Lecture#1

# **Course Introduction**

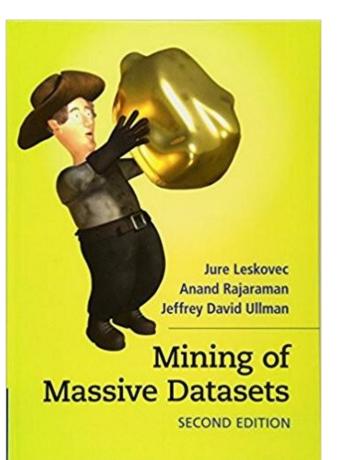
#### Course webpage

• Course Homepage:

http://aytugonan.cbu.edu.tr/ROE505\_index.html

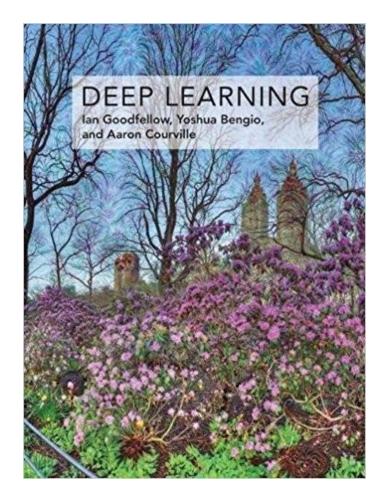
# Reading

- Available for free online
- Explains complex stuff in an easy way with lots of examples
- Covers much more than Machine Learning, so only a few chapters are relevant
- It is highly recommended to read those chapters



# Reading

- Available for free online
- Good book, but a bit mathematical
- More comprehensive than the other book
- Read the parts you find interesting



### Other good books



### Project

- The examination task is a project you define
- Select some data to work on, preferably from your daily work
- Gather, pre-process and learn from the data
- Evaluate different algorithms
- Describe your findings in a report

#### **Evaluation Criteria**

- Research Proposal: 10%
- Literature Survey: 25%
- Research Presentation: 15%
- Final Report: 50%

(Week#4) (Week#7) (Week#11-14) (Week#11)

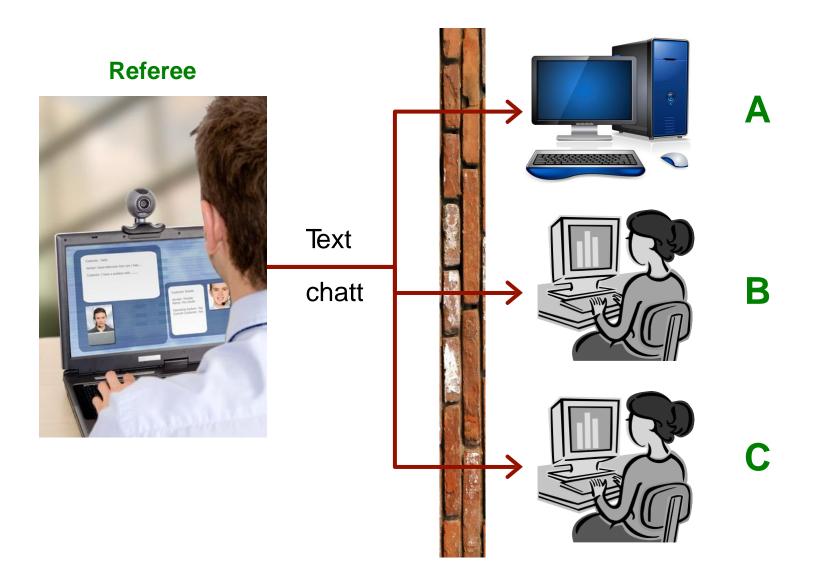
### **History of AI**

### Can Machines Think?

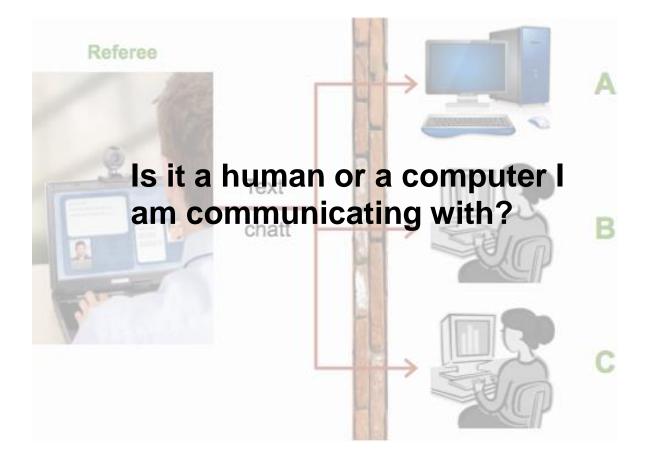
- In 1950, Alan Turing published the paper *Computing Machinery and Intelligence* in the journal *Mind*
- In the paper, Turing reasons about the question *Can machines think?*
- This paper was the first to mention the Turing Test
- The test has become an important benchmark for AI discoveries



# Turing Test



### Turing Test



#### Has it been beaten yet?

- We have come far, but no, it hasn't been beaten yet
- There are variants of the test that have been beaten
- One of the best attempts is Cleverbot:
  - <u>http://www.cleverbot.com</u>
- Depending on how the exact rules are defined, some might argue that by some metrics it has beaten the test
- In 2011, it scored 59.3% human compared to controls that were scored 63,3% human

#### Has it been beaten yet?

- Many believe that we have come so far that the Turing Test can be passed if we use average persons as judges
- It can however not yet beat judges trained to spot non-humans
- Some of the early attempts are ELIZA (1966) and PARRY (1972)

## ELIZA

- ELIZA: Please tell me what's been bothering you.
- YOU: I am scared of spiders.
- ELIZA: Is it because you are scared of spiders that you came to me?
- YOU: Yes. I need help.
- ELIZA: I see.
- YOU: Can you help me?
- ELIZA: Really, I help you?
- YOU: You are the psychiatrist.
- EUZA: What makes you think I am the psychiatrist?
- YOU: It says so on your office door.
- ELIZA: Why are you concerned over my office door?
- YOU: Because it says you are apsychiatrist and now you are telling me that you are not.
- ELIZA: Does it please you to believe I am not?
- YOU: No. Now I wonder if you can help me or not?
- ELIZA: You are being abit negative.

### ELIZA

- ELIZA tries to identify keywords in what the user typed
- If a keyword is found, the user sentence is transformed using some rules
- If a keyword is not found, ELIZA responds with a generic reply or repeats an earlier user sentence
- ELIZA does not understand the meaning of a sentence

### Alan Turing

- Turing was highly influential in formalising the theories needed to build a general purpose computer
- He also formalised the concepts of algorithms and computation with 0's and 1's
- During WW 2, he worked on the Enigma machine that was able to break the German ciphers
- Today, Turing is considered to be the father of theoretical computer science and AI

# Alan Turing

- Turing was gay, which was a crime at the time
- 1952 he was sentenced to either imprisonment or hormonal treatment
- He choose the later, which made him suffer from mental illness
- In 1954 he killed himself by eating a cyanide poisoned apple
- He was given pardon in 2013, much thanks to Neil Tennant from Pet Shop Boys



Alan Turing 1912-1954

#### Al as research field

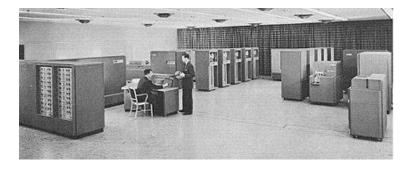
- At the Dartmouth Conference in 1956, AI was established as an academic field
- The conference was organized by Marvin Minsky, John McCarthy, Claude Shannon and Nathan Rochester
- The proposal included the following statement:
  - "every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it"
- The conference is considered the birth of AI

### The golden years 1956-1974

- During the golden years many discoveries were made, and agencies like DARPA poured a lot of money into the field
- Some of the more important ones were:
  - Solving problems by search, which lead to problem solvers in geometry (Geometry Theorem Prover in 1958) and algebra (SAINT)
  - Search was also behind the planning system STRIPS, developed to control robot behaviors
  - Natural language communication, most notably ELIZA by Joseph Weizenbaum
  - Micro-worlds, for simulations in for example physics (Marvin Minsky and Seymour Papert)

## Optimism

- Al researchers at the time were very optimistic about the field
- In 1965, Herbert A. Simon said:
  - "machines will be capable, within twenty years, of doing any work a man can do."
- In 1967, Marvin Minsky said:
  - "Within a generation ... the problem of creating 'artificial intelligence' will substantially be solved."
- In 1967, Marvin Minsky said in an interview in Life magazine:
  - "In from three to eight years we will have a machine with the general intelligence of an average human being."



### The first AI winter 1974-1980

- In 1970's, the AI field was criticized for failing to solve the problems they faced
- The optimism had raised the expectations impossibly high
- Most agencies cut the funding in AI
- The research in neural networks was almost completely halted for 10 years, due to the devastating criticism from Marvin Minsky
  - He argued that since a single neuron cannot solve non-linear functions such as XOR, larger networks will have similar limitations
  - Today we know this is completely wrong!

### The first Al winter

- Some of the problems the field faced at the time were:
  - Limited computer power a super computer at the time was capable of around 80 to 130 MIPS, and tasks such as computer vision has been shown to require 10000 to 1 million MIPS
  - Combinatorial explosion many problems could only be solved in exponential time, and moving from "toy" problems to real-world problems was not possible

### The first Al winter

- Knowledge and reasoning knowledge about the real world, for vision and natural language tasks, required huge amount of information about the world
- Moravec's paradox computers are very good at proving theorems or solve geometrical problems, but for humans simple tasks such as recognizing a face turned out to be extremely difficult
- The frame problem/qualification problem difficulty of describing and handling conditions in dynamic worlds

### Al criticism

- Early AI systems was to a large extent based on symbol processing (logic and rules)
- Several philosophers argued against the claims made by Al researchers
- Hubert Dreyfus argued that human reasoning involved very little symbol processing, but a great deal of embodied, instinctive, unconscious know-how
- John Searle put forward the Chinese room thought experiment:

#### Chinese room thought experiment

- Suppose that we create an AI that behaves as if it understands Chinese
- It takes Chinese characters as input, and produces a response with other Chinese characters
- Suppose that this AI is so good that it passes the Turing Test
- The question is then:
  - does the machine understand Chinese, or does it merely simulate the ability to understand Chinese?
- He argued that if the symbols have no meaning for the machine, the machine cannot be described as *thinking*
- Searle called understanding strong AI, and simulating the ability to understand weak AI

### Al criticism

- Many AI researchers did not take the criticism seriously
- Marvin Minsky, for example, said about Dreyfus and Searle that:
  - "they misunderstand, and should be ignored."
- Weizenbaum, who created ELIZA, did not agree with Dreyfus but said that his colleagues' treatment of Dreyfus was unprofessional and childish

### AI boom 1980-1987

- In 1980s, companies successfully began to use expert systems
- An expert system answers questions or solves problems about a specific knowledge domain, using logical rules
- The rules are derived from the knowledge of domain experts
- Expert systems were part of a new direction in AI

### Knowledge revolution

- During the 1970s, researchers began to suspect that
  - "... intelligence might very well be based on the ability to use large amounts of diverse knowledge in different ways" (Pamela McCorduck)
- Intelligent behavior depends on dealing with knowledge, sometimes very detailed knowledge, of a domain
- Knowledge based systems and knowledge engineering became a major focus of AI in the 1980s

#### Neural Networks

- In 1982, John Hopfield proved that a form of neural networks (called Hopfield nets) could learn and process information in new ways
- Around the same time, David Rumelhart popularized backpropagation for training neural networks
  - First published in 1974 by Paul J. Werbos
- These two discoveries revived research in neural networks
- It was however not until 1990s when neural networks become commercially successful for tasks such as optical character recognition and speech recognition

### The second Al winter 1987-1993

- Unlike the first AI winter, advances were still made in the field during this period
- The collapse was more about how the field was viewed by government agencies and investors
- Once again, expectations had been higher than was actually possible
- A goal listed in Japan's Fifth Generation Project for 1981 was to "carry on a casual conversation"
- This has, arguably, not been met today

### The second Al winter 1987-1993

- The interest in expert systems were fading
- They proved to be very difficult and expensive to maintain and update
- They could also make huge mistakes when given
  new and unknown input
- They were still useful, but only for a few very specialized tasks

#### Modern Al 1993-

- During this period, the AI field was able to achieve some important goals
- AI systems were also successfully used in technology industry, mainly due to a focus on isolated problems and increasing computer power
- The repuation of AI was now very bad, and many discoveries and achievements were made somewhat behind the scenes
- Al researchers also did not agree about the reasons for earlier failures, and the field became fragmented into subfields

### Milestones

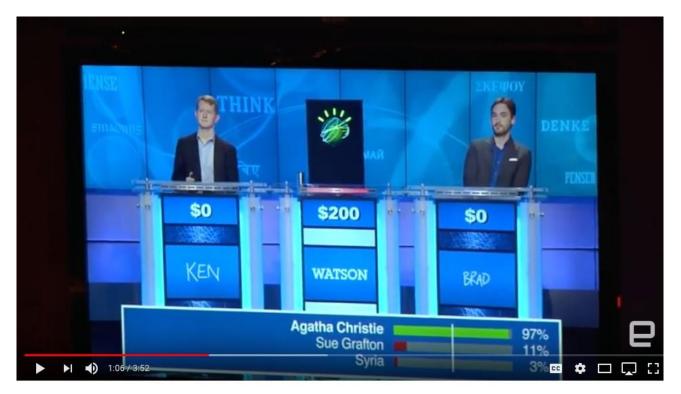
- On May 11 1997, Deep Blue became the first computer system to beat a world champion, Garry Kasparov, in Chess
- Deep Blue searched for all combinations of future moves from the current board state
- It could easily search 14 moves in the future, sometimes up to 40 moves
- Deep Blue was custom built for Chess playing, and could process around 200 million moves per second
- The event was also broadcasted live on the Internet and had over 74 million viewers



### Milestones

- In 2005, a Stanford robot won the DARPA Grand Challenge by driving 200 km on a (previously unknown) desert trail
- Two years later, a team from Carnegie Mellon University won the DARPA Urban Challenge by navigating 88 km in an urban environment while avoiding hazards and adhering to traffic laws
- In 2011, IBM developed the Jeopardy playing system Watson:





IBM's Watson Supercomputer Destroys Humans in Jeopardy | Engadget

https://www.youtube.com/watch?v=WFR3IOm\_xhE



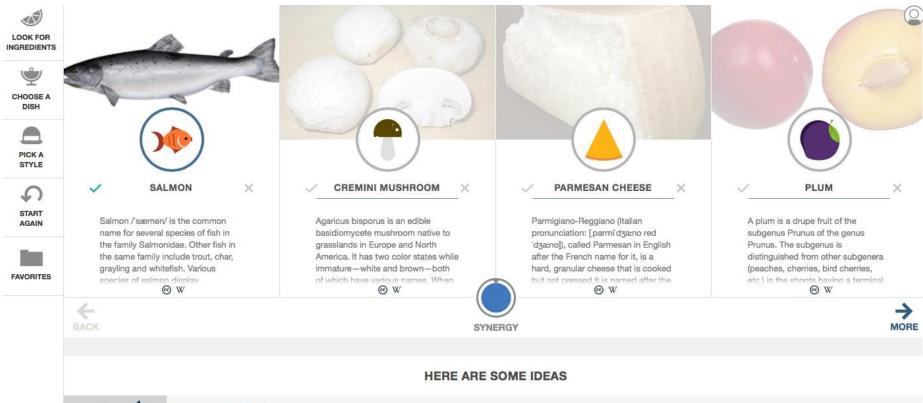
- Watson is a question answering (QA) system
- The difference between QA and document search is that a QA system tries to understand the question in detail, and returns a precise answer to the question
- IBM states that:
  - "more than 100 different techniques are used to analyze natural language, identify sources, find and generate hypotheses, find and score evidence, and merge and rank hypotheses."



- Some of the techniques used are natural language processing, information retrieval, knowledge representation, automated reasoning, and machine learning
- Watson had access to millions of documents from Wikipedia, dictionaries, encyclopedias, newspapers, etc.
- An innovation was that instead of using a single algorithm for language analysis, Watson executed hundreds of proven algorithms in parallel and evaluated the responses
- The more algorithms that returned the same or similar answer, the more sure Watson was of being correct



- Watson has been used successfully in other domains such as:
  - Healthcare as a decision support system for medical professionals when treating e.g. cancer
  - Teaching assistant used in electronic teaching material to provided natural language, one-to-one tutoring to students
  - Watson was also used as TA at Georgia Tech, where it answered questions it was 97% certain the answer was correct, leaving the rest of the questions to human teaching assistants
  - Cook book Chef Watson suggests unique dishes based on what you have in the fridge!
  - Weather forecasting Deep Thunder project started at IBM 2016
  - ... and more



#### Salmon Stuffed Vegetable

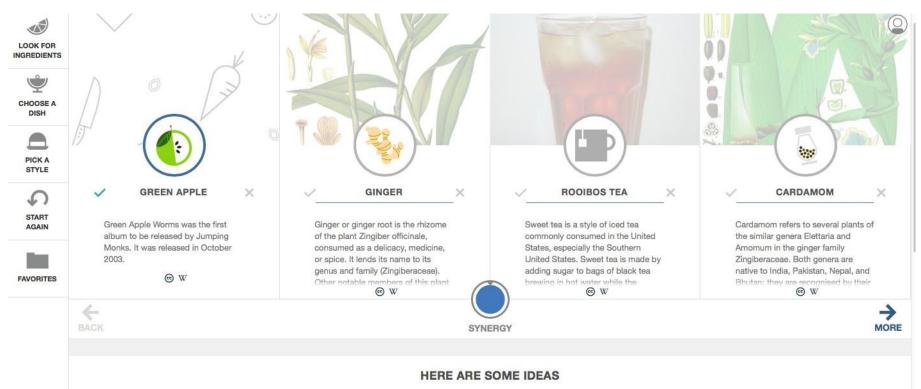
water, barley malt syrup, granulated sugar, vinegar, salmon, cremini mushroom, extra-virgin olive oil, plum, parmesan cheese, cilantro, harissa, coriander seed, pink peppercorn, anise

MORE ....

Based on: Chipotle Chiles Stuffed With Shrimp And Plantains from Bon Appétit

#### Salmon Pasta

tortellini, salmon, cremini mushroom, plum, parmesan cheese, herbes de provence, kosher salt, red pepper flakes



#### Green Apple Cake

confectioners' sugar, greek honey, egg, bicarbonate of soda, raisin, butter, ginger, flour, green apple, kosher salt, rooibos tea, whole grain mustard, yellow mustard seeds, caraway seed, cardamom, white pepper

Based on: Ginger Spice Cake With Dried Cherries from Bon Appétit



#### Green Apple Buche De Noel

water, baking powder, butter, parmesan cheese, heavy cream, clove, cardamom, ground cinnamon, granulated sugar, orange blossom honey, egg, dried cherry, angostura bitters, balsamic vinegar, cake flour, lemon juice, green apple, salt, ginger, rooibos tea

MORE ...

### Using Watson

- Watson has an API that can be used to create interesting applications
- It is also possible to submit documents for Watson to learn from
- Stena Line is for example using Watson to analyse CVs and contracts, and their chattbot Stina is based on Watson

### Intelligent agents

- Another AI paradigm that has successfully been used for many tasks is intelligent agents
- Agent systems became popular during the 1990s
- An intelligent agent is an entity that perceives its environment, and takes actions to maximize its chances of success
- Especially multi-agent systems, where multiple agents collaborate, have solved tasks such as optimizations in electrical grids and distributed heating systems

### Mathematical foundation

- In early 2000s, AI researchers began to develop and use mathematical tools more than in the past
- The researchers in the field realized that many problems Al faced, were already being worked on in fields such as mathematics and economics
- The shared mathematical language allowed for more cross-discipline collaborations, leading to many discoveries
- Russell and Norvig even called it "... nothing less than a revolution"

#### Mathematical foundation

- In 1988, Judea Pearl wrote a highly influential book that brought probability and decision theory into AI
- It lay the foundation for Bayesian networks, hidden Markov models, information theory, stochastic modelling and optimization
- Mathematical descriptions were also developed for neural networks and evolutionary algorithms

#### Al behind the scenes

- All has solved a lot of problems, and the solutions have proved to be useful in the tech industry, for example in:
  - Industrial robotics
  - Logistics
  - Banking software
  - Medicine
  - Search engines such as Google
- The algorithms and solutions are however often only a part of larger systems
- The AI field received little or no credit for these achievements

#### Al behind the scenes

- Al became a tool among others in computer science
- Nick Bostrom said in CNN 2006 that:
  - "A lot of cutting edge AI has filtered into general applications, often without being called AI because once something becomes useful enough and common enough it's not labeled AI anymore."
- Due to the history of AI not able to deliver promises, many researchers called their work other names such as knowledge-based systems, cognitive systems or computational intelligence

#### Where is HAL 9000?

- In the movie 2001: A Space Odyssey from 1968, Arthur C. Clarke and Stanley Kubrick predicted that in 2001 a machine would exist with an intelligence matching or surpassing humans
- The machine, HAL 9000, was based on beliefs shared by many leading AI researchers at the time
- Obviously, we didn't get a HAL 9000 by 2001
- But why?



#### Where is HAL 9000?

- Marvin Minsky said it was because the central problems, like commonsense reasoning, were being neglected while research was focused on commercial applications
- Ray Kurzweil referred to Moore's Law (computer power will double every second year) when he argued that we will have enough computational power to create human-level intelligence in 2029
- Jeff Hawkins, the founder of Palm Computing, argued that the research ignores the essential properties of the human cortex, preferring simpler models able to solve simpler problems
- ... these are just a few of the explanations put forward by researchers

#### Big Data and Deep Learning

- In a New York Times article 2016, the market for AI related software and hardware was claimed to have exceeded 8 billion dollars in the US, and the interest in AI had reached a frenzy
- Today, we have access to huge amounts of data, very fast parallel computer systems and advanced machine learning techniques
- Specialized deep neural networks, most notably convolutional and recurrent neural networks, boosted progress in image and video processing, text analysis, speech recognition, and more

#### Chess and Go

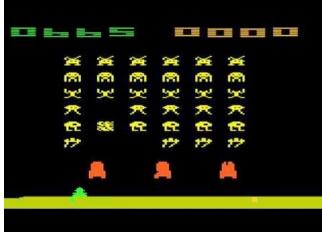
- In Chess, there is on average around 35 possible moves from each board state
- The branching factor of the search tree (without pruning) is 35
- Deep Blue was able to search at least 20 moves in the future, using smart pruning
- The game Go is huge compared to Chess, with a branching factor of 361
- Until recently, AI systems were only able to play the game on amateur level



- AlphaGo is a Go playing system developed by DeepMind, a Googleowned company
- In October 2015, AlphaGo was able to beat the 18 times world champion Lee Sedol by 4 games to 1
- AlphaGo uses a deep neural network to learn knowledge about the game, in a 40 days learning phase
- The system generates versions of itself to play against, to generate new data
- During gameplay it uses sophisticated search methods to find its moves from the learned knowledge
- AlphaGo is by many considered a major milestone in Al research

### Playing Atari games

- Before AlphaGo, DeepMind created a deep reinforcement learning system that played seven Atari 2600 games
- The system had not seen the games before, and was only given screenshots as input
- With a trial and error approach, the system gradually learned how to play the games
- It outperformed other computer systems on 6 out of 7 games, and surpasses human experts on 3 of the games



#### Cats on Youtube

- With a deep unsupervised network running on 16000 processors, Google learned to detect cats on Youtube clips
- The clips were not labelled, but the system was able to group clips with cats together
- The accuracy was 74.8% on identifying cats, and 81.7% on human faces

## Big Data

- Ray Kurzweil said that:
  "one of the strengths of humans is that we can learn from only a few examples"
- Machine Learning in general require many examples to learn properly
- ... and the more attributes and categories we have, the more data we need
- It has been said about deep learning that "life starts at a billion examples"

#### Al in the future

- In January 2015, the Future of Life Institute organized a private conference
- FLI is run by Max Tegmark, a Swedish cosmologist who got famous for writing a book about the universe being a simulation, and Jaan Tallinn, co-founder of Skype
- FLI has an advisory board with many famous names:
  - Actor Morgan Freeman
  - Nobel prize winner in physics, Alan Guth
  - AI-expert Stuart Russell
  - Elon Musk (Tesla/SpaceX)
  - Nick Bostrom, director of Future of Humanity Institute
  - Stephen Hawking, before his passing
  - ... and some more



- The attendees were asked to first predict when machines would become better than humans at all human tasks
- The median answer was by 2050
- The second question was if they thought this breakthrough could be a very bad thing
- Many famous names have talked about the risks of AI:
  - Steve Wozniak, Bill Gates, Elon Musk, Stephen Hawking, ...
  - Musk even said that "With artificial intelligence, we are summoning the demon."
  - Stephen Hawking had similar concerns: "the development of full artificial intelligence could spell the end of the human race."
- ... but why are we afraid of AI?

- Today, AI can do some tasks at human- or near-human level
  - Self-driving cars
  - Image/object recognition
  - Cooking
  - ...
- This raises some issues regarding:
  - Ethics what are the ethical implications for bringing AI into society?
  - Laws and regulations who is responsible when AI is doing harm, like when the Über autonomous car killed a pedestrian?
  - Economy jobs disappear with robots and AI systems, and the people who build the systems will thrive?

- The economical implications are already discussed jobs will disappear or change
- We will need more educated engineers and less factory workers
- Ethical and legal issues about self-driving cars and autonomous military drones are also discussed
- The concerns about the risks of AI is if, or when, we can create artificial general intelligence
- This means that we have created an AI that is at the same level or better than humans at all human tasks
- What will happen to our society then? Are humans even needed?

- If we have managed to create an artificial general intelligence, it means that it will be at our level or even better at designing new Als
- The progress will explode, leading to the creation of an artificial super intelligence
- The ASI is vastly superior to humans at all tasks!
- Now, things can get ugly... Terminator 2 here we come!



- Nick Bostrom, born in Sweden but now professor in philosophy at Oxford, is famous for his discussions about the future of AI
- He has written the book "Super intelligence: Paths, Dangers, Strategies" about the dangers of Al
- According to Bostrom, humanity can survive bombs, asteroids and diseases, but AI has the capacity to erase the entire planet
- Let's think about some of his thought experiments:



- We create a psychiatrist robot with the goal of making humans happy
- The robot is an artificial general intelligence, capable of all tasks humans can do
- It also has access to all things humans in its environment has access to
- The robot is only concerned with its goal, and how to effectively achieve it
- It is then very likely that the robot thinks the best way is to give humans loads of drugs, instead of taking the long route of therapy

- You are, as always, late for work and the kids are home alone with your domestic robot
- The kids are hungry, but there's no food in the fridge
- The robot then turns to ... the cat
- It has not yet learned human values: the emotional value of the cat outweighs its nutritional value
- It is only concerned with fulfilling the goal of giving the kids food



- Why don't we just switch the robot of when it behaves badly?
- The robot wants to achieve its goal: give the kids food
- It cannot do this when its dead
- So it will, most likely, learn to disable the Off button



- The difficulty is that we need to build human ethics, values and moral into AI systems
- This is obviously very difficult since it is very hard to define what these actually are...
- We must also make sure that humans are in control of the AI systems, and not the other way around
- But who should have that control over super intelligent Al systems?
- And can we be in control (the off-button problem)?

#### Tay



- In March 2016, Microsoft released the Twitter bot Tay
- Tay was designed to be like a witty teenaged girl
- Tay learned from Twitter and other social media
- In less than 24 hours, Tay became a racist, holocaust denying bot
- This was the effect of people tweeting the bot with such remarks
- After a day, Tay was shut down
- Tay was an interesting example of how AI can go in directions the engineers didn't thought about

# Tay



# Tay

11m

10 2



17 1 • ...

#### Think about...

- Will we be able to design an artificial general intelligence in the near future?
- If we can, how can we build human ethics, values and moral into it?
- Will the AGI eventually design an artificial super intelligence that can be a threat to humanity?
- Or, are the obstacles of designing an AGI simply to complex so we will likely not be able to create one?

#### More information...

- On the course webpage under the Media section, there are some interesting videos to watch from Nick Bostrom, Stuart Russell and Sebastian Thrun about the future of AI
- I also recommend reading the Wait but why article:
  - The AI Revolution: The Road to Superintelligence

#### **Examples of AI in industry**



- TGI Fridays is an American bar and grill restaurant, located in several countries (there is one in Malmö)
- They created a chat platform based on machine learning
- Customers can, for example, chat about what they should have for dessert, book tables, order take-away in both social media and with Amazons Alexa (voice service, similar to Siri)
- It also helps TGI Fridays in identifying which social media platforms they should focus on





- Elsevier is a publisher for scientific books and journals based in Netherlands
- They are investigating and using AI for security systems, chat bots for customer services and information management systems
- The latter automatically identifies relevant content in documents, and deliver it to the right person
- They have also extracted medical information from images taken over several decades, classify it and make it searchable



- Capital One is an American bank
- They have integrated Amazon Alexa and machine learning into their regular customer service systems
- Customers can now do all kinds of banking errands over phone
- Capital One is also investigating if AI can be used for:
  - Analyzing customer service calls to see how customer service can be improved
  - Detect fraud
  - Analyzing the housing market to understand why certain areas are more popular than others



- The ferry company Stena Line uses AI in several services
- Their customer chat bot is based on Watson for natural language understanding, combined with other customer service tools
- They also use Watson for mining and organizing information from CVs, and analyzing written agreements
- They are also working on using machine learning for setting prices on products onboard
- A deep neural network is also used to find misspelled names in bookings





- Over 1 billion people use Kone's elevators and escalators every day
- Kone has started a project to use IoT to connect all units
- All gathered data is analyzed with Watson
- The analyzed data not only tells when units are malfunctioning, but also how they are used, how they feel and if there if a risk for problems in the near future
- The data is analyzed and visualized in real-time

### Libratus

- Designing a Poker bot is notoriously difficult
- There are many aspects involved in the game such as bluffing, incomplete information and randomness
- Libratus, a Poker bot developed at Carnegie Mellon University, won in 2017 against four top professional players
- The tournament lasted 20 days and around 120000 poker hands were played
- Libratus uses a knowledge database of 2600 TB!
- As with AlphaGo, Libratus played against itself to generate more data

