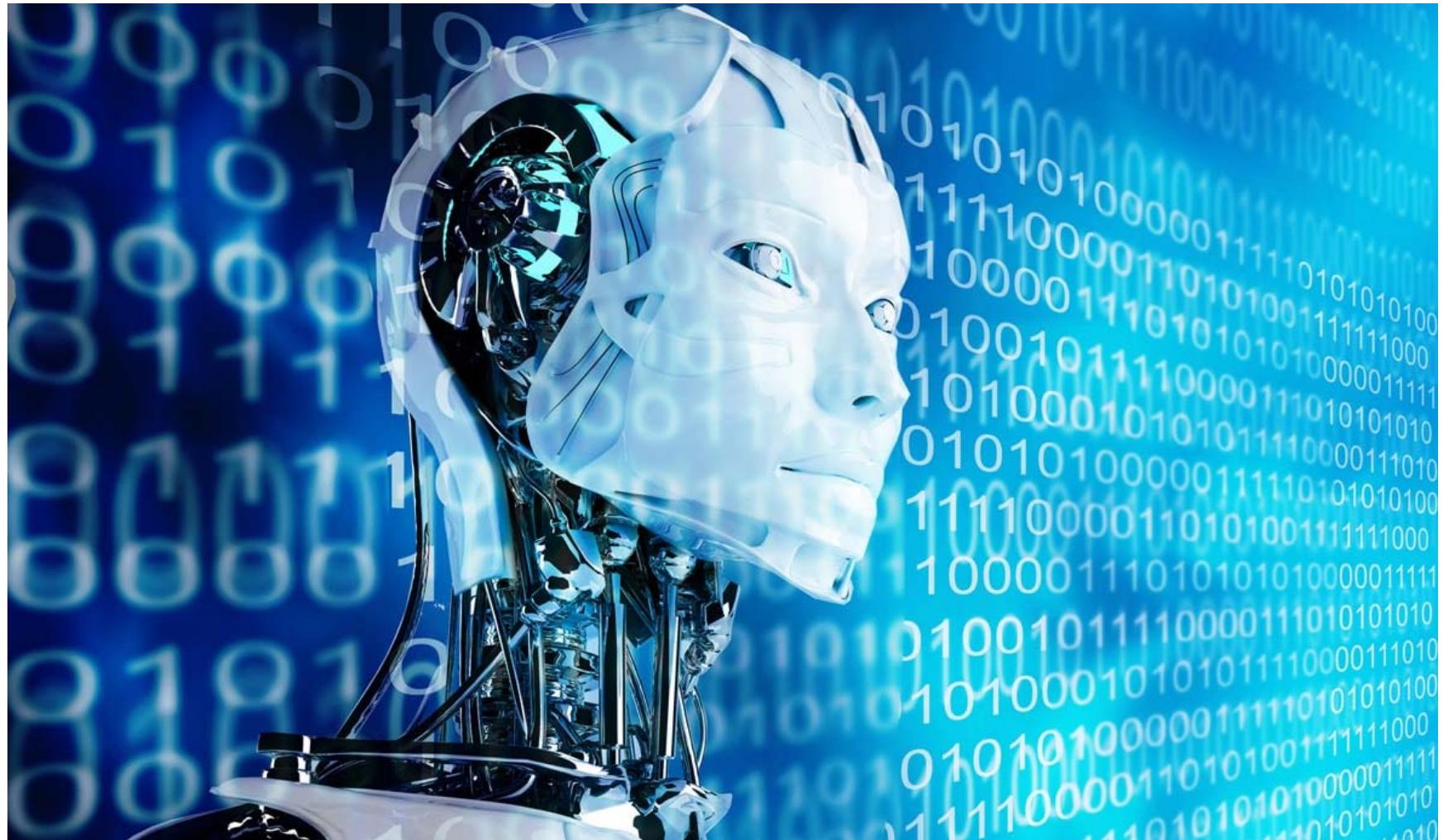




Understanding AI



Artificial intelligence



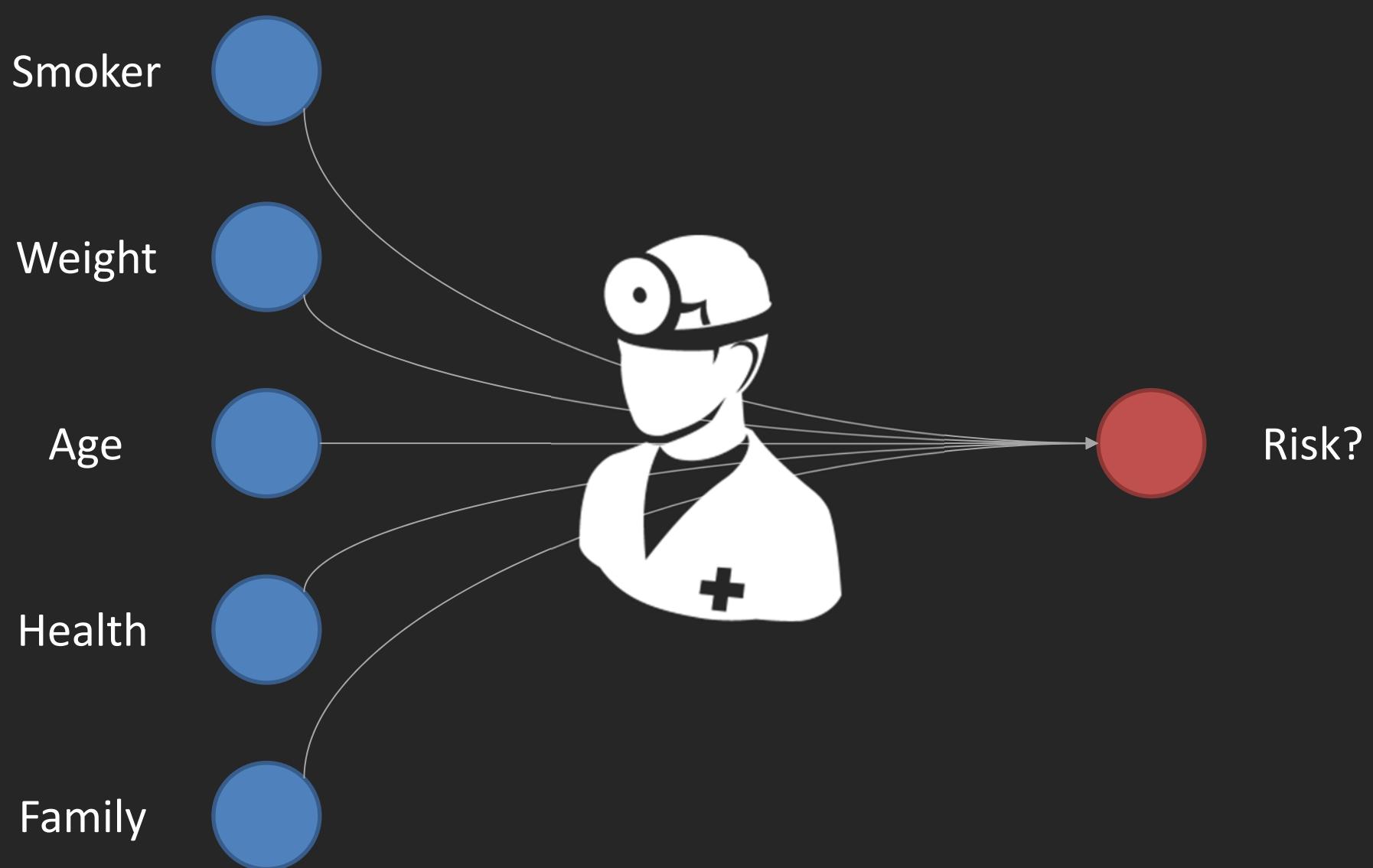
Agenda

1. Supervised learning
2. Reinforcement learning
3. Society and AI

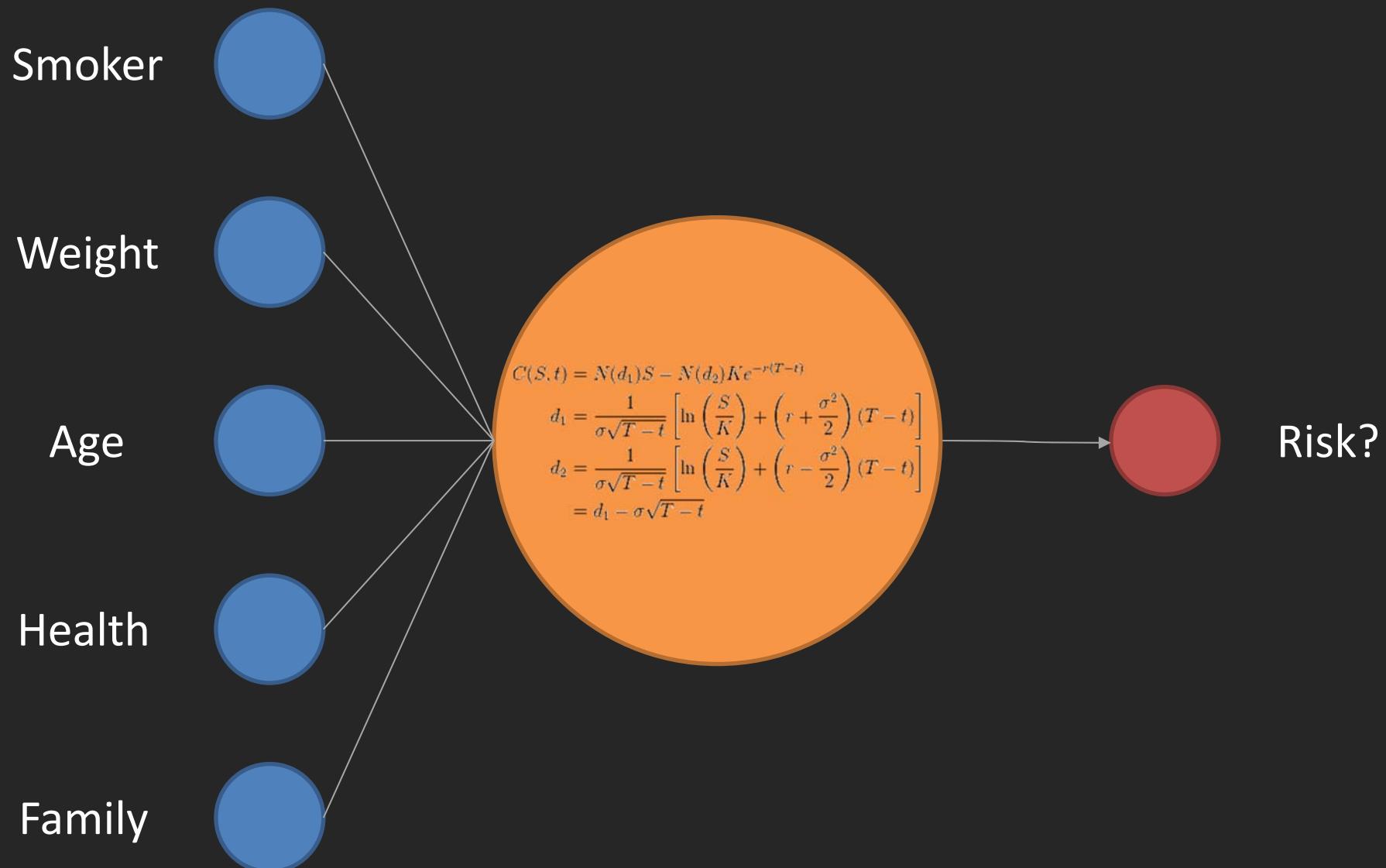


SUPERVISED LEARNING

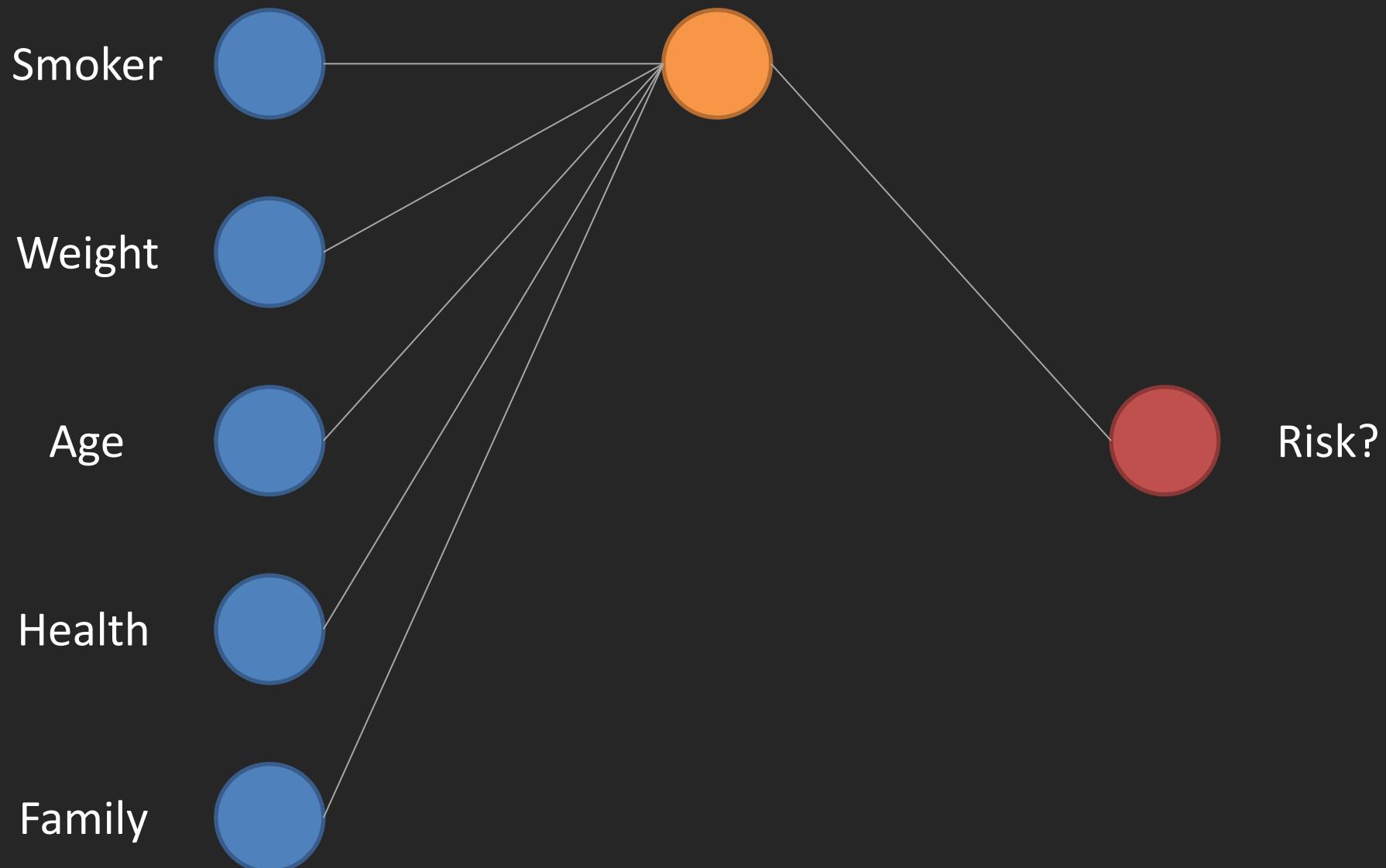
Heart incident



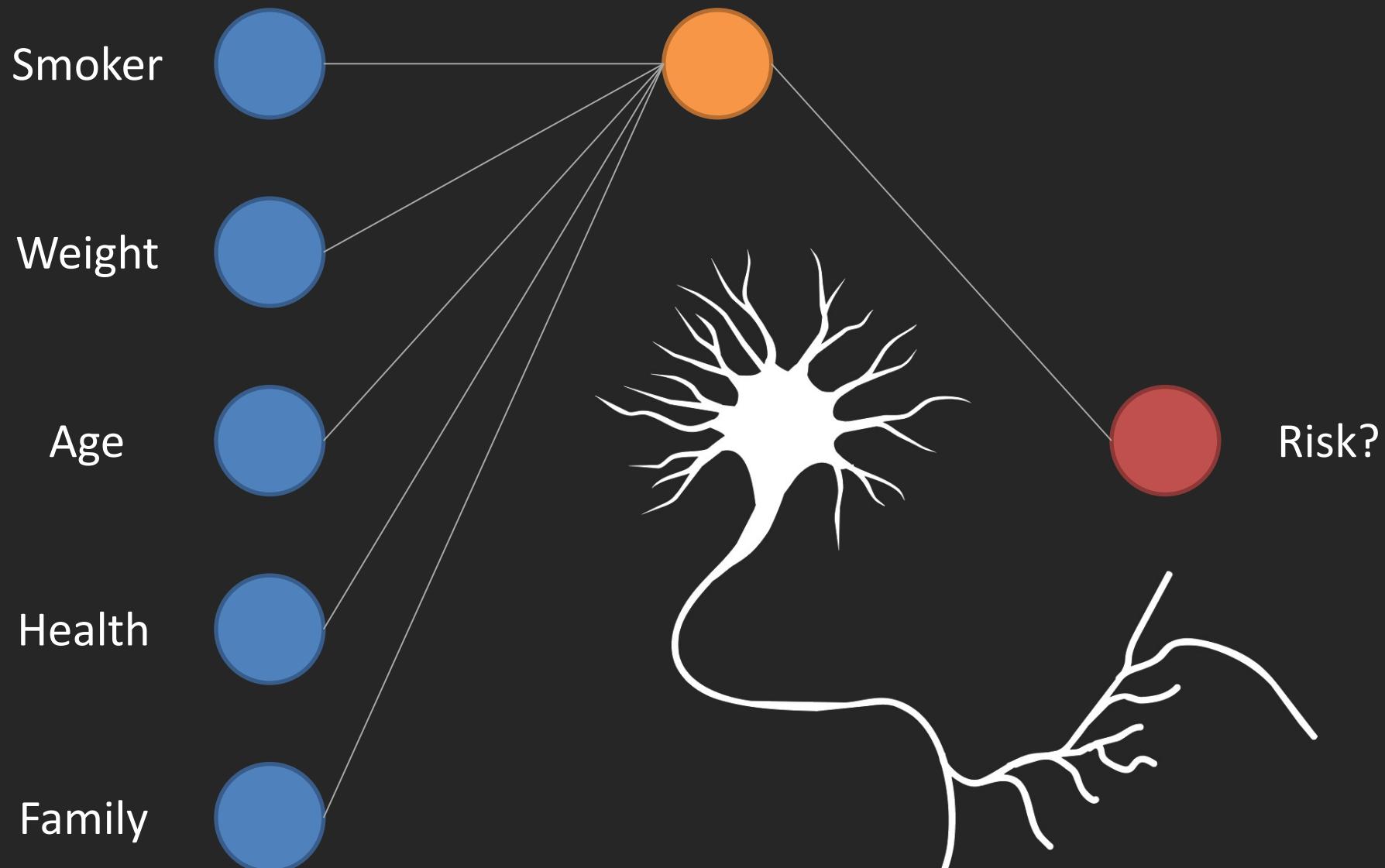
Mathematical model



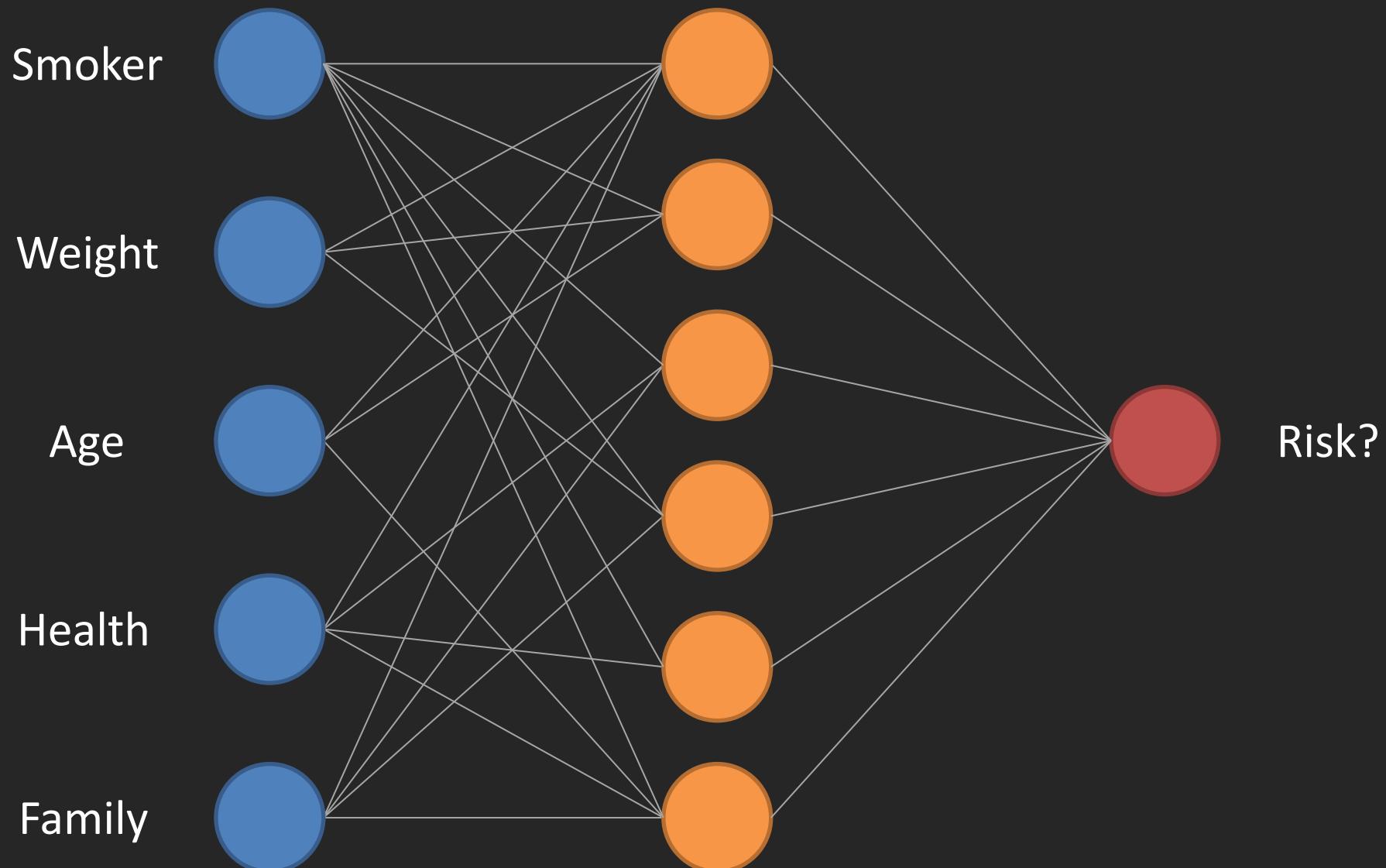
Neural network (1975)



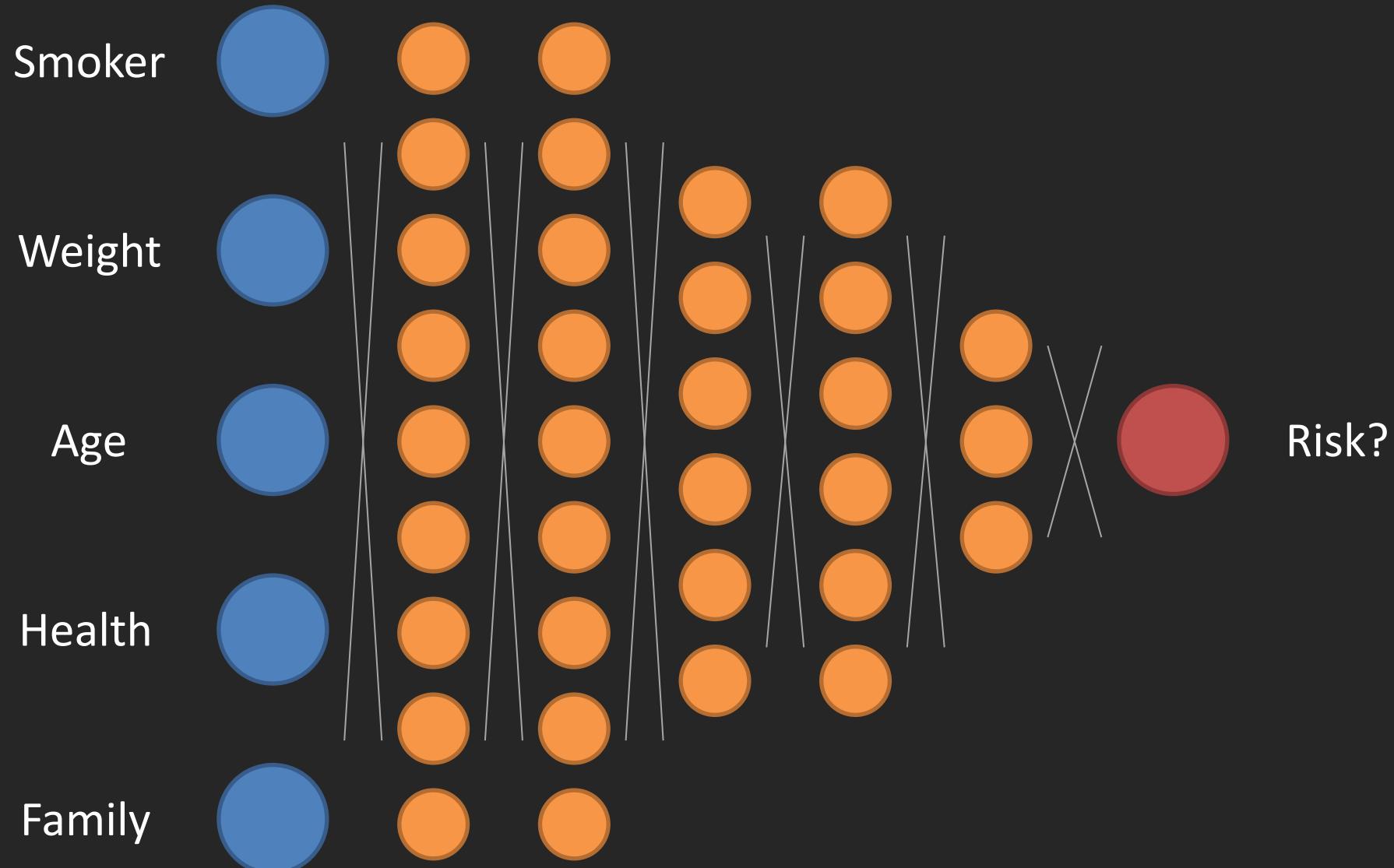
Neural network (1975)



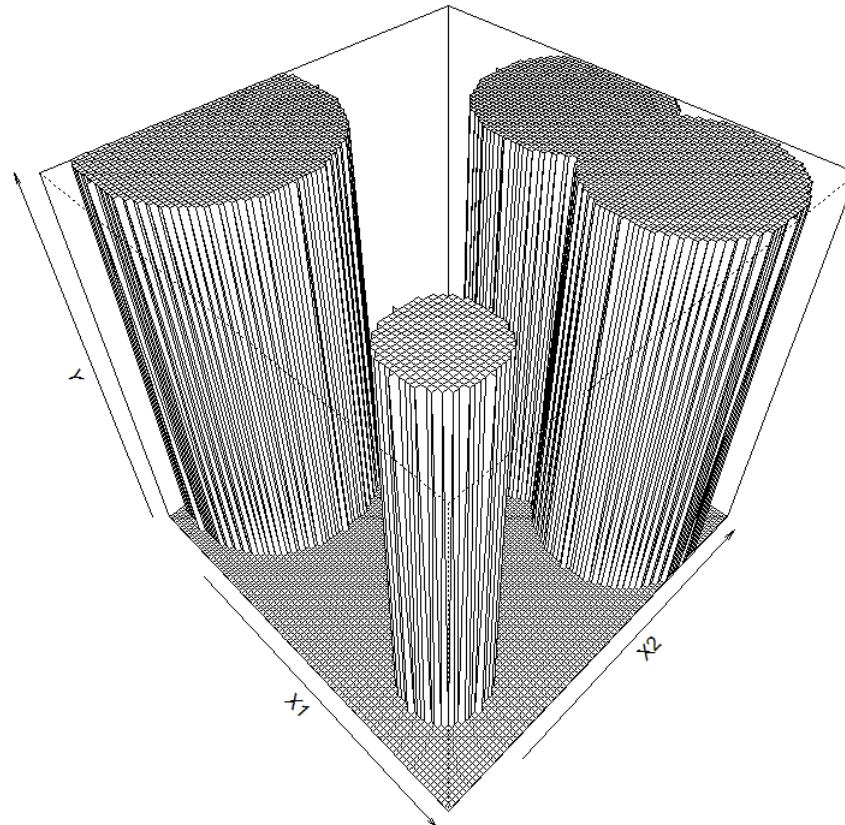
Neural network (1975)



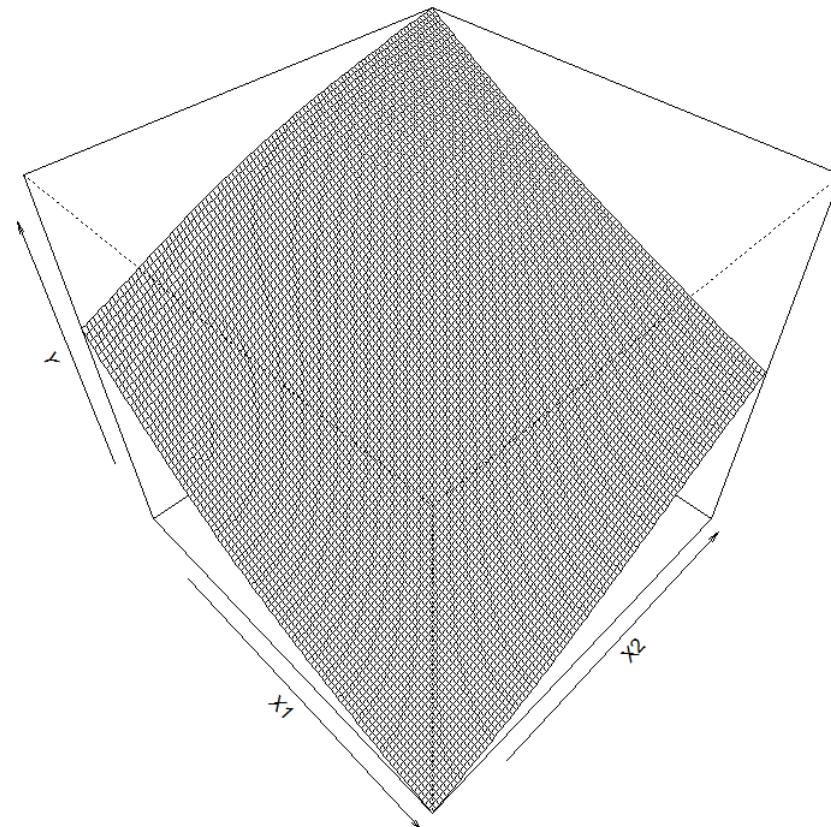
DEEP neural network (2012)



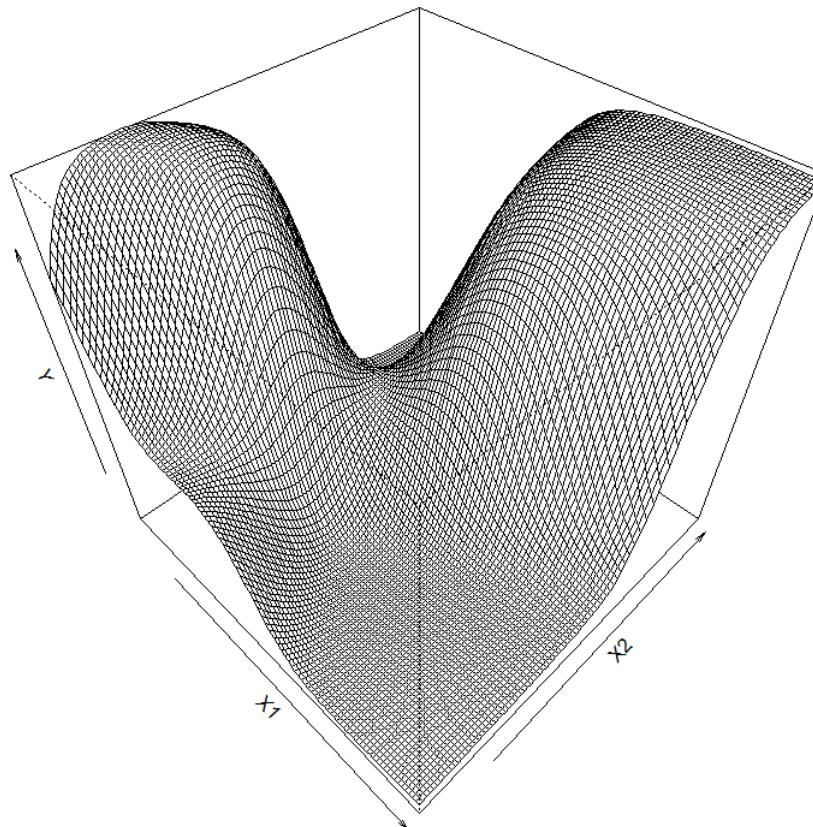
Learning



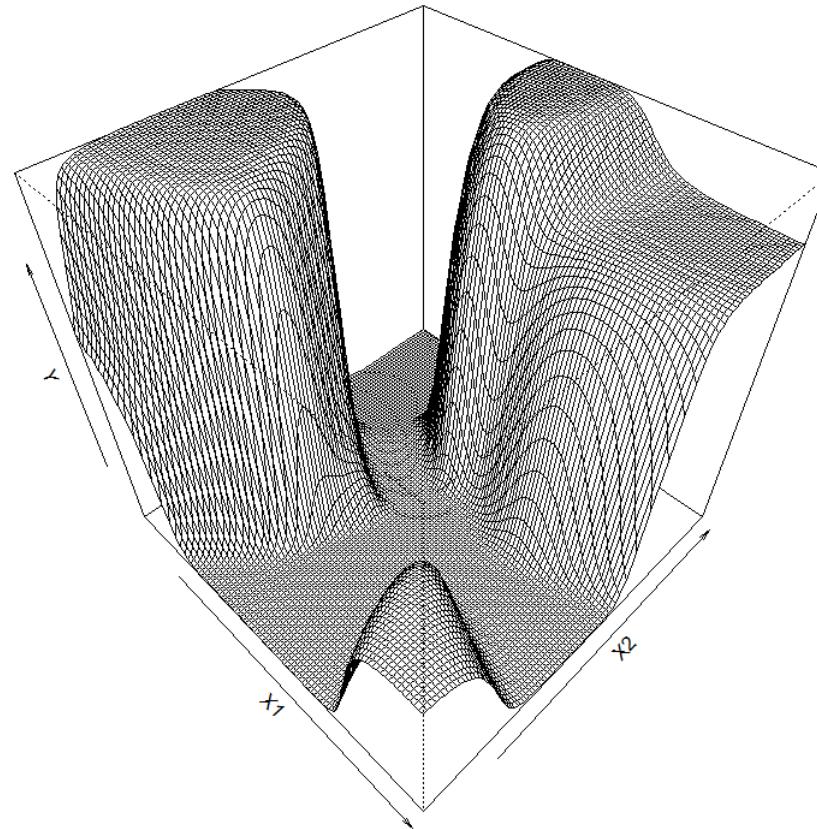
Learning



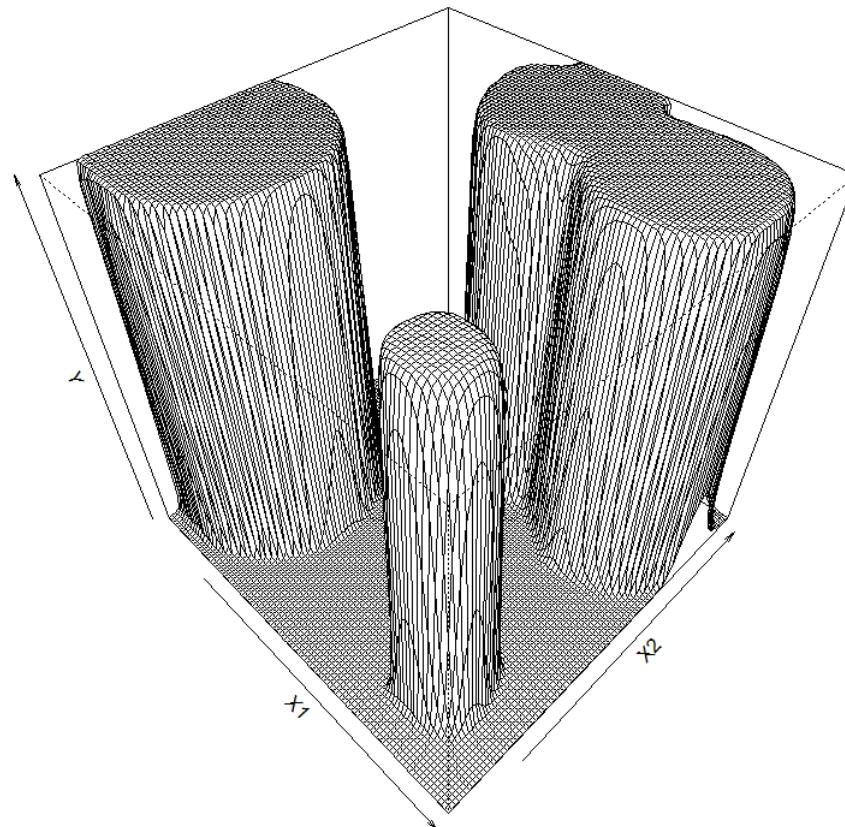
Learning



Learning



Learning



Learning

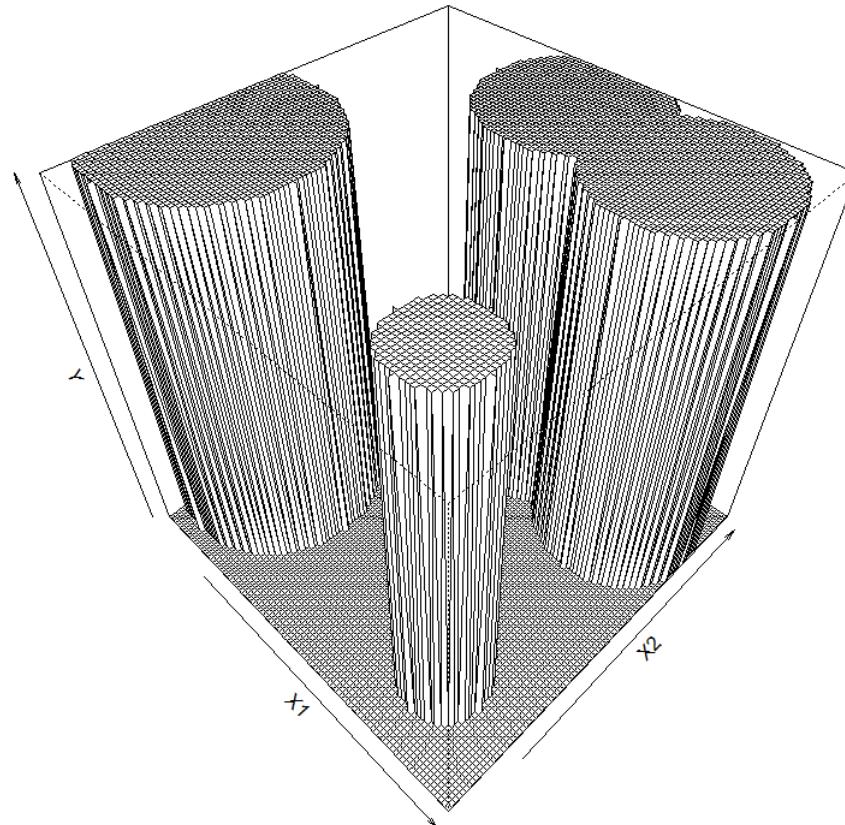
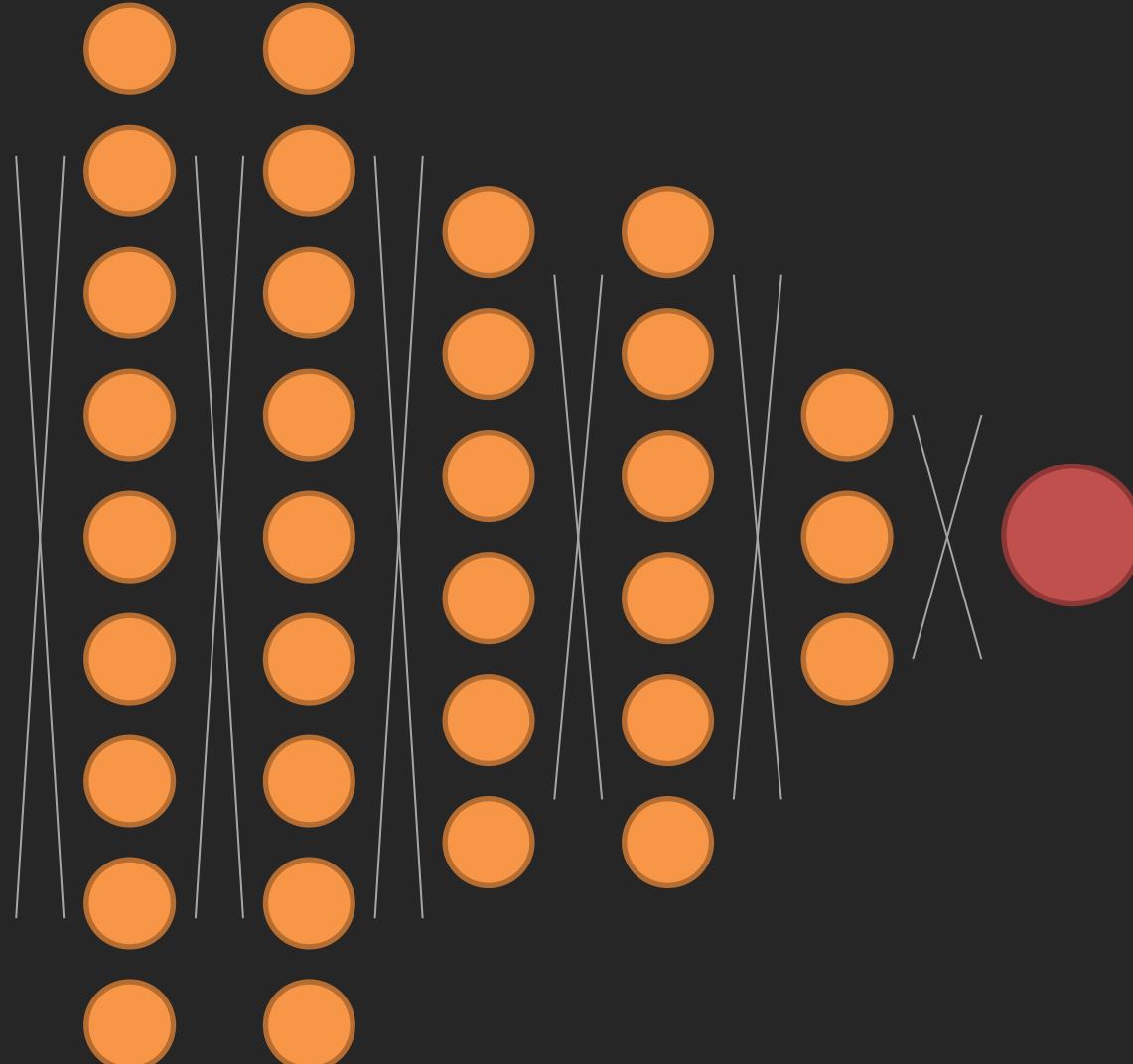
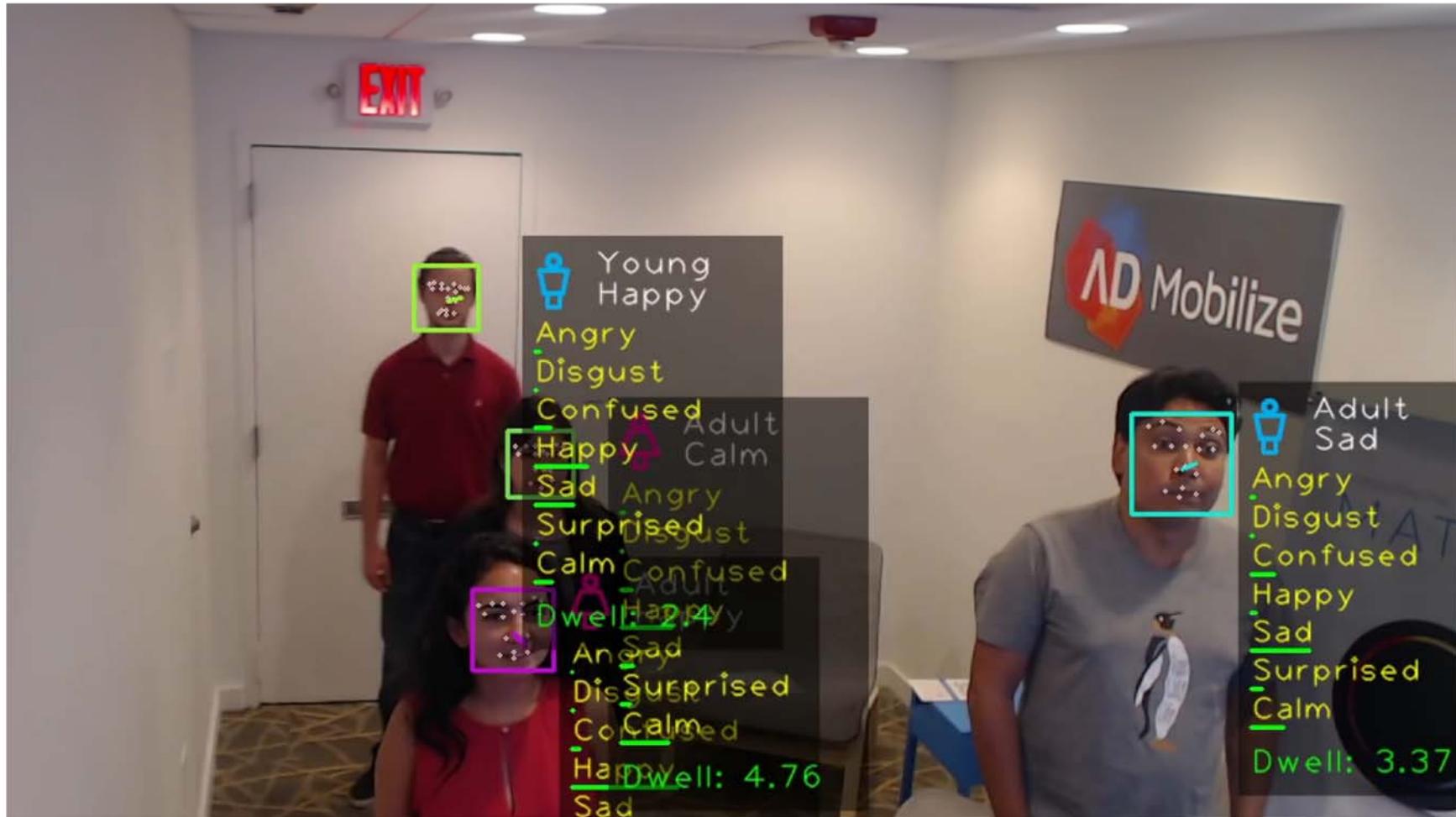


Image recognition

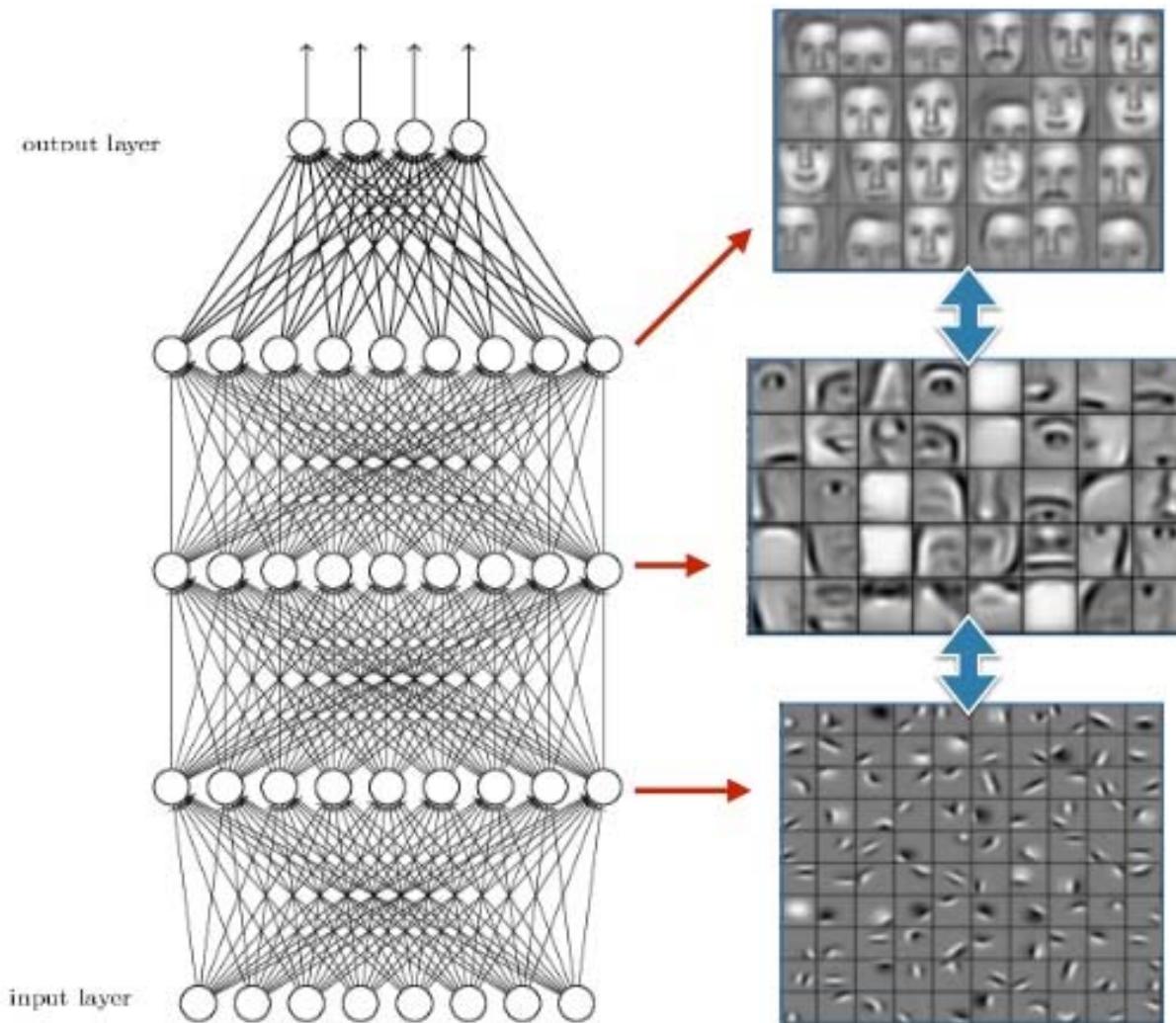


Cat?

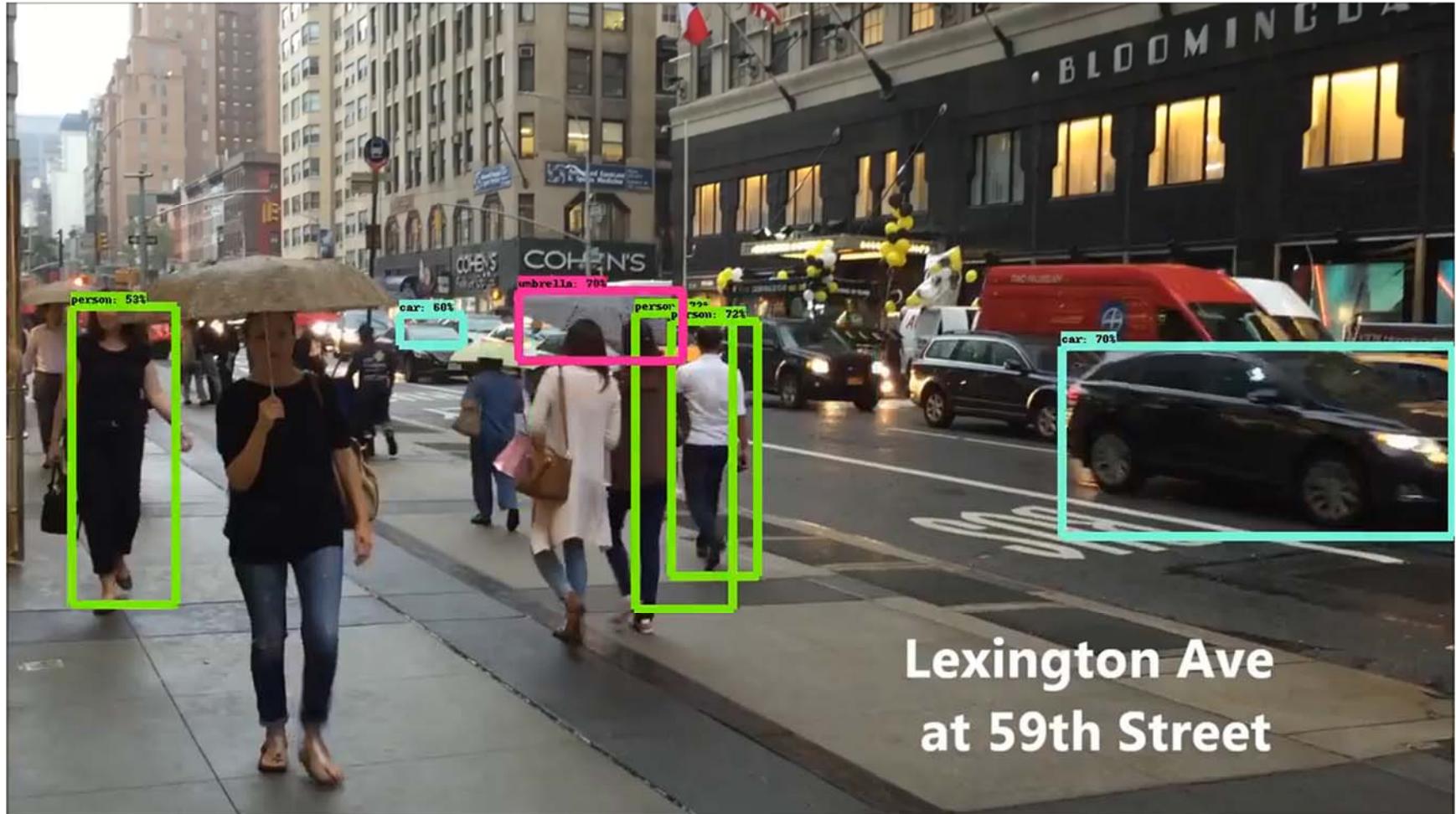
Facial recognition



How does it work?



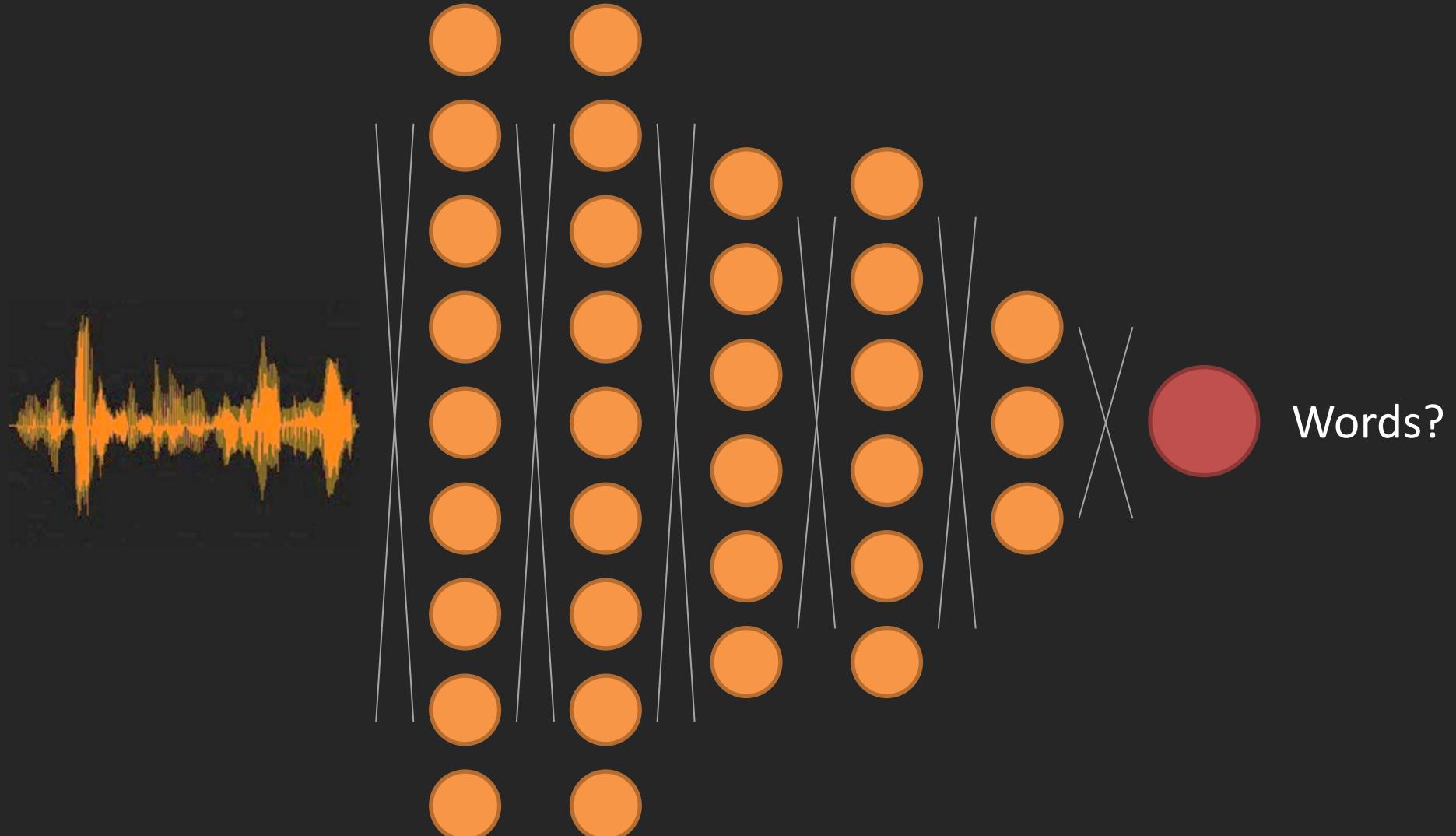
Object detection



Movement detection



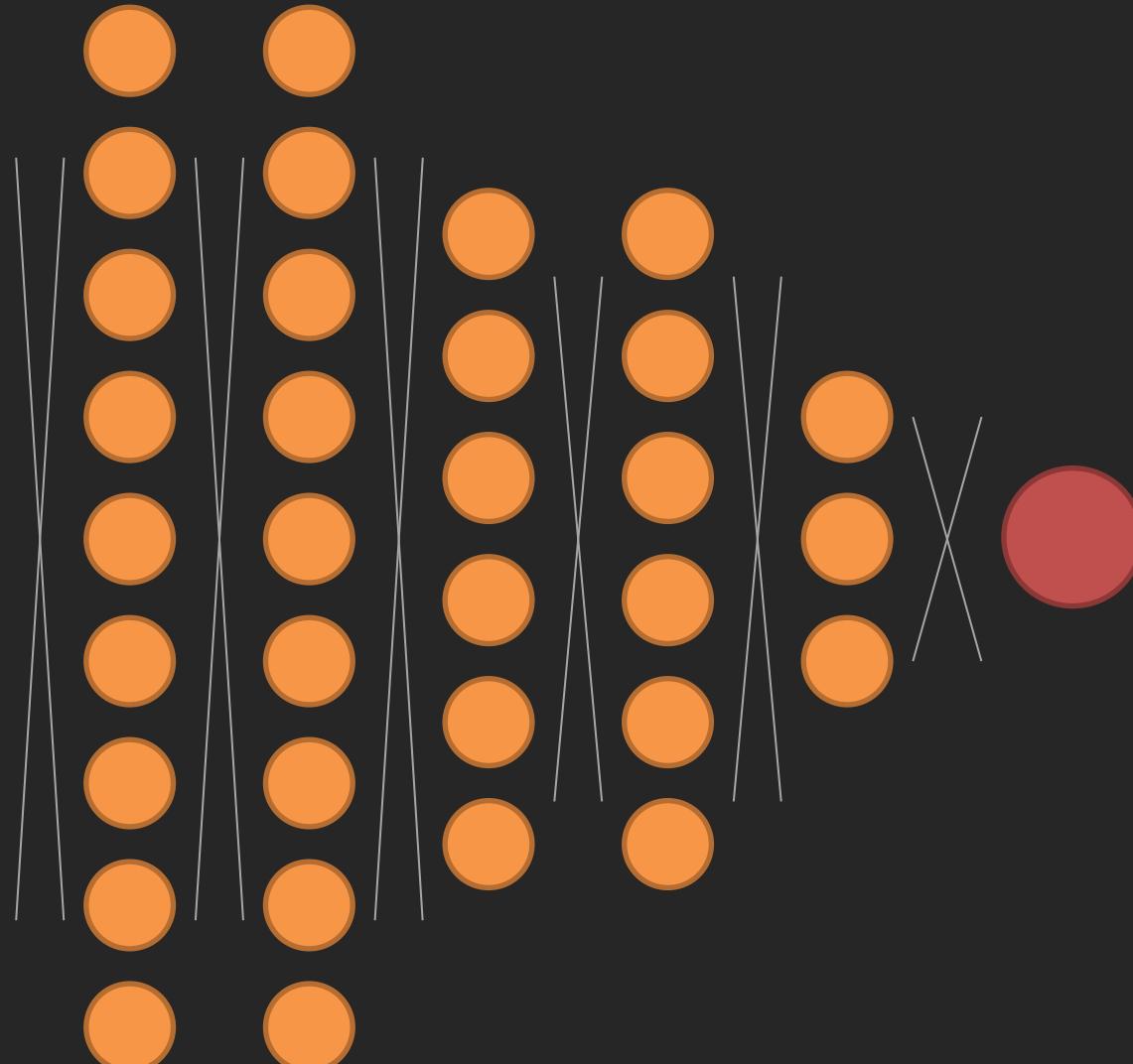
Voice recognition



Voice recognition



Sequence predictions



next...

iPhone



Bach

A musical score for four voices: Soprano, Alto, Tenor, and Bass. The music is in common time ($4/4$) and has a key signature of two flats. A red vertical line on the left indicates the beginning of the piece. The vocal parts are arranged vertically from top to bottom: Soprano, Alto, Tenor, and Bass. The Soprano part begins with a dotted quarter note followed by eighth notes. The Alto part follows with eighth notes. The Tenor part begins with a half note followed by eighth notes. The Bass part begins with a quarter note followed by eighth notes. The music continues with a series of eighth and sixteenth note patterns across all voices.

Shakespeare

When I in dreams behold thy fairest shade
Whose shade in dreams doth wake the sleeping morn
 The daytime shadow of my love betray'd
Lends hideous night to dreaming's faded form
 Were painted frowns to gild mere false rebuff
Then shoulds't my heart be patient as the sands
 For nature's smile is ornament enough
When thy gold lips unloose their drooping bands
 As clouds occlude the globe's enshrouded fears
 Which can by no astron'my be assail'd
Thus, thyne appearance tears in atmospheres
 No fond perceptions nor no gaze unveils
Disperse the clouds which banish light from thee
 For no tears be true, until we truly see

Horror

I remember his face
in a look of horror,
and it was agony and malice.
I was trapped.
I was trapped in this hospital bed.

Translation

The screenshot shows the Google Translate interface. At the top, it says "Google Translate". Below that is a navigation bar with "Text" selected (highlighted in blue) and "Documents". The main area has two language dropdowns: "FRENCH" on the left and "ENGLISH" on the right. A double-headed arrow icon is between them. Below these are two text boxes. The left text box contains the French sentence: "La répétition est la mère de la pédagogie, et la grand-mère de l'ennui". The right text box contains the English translation: "Repetition is the mother of pedagogy, and the grandmother of boredom". There are small icons for audio, character count (70/5000), and a star for favoriting. At the bottom right of the main area, there are icons for copy, edit, and share. A "Send feedback" link is at the very bottom right.

≡ Google Translate

Text Documents

DETECT LANGUAGE FRENCH ENGLISH SPANISH ↕ ENGLISH FRENCH SPANISH ↕

La répétition est la mère de la pédagogie, et la grand-mère de l'ennui

Repetition is the mother of pedagogy, and the grandmother of boredom

70/5000

Send feedback

Examples of applications

- NSA (TSA)
- Vehicle identification (license plates)
- OCR and document classification (accounting)
- Fraud detection (CPAM + IBM)
- Risk analysis (medical, insurances, audit, loans)
- Biometric security (iPhone)
- Self-driving cars
- Marketing (online marketing)
- Medical imagery... (radiology, melanoma)

Challenge #1



Models are so **complex**
it's often hard to explain
the “why”

Wisconsin v. Loomis



Gravity + Danger to society = Severity



U.S. Customs and Border Protection



Challenge #2



AI makes **mistakes**

These mistakes are all the more problematic that they are hard to detect

Error #1



Error #2



Challenge #3



AI learns from data

Even from **biased data**

Biais #1



Biais #2



Biais #3

20:32

08:59

11:41

11:45

Challenge #4



All predictions are probabilistic

Enemy combatants



Business opportunities

- Stable environment
- A lot of data
- Structured and quantified objectives
- Complex, misunderstood relationship between “inputs” and “outputs”
- No need for transparency

What about Sophia?

What about Sophia?





REINFORCEMENT LEARNING

Reinforcement learning



Rewards or penalties

- Rewards to seek:
 - Win a game
 - Win money, points
 - Reach a destination
 - Move an object back to its place
 - ...
- Penalties to avoid:
 - Lose a game
 - Bump into an obstacle
 - Fall
 - ...

Deep Blue vs. Kasparov (1997)



AlphaGo vs. Lee Sedol (2016)



DeepStack vs. pros (2016)



Learn to walk



Atlas



Meet **Atlas**, the latest humanoid robot from
Alphabet Inc.'s Boston Dynamics.

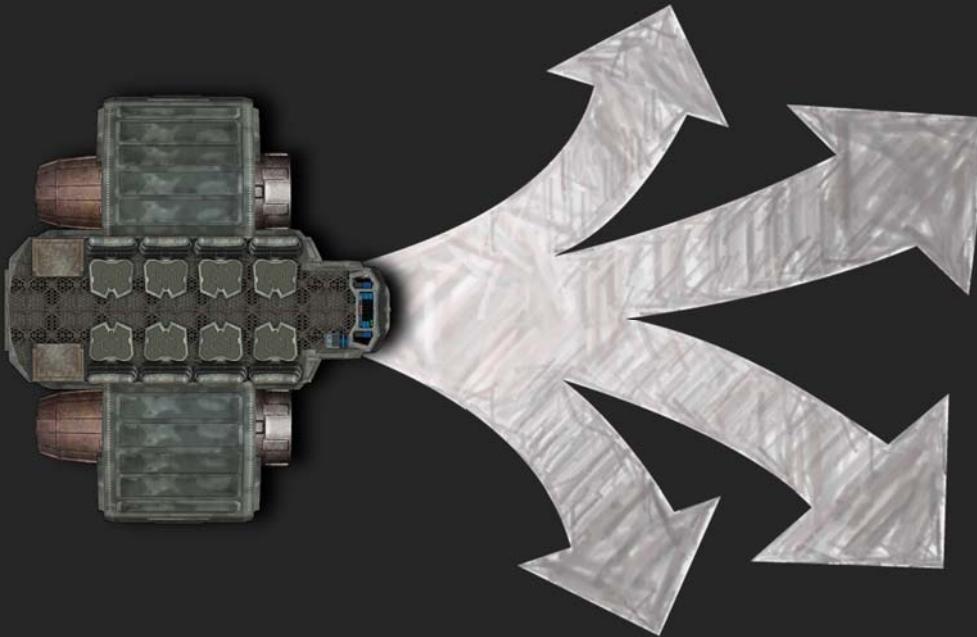
Boston Dynamics **WSJ**

Handle



Boston Dynamics

Rewards?



“To the airport ASAP...”



Challenge #5



What **goal** do we really want to achieve?

How to quantify the unquantifiable?

(ethics, happiness, safety, life, freedom...)



SOCIETY AND AI

Terminator



Robocop



I, Robot



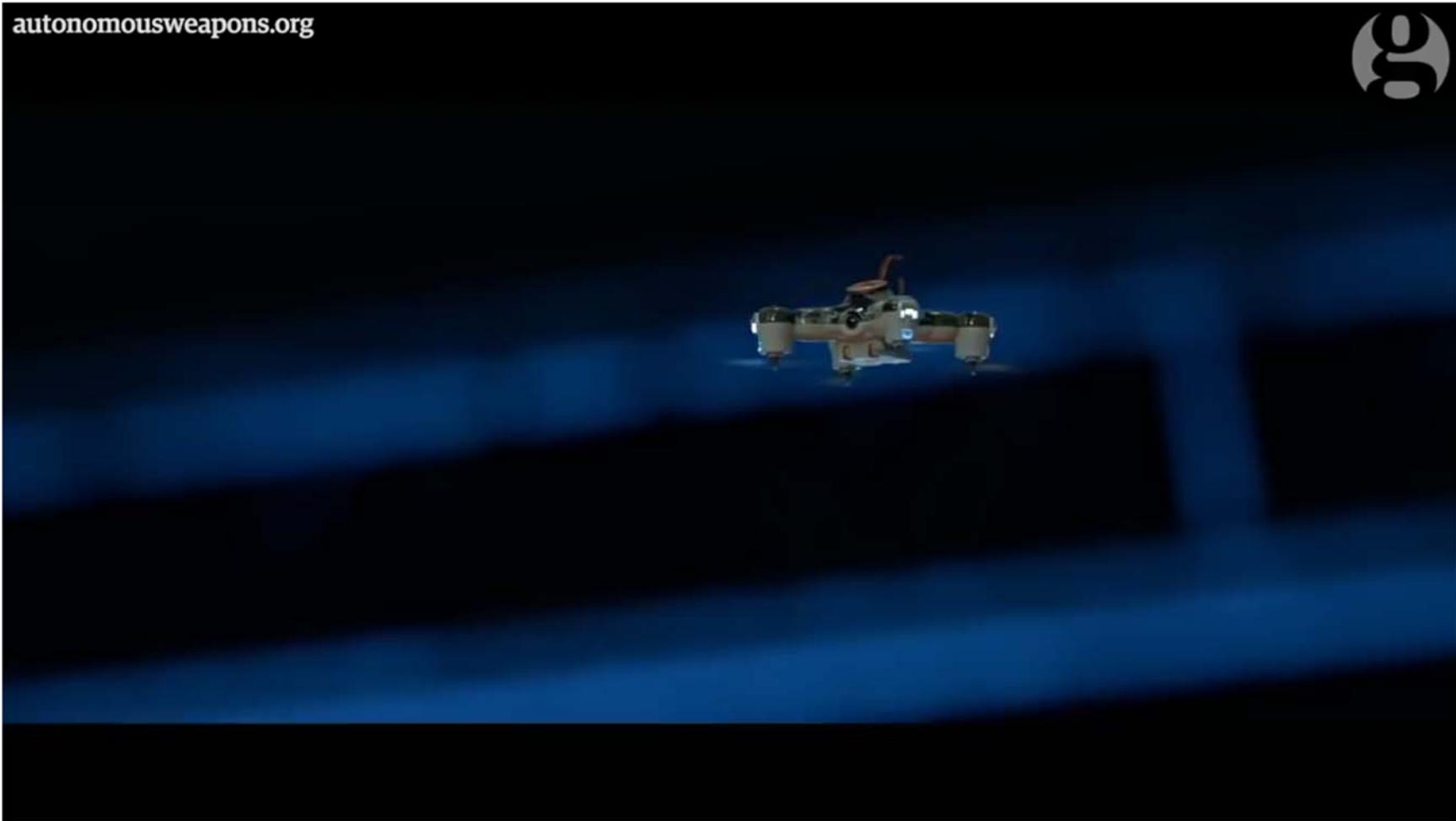
The three robotic laws

Asimov, 1942:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

Killer drones

autonomousweapons.org



Challenge #6



Killer robots

Job losses or...

<https://willrobotstakemyjob.com/>

| | |
|-----------------|-----|
| Psychiatrist | 4% |
| Police officers | 10% |
| Managers | 25% |
| Cameramen | 60% |
| Slaters | 90% |
| Cooks | 92% |
| Accountants | 94% |
| Drivers | 98% |

Challenge #7

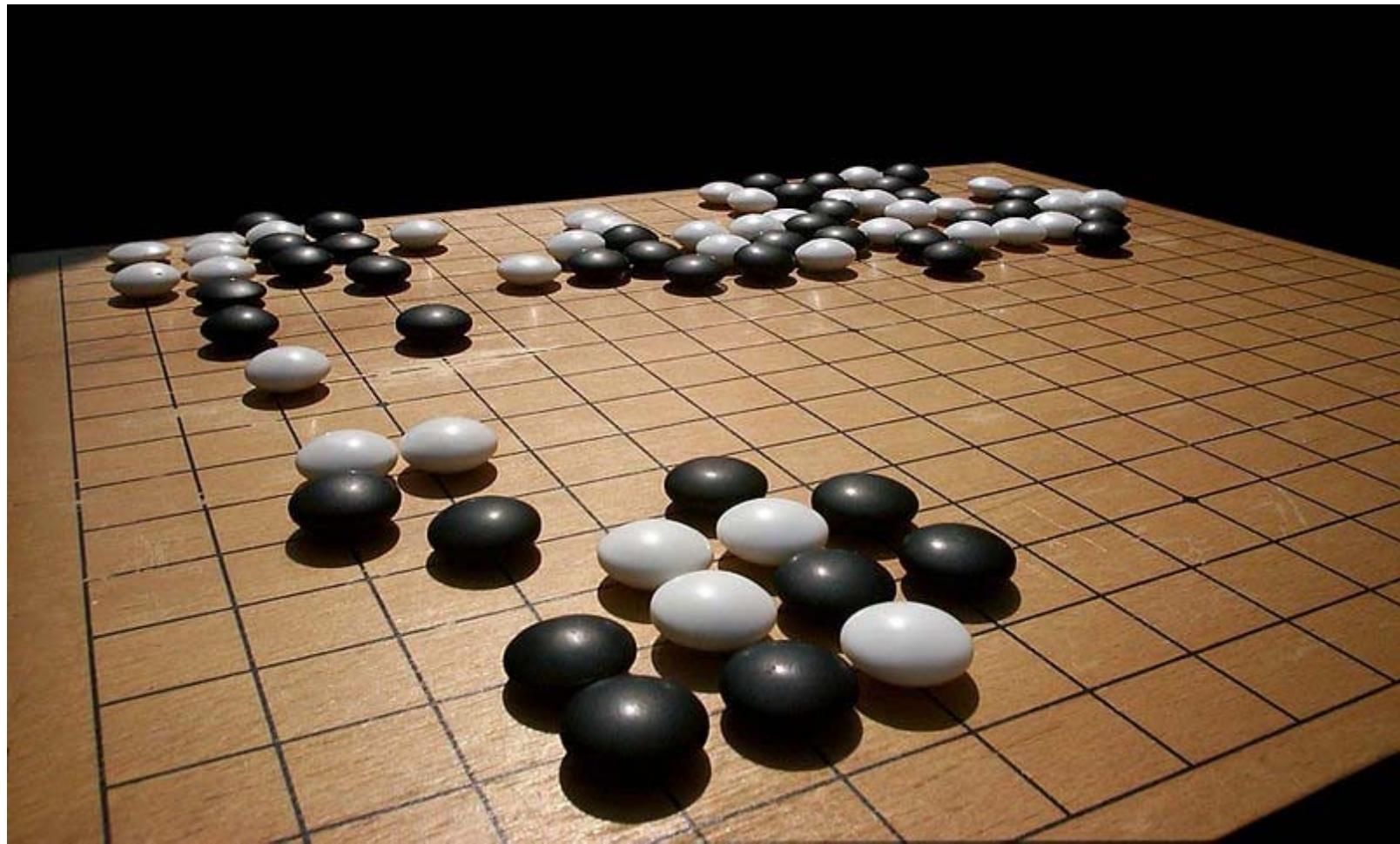


How to prepare a society
where there will be much **fewer**
jobs?

Challenge #8



Tipping point, AI singularity,
and Artificial General
Intelligence
(or « strong » AI)





That's all folks!

(I mean, really... It's over)

