOUR DIPLOMA TODAY!
If you are looking for a fast and cheap way to get a diploma, this is the best way out for you. Choose the desired field and degree and call us right now: For US: 1.845.709.8044 Outside US: +1.845.709.8044 "Just leave your NAME & PHONE NO. (with CountryCode)" in the voicemail. Our staff will get back to you in next few days!

ALGORITHM

Naïve Bayes

Rule mining

Some Examples – Recommendation System

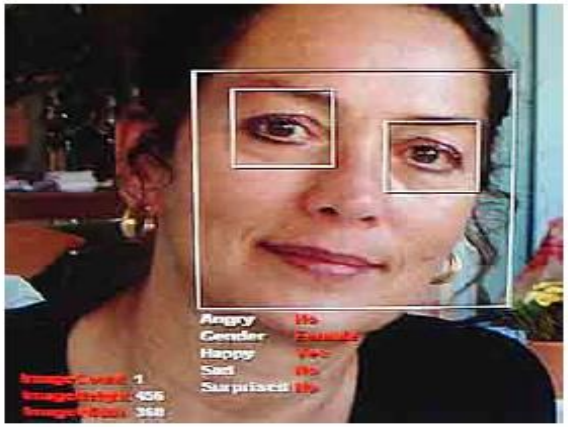
ALGORITHMS

Collaborative Filtering

Nearest Neighbour

Clustering

Some Examples – Face Detection

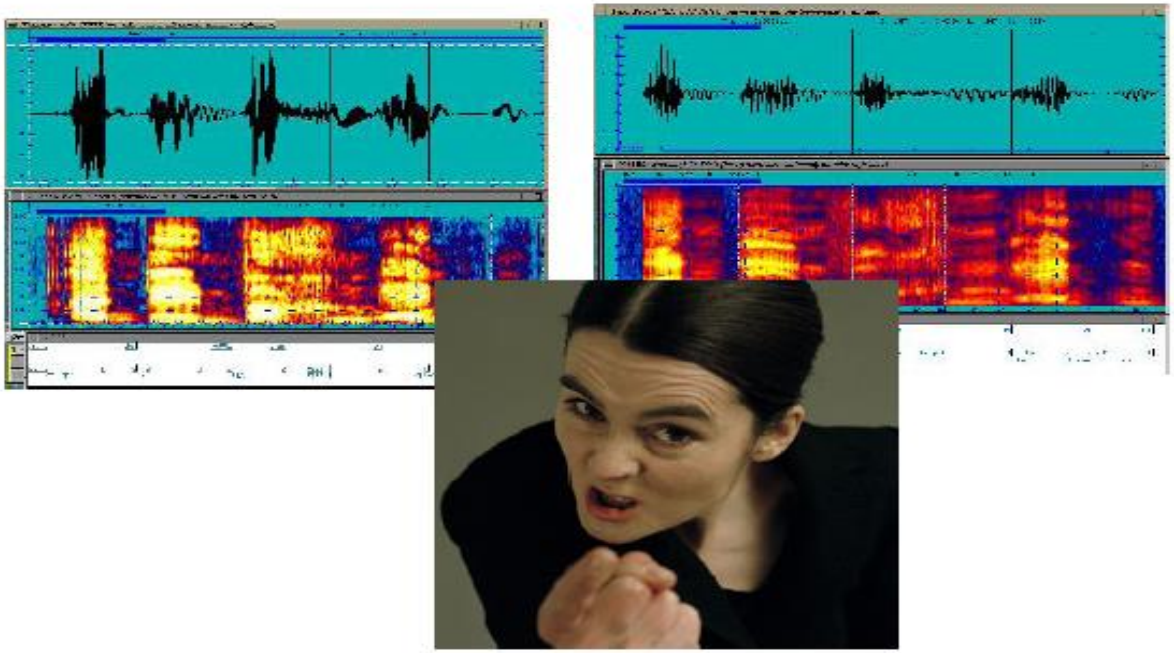


ALGORITHMS

Decision Trees

Adaboost

Some Examples – Speaker Recognition



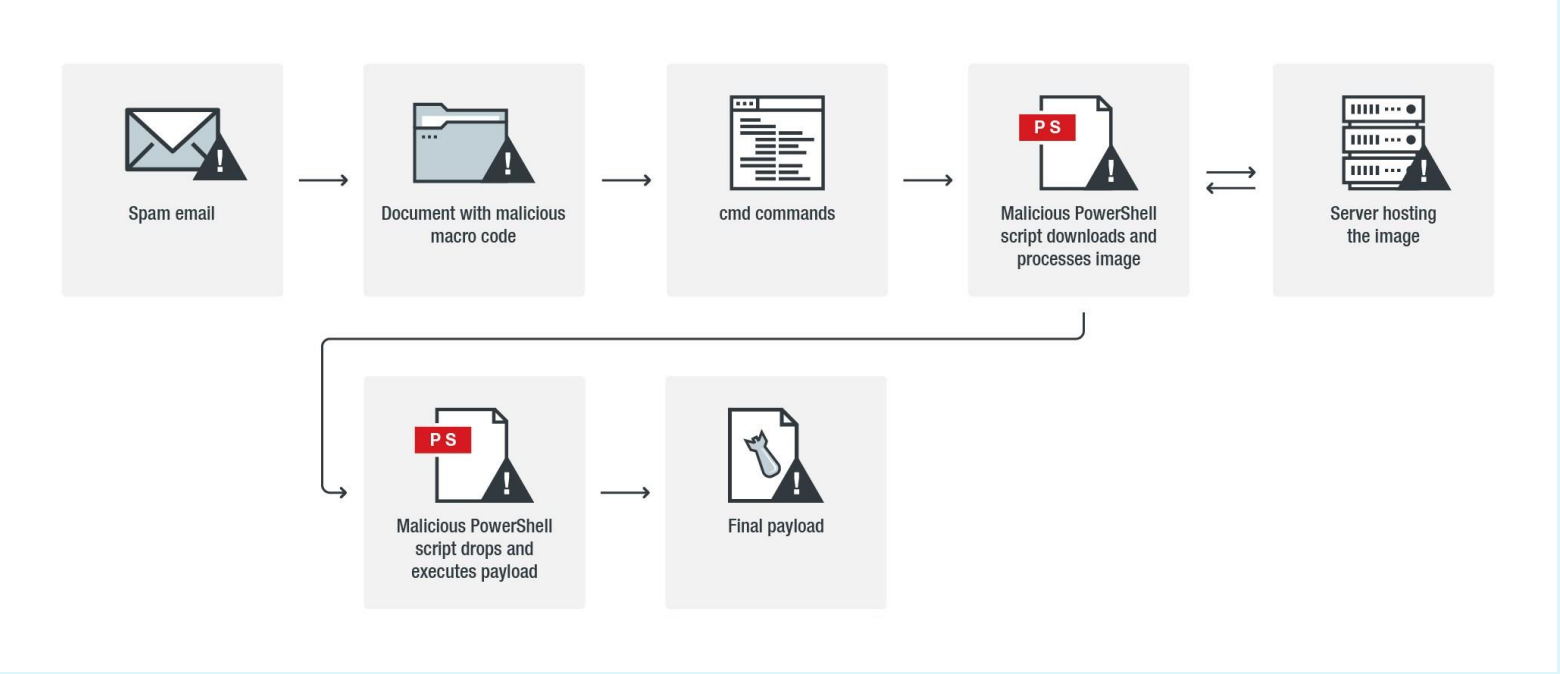
ALGORITHMS

- **Feature Extraction**
- **Probabilistic Classifiers**
- **Support Vector Machines**
- **+ many more….**

Some Examples – Hacking using AI

Machine learning based Steganography Malware

- Malware abuse uses a publicly available script called (Invoke-PSImage) that helps to *embed* malicious scripts in the pixels of a image (or video clip).
- Later attackers approaching the victims via spam email campaigns that contain a document with embedded malicious macro code.
- Attackers can use various social engineering technique lures to trick the user to download the attachment and click on it.



Some Examples – Social Media and Fraud Prevention

facebook

amazon.com

ebay

Lloyds TSB

ALGORITHMS

- Support Vector Machines
- Collaborative filtering
- Rule mining algorithms
- Many many more….

History

- 1940s
    - Human reasoning / logic first studied as formal subject within mathematics (Claude Shannon, Kurt Godel et al)
- 1950s
    - Turing Test is proposed
    - *1956 – Dartmouth Conference coins the phrase Artificial Intelligence*
    - *1959 – Arthur Samuels wrote a program that learnt to play checkers*
- 1960s
    - Funding increased for A.I. especially from Military
- 1970s
    - A.I. “winter” – Funding dried up

History (cont.)

- 1980s
  - Revival through bio-inspired algorithms: Neural Networks, Genetics Algorithms etc.
  - Rule based “expert systems” used in medical and legal professions
- 1990s
  - A.I. diverges into separate fields – Computer Vision*, Automated Reasoning, Planning System, Natural Language Processing, **Machine Learning**….
- 2000s

  ML merging with statistics continue. First commercial strength applications appear: Google, Amazon etc
- 2010 – present:
  - Emergence of very high performance machine (GPUs)
  - Deep learning

Accuracy Measures of an ML based system

Possible outcomes of a AI - Machine Learning System are:

- ***Genuine Accept***

  The genuine user was identified correctly

- ***Genuine Reject***

  The imposter user was rejected

- ***False Accept***

  Also know as “False Match” or “Type II Error”.

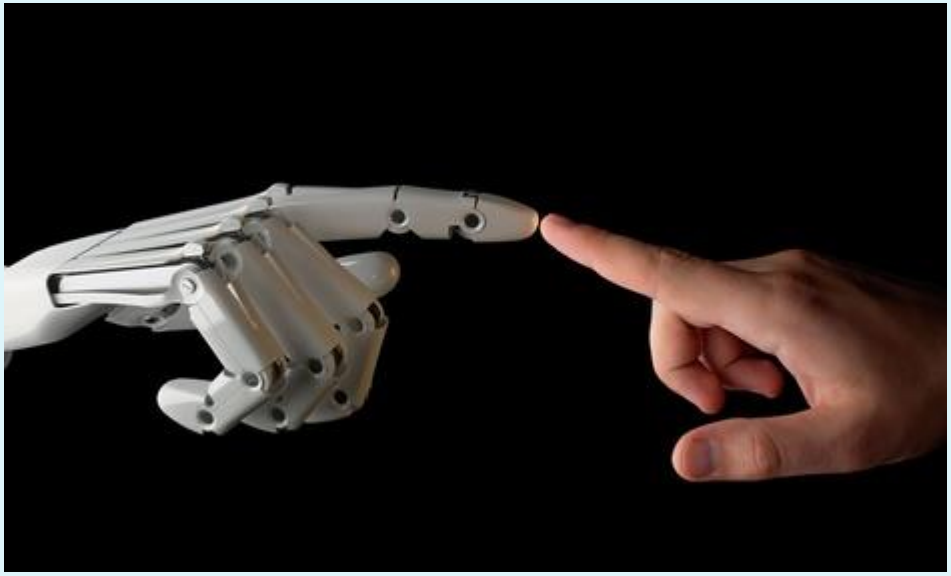
  An imposter is accepted as a genuine user.

- ***False Reject***

  Also know as “False Non-Match” or “Type I Error”.

  The genuine user was rejected as imposter.

AI:
Is it for us or against us?

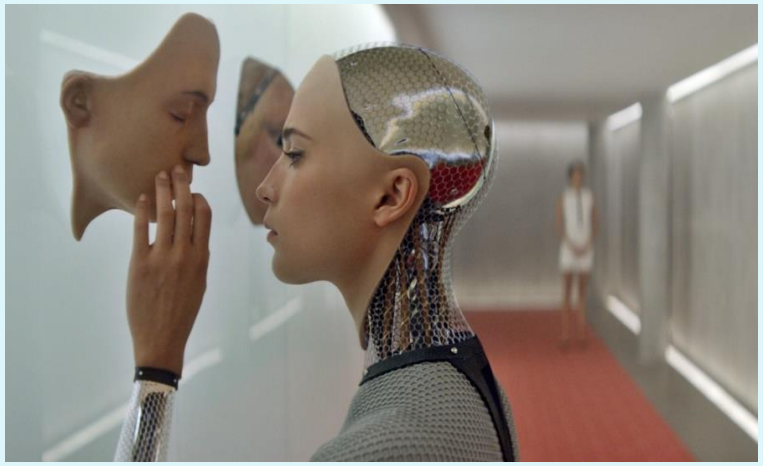


Purpose

The purpose of this presentation is to add to your confusion about artificial intelligence (AI)

Is it a good thing or a bad thing?

It’s Only the Stuff of Movies – Right?



What Some Smart People are Saying About AI

"The future is scary and very bad for people."

Steve Wozniak
Apple co-founder

AI is a "demon" that is "potentially more dangerous than nuclear weapons"

Elon Musk
Tesla chief executive

"I don't understand why some people are not concerned"

Bill Gates
Microsoft co-founder

" … full artificial intelligence could spell the end of the human race"

Stephen Hawking
British theoretical physicist

Some Other Quotes

'Eventually, I think human extinction will probably occur, and technology will likely play a part in this,' DeepMind's Shane Legg (DeepMind is part of Google)

How can an AI system behave carefully and conservatively in a world populated by unknown unknowns - Tom Dietterich, president of the AAAI

"It [AI] would take off on its own, and re-design itself at an ever increasing rate," – Stephen Hawking (on the consequences of creating something that can match or surpass humans)

“Humans, who are limited by slow biological evolution, couldn't compete, and would be superseded.“– Stephen Hawking

Take Home

Artificial Intelligence (aka Machine Intelligence) has been around for some time with no one claiming potentially dangerous consequences (outside science fiction) – so what's changed?

Are there dangers - What are they?

What can be done about them?

Artificial Enhancements

Not a new thing . . .

Strength – tractor replaced horse-drawn plow that replaced human labor

Speed – Automobile replaced the horse that replaced walking

Sight – telescopes & microscopes enhance human visual capabilities

Hearing – non-electronic amplification (e.g., gramophone) electronic amplification (electric speakers)

These are generally regarded as good things

What about enhanced intelligence?

Can a machine be intelligent? Can it "think"?

The Turing Test was introduced by Alan Turing in his 1950 paper Computing Machinery and Intelligence.

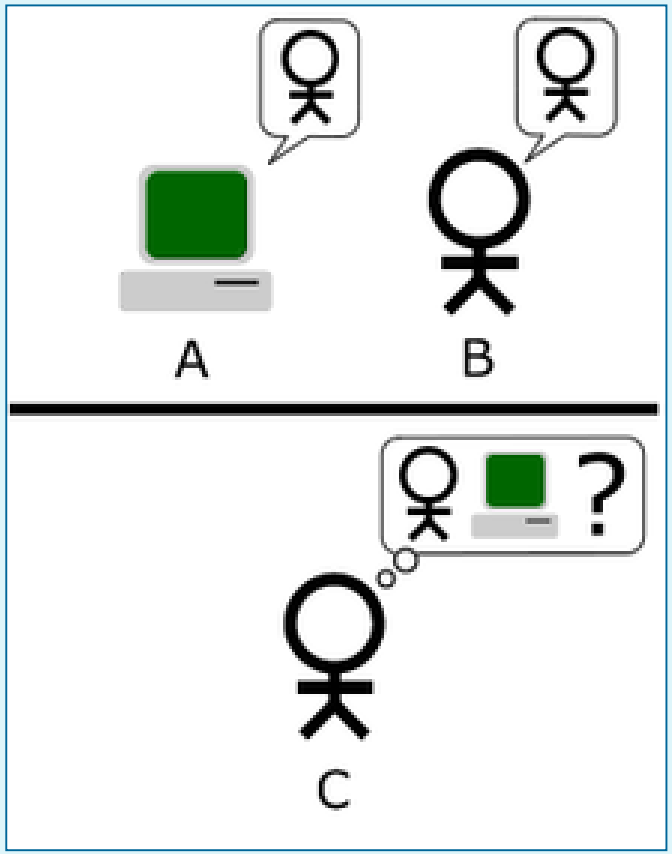
“I propose to consider the question, ‘Can machines think?’”

Since “thinking” is difficult to define, Turing chooses to “replace the question by another, which is closely related to it.”

“Are there imaginable digital computers which would do well in the imitation game?”

http://en.wikipedia.org/wiki/Turing_test

The Imitation Game



A human judge engages in a natural language conversation with one human and one machine, each emulating human responses.

All participants are separated from one another.

If the judge cannot reliably tell the machine from the human, the machine is said to have passed the test.

http://www.turinghub.com/ - take a Turing test on-line

Are We There Yet?

Cleverbot's software learns from its past conversations, and has gained high scores in the Turing test, fooling a high proportion of people into believing they are talking to a human.

http://www.cleverbot.com/

Is the Turing Test Enough?

northumbria UNIVERSITY

“. . . it's not enough to have a human be deceived for a machine to be real, The machine needs to convince the human to do things for it -- to fall in love with it, to serve its own purposes.”

- Tim Tuttle, a former MIT AI researcher and the CEO of the predictive-intelligence company Expect Labs

IBM's Deep Blue is better at chess than any human and Watson proved it could outsmart Jeopardy world champions, but they don't have any consciousness of their own.

It's worth noting that neither of those supercomputers has gone through the Turing test, though inventor and futurist Ray Kurzweil believes Watson could be retooled to pass it easily.

The Singularity

Will artificial intelligence surpass human intelligence? If so . . . When?

The “singularity” - the point in time in which artificial intelligence exceeds human intellectual capability.

Kurzweil predicts the singularity to occur around 2045, others predict some time before 2030.

https://en.wikipedia.org/wiki/Technological_singularity

What's the Problem?

Who cares if machines are smarter than people

What are the Dangers?

- Automation putting us all out of work - we will be working for robots
- Loss of human control of our lives - Robots that surpass humans in strength, speed, agility, endurance, decision making, intelligence

- Killer robots – militarization of robots (e.g. drones) with AI

- Robot emotions – will they have empathy

- Will goal seeking intelligent machines, seek the same goals as we do? Will their goals “evolve” in a negative direction?

- Everybody knows everything – the drones are watching you!

What are the Benefits?

'The potential benefits are huge, since everything that civilization has to offer is a product of human intelligence; we cannot predict what we might achieve when this intelligence is magnified by the tools AI may provide, but the eradication of disease and poverty are not unfathomable‘ – Elon Musk, Stephen Hawking

AI is Becoming Ubiquitous

Are They Taking Our Jobs?

Industries that robots will transform by 2025 [1]

from BusinessInsider.com

Automotive - 10% of cars will be fully autonomous and many will drive themselves. Japan is testing "robot taxis" for transportation during the 2020 Olympics in Tokyo.

Agriculture - Farm will increasingly use AI technology and big data analytics to optimize crop output. More driverless tractors, drones and milk bots.

Service - Personal robots will take on easy, dangerous or repetitive jobs. Mowing your lawn, cleaning your windows, washing dishes.

Are They Taking Our Jobs?

Financial - Up to $2.2 trillion in investments will be made through AI-enabled computers that can learn markets

Healthcare - Robot assistance in critical surgery, elderly care, disabled patient assistance. In 2000 there were 1,000 robot-assisted surgeries performed, with 570,000 in 2014

Manufacturing - 10% of worldwide manufacturing tasks are automated. In 10 years that will increase to 45% as robots get cheaper.

Aerospace and Defense - 90 countries now operate drones, 1/3 are armed. The number of commercial and military drones will triple over the next 5 years. Autonomous military vehicles and land robots are under development.

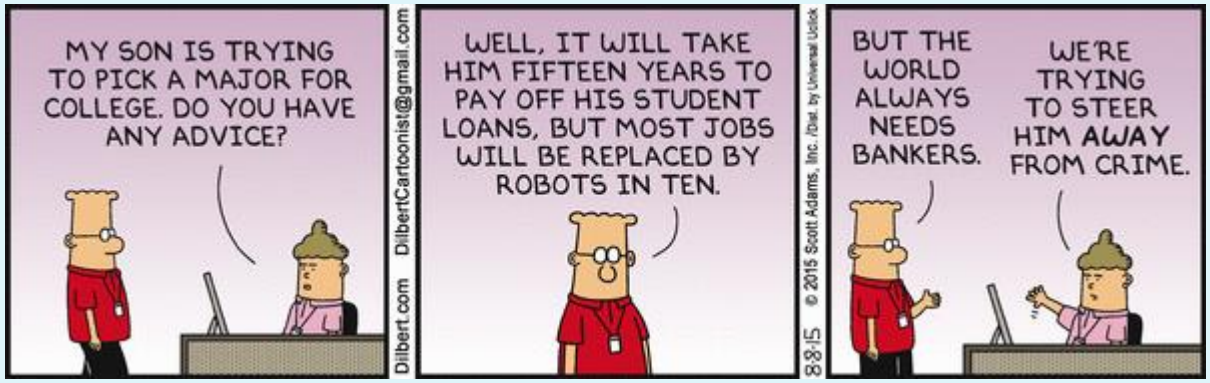
Is Your Job at Risk?

Robots could steal 80 million U.S. jobs

80 million jobs in the United States are at risk of being taken over by robots in the next few decades, a Bank of England (BoE) official warned

In a speech at the Trades Union Congress in London, the bank's chief economist, Andy Haldane, said that up to 15 million jobs in the U.K. were at risk of being lost to an age of machines, which is around half of the employed population.

Is Your Job at Risk?



Jobs with the highest level of being taken over by a machine in the U.K. included **administrative**, **production**, and **clerical tasks**. Haldane (Bank of England (BoE) official) gave two contrasting examples of risk, with **accountants** having a 95 percent probability of losing their job to machines, while **hairdressers** had lower risk, at 33 percent.

With robots being more cost-effective than hiring individuals in the workplace over the long term, jobs with the lowest wages were also at the highest risk of going to the machines.

Will We be Working for Robots?

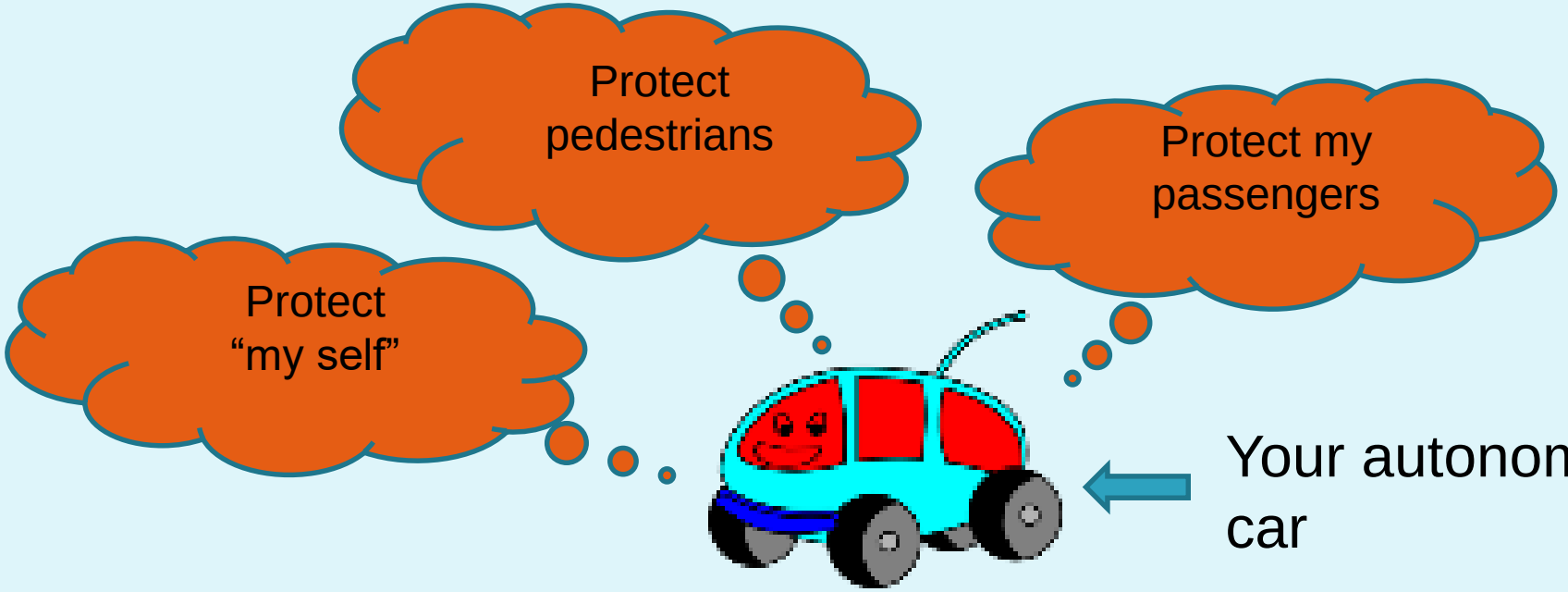
Apply now for the job of the future: "Robot helper"

AI machines can learn from experience and from the humans around them. Which means that, as AIs take on a growing role in the workplace, a new role is opening up for humans: The robot's assistant

AI trainers who work as "robot's helpers" already exist at several tech companies: **Facebook**, virtual assistant start-up **Clara Labs**, and **Interactions**, a company that builds AI to handle customer service calls.

“Man in the Loop” – Maybe Not

Can AI Machines make better decisions than us?



How well will an autonomous vehicle resolve conflicting priorities?

Should your car be making these decisions?

Smarter War Toys

- In March 2014, the Russian Strategic Missile Forces announced it would deploy armed sentry robots that could select and destroy targets with no human in or on the loop at five missile installations.

- China's Harbin Institute of Technology, unveiled at the Beijing 2015 World Robot Conference. The robots can wield anti-tank weapons, grenade launchers, or assault rifles.

The Armata (T14 tank) now requires three crew members. *"Then it will be two and then without them at all,"* Vyacheslav Khalitov, the company's deputy director said.

Autonomous Weapons

The Taranis (BAE Systems) drone

- Uses a programmed flight path to reach a preselected area.
- Automatically identifies and targets the threat within that search area.
- Sends data back to its home base, where the information is verified by the human operator.

- Human OK‘s attack and a remote pilot essentially pulls the trigger, and the Taranis fires before flying back to the base on its own

Because the Taranis is a prototype, it doesn't currently carry missiles, but future generations will likely carry weapons

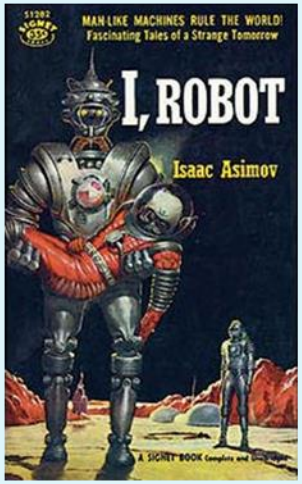
Taranis is a technology demonstrator, much like the U.S. Navy's X-47B.

What Can We Do?

Build in “defects” or “safeguards”

Or . . . Implement “The Three Laws of Robotics”

A set of rules devised by the science fiction author Isaac Asimov, introduced in his 1942 short story "Runaround”



"I Robot - Runaround". Via Wikipedia - https://en.wikipedia.org/wiki/File:I_Robot_-_Runaround.jpg#/media/File:I_Robot_-_Runaround.jpg

What Can We Do?

northumbria UNIVERSITY

Autonomous weapons are a problem

More than 16,000 AI researchers have signed an open letter to the UN, urging government leaders to take action against the creation of semiautonomous and autonomous weapons.

It's often unclear where the human comes into the decision process of targeting and firing an intelligent weapon

Starting a military AI arms race is a bad idea, and should be prevented by a ban on offensive autonomous weapons beyond meaningful human control

http://www.businessinsider.com/british-taranis-drone-first-autonomous-weapon-2015-9

The End

I thank the following for the contents of this presentation:

Dr Richard Jiang, Ismahane Cheheb, Chirine Riachy.
The document is also based on a paper from Space Coast of INCOSE: International Council of Systems Engineering

Thank you for your attention.

Any Questions?