Machine Learning

Are machines the next Einstein?

Institute of Engineering and Technology
Susanna Green
“Even though I am doing an Astrophysics PhD… I oscillate between being a software engineer, machine learning engineer, and a data scientist. Plus I am a content creator!”

– Susanna Green –
NEW video this Friday!
How do YOU think and learn?
“Did you know the average person has about 6,200 thoughts per day”

Psychologists at Queen’s University in Kingston, Ontario, 2020 (Credit: Bigthink.com)

A. I have more thoughts than this!!!
B. I have about this many thoughts.
C. I have very few thoughts a day…
Have you ever thought about your thoughts? Like, what even is a thought?

Like… what are your thoughts made of?
“Thoughts are generated in the brain, which is composed of 100 billion neurons that transmit impulses through the synapses. Thoughts are electrochemical reactions.” - ScienceABC
Biological neuron

Stimulus → Enough stimulus to send on signal → Signal passed onto the next neuron
In 1957, Frank Rosenblatt created the first artificial neuron, called the perceptron.
Biological neuron v.s. artificial neuron (perceptron)?

Credit: Wiki
Credit: knime.com
Biological neuron v.s. artificial neuron (perceptron)?

- Biological neuron:
  - Cell body
  - Axon
  - Axon hillock
  - Golgi apparatus
  - Endoplasmic reticulum
  - Nucleus
  - Mitochondrion
  - Dendrite
  - Dendritic branches
  - Synaptic terminals

- Artificial neuron (perceptron):
  - $x_1, x_2 = \text{inputs}$
  - $w_1, w_2 = \text{weights}$
  - $b = \text{bias}$
  - $f = \text{activation function}$
  - $y = \text{output}$

$$f \left( b + \sum_{i=1}^{n} x_i w_i \right)$$
Biological neuron v.s. artificial neuron (perceptron)?

V.S.

\[
\begin{align*}
x_1, x_2 &= \text{inputs} \\
\sum f(b + \sum_{i=1}^{n} x_i w_i) &= \text{activation function} \\
y &= \text{output} \\
w &= \text{weights} \\
b &= \text{bias}
\end{align*}
\]
“Neurons that fire together, wire together.”

- Donald Hebb -
You are here!

My second home!
Biological neuron v.s. artificial neuron (perceptron)?

V.S.

\[ y = f \left( b + \sum_{i=1}^{n} x_i w_i \right) \]

where:
- \( y \) = output
- \( w \) = weights
- \( b \) = bias
- \( x_i \) = inputs
- \( f \) = activation function

Input → Output
Biological neuron v.s. artificial neuron (perceptron)?

V.S.

\[ x_1, x_2 = \text{inputs} \]

\[ w = \text{weights} \]

\[ b = \text{bias} \]

\[ y = \text{output} \]

\[ f = \text{activation function} \]
Biological neuron v.s. artificial neuron (perceptron)?
Humans think when neurons are connected (neural pathway), similarly machines ‘think’ when artificial neurons are connected.
Neural networks are one way that a machine ‘thinks’ and ‘learns’...
YOU v.s. MACHINE

Dog

Cat
YOU v.s. MACHINE

Credit: upadhyay-shivam.medium.com
What is Machine Learning?
The machine learns the pattern rather than you explicitly telling the machine the pattern
Humans learn in different ways... 

...so do machines.
Supervised learning you tell the computer what to learn (i.e. a cat or dog) while unsupervised learning you don’t tell the computer what to learn.
Supervised learning is when you give the computer the ingredients and final dish, then tell it to find a recipe.

Unsupervised learning is when you give the computer the ingredients and ask it to create a recipe and new dish.
Let’s use a classification task to explain this...
Supervised Learning

Unsupervised Learning

Labelled Data

Image Credit: thabisfotowelt, Pixabay
Image Credit: GreenCross Vets
Supervised Learning

Unsupervised Learning

Labelled Data

Unlabelled Data

Image Credit: thabisfotowelt, Pixabay
Image Credit: GreenCross Vets
NEURAL NETWORKS

Labelled Data

Unlabelled Data
Disclaimer: Both require machine learning engineers to check the results.
How does a Neural Network ‘learn’?
Like humans, neural networks learn by their mistakes…

…but we call it ‘Optimization’!
Optimization

Neural Network

Data preparation

Data

Evaluate the Loss Function

Minimise the Loss Function

Loss Function

Minimise the Loss Function
Optimization

Neural Network

Data preparation

Evaluate the Loss Function

Minimise the Loss Function
Optimization

Neural Network

Data preparation

Evaluation the Loss Function

Minimise the Loss Function
Optimization

Neural Network

Data

Data preparation

Evaluate the Loss Function

Minimise the Loss Function

Image by chandlervid85 on Freepik
"The Loss Function is the choice of the error function which is how the optimization algorithm is told how good the model is…"
"The Loss Function is the choice of the error function which is how the optimization algorithm is told how good the model is…"
Optimization: Loss Function

“As a machine learning engineer, it is your responsibility to choose the loss function that best suits your problem.”

For example:
Optimization: Loss Function

“As a machine learning engineer, it is your responsibility to choose the loss function that best suits your problem.”

For example:
Optimization: Loss Function

“As a machine learning engineer, it is your responsibility to chose the loss function that best suits your problem.”

For example:

1) Mean Squared Error (MSE)
2) Mean Squared Logarithmic Error (MSLE)
3) Mean Absolute Error (MAE)

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2
\]

\[
MSLE = \frac{1}{n} \sum_{i=1}^{n} (\log(Y_i) - \log(\hat{Y}_i))^2
\]

\[
MAE = \frac{1}{n} \sum_{j=1}^{n} |y_j - \hat{y}_j|
\]
“Optimizers are responsible for changing the weights, bias and learning rate of the neurons in the neural network to reach the minimum loss function.”
“...Optimizers are responsible for changing the weights, bias and learning rate of the neurons in the neural network to reach the minimum loss function.”
Optimization

Neural Network

Data preparation

Data

Evaluate the Loss Function

Minimise the Loss Function

\[ \theta_t = \theta_{t-1} - \alpha_t \nabla_{\theta_{t-1}} L_t \]

New weights and bias

Learning Rate

Loss Function

Current weights and bias
In order for the neural network to learn, the neural network architecture may need to be adjusted…
it is up to a machine learning engineer to adjust the neural network architecture so that the machine ‘learns’.
E.g. vary the width of the neural network...
E.g. vary the depth of the neural network…
Vary the depth of the neural network…

**Deep Learning** is defined as a neural network with many layers.
Different neural network architectures enable the machine to ‘learn’ different data types.
In the modern era we have lots of different types data (text, images, audio)...

... this has led to an explosion of different neural networks!
Keeping up with the Neural Networks

- Convolutional Neural Networks
- Graph Neural Networks
- Recurrent Neural Networks
- Transformer
- Deepsets
- Multi-Layered Perceptron
- Convolutional Neural Networks

Image Credit: hdqwalls.com
Tell me more, Tell me more…
Convolutional Neural Networks
Convolutional Neural Networks

It is an artificial neural network that features one or more convolutional layers.
How do **Humans** process **visual information**?
Did you know that it takes your brain 100 milliseconds to process an image.

True or False?
Did you know that it takes your brain 100 milliseconds to process an image.

False
Did you know that it takes your brain 100 milliseconds to process an image.

"A team of neuroscientists from MIT has found that the human brain can process entire images that the eye sees for as little as 13 milliseconds”
- Anne Trafton

False
The visual cortex is where your brain processes visual information.
How do we know this?
In the 1950s at Johns Hopkins University, the David Hubel and Torsten Wiesel performed pioneering research of how visual information is processed in the brains of mammals.
David and Torsten set up the following experiment…
...and concluded that the cat’s primary visual cortex neurons were only stimulated by certain angles of orientation.

Figure 1. A neuron in the primary visual cortex responds selectively to line segments. Credit: Bin Li et al. (2022)
This is **Neocognition**, is the first convolutional neural network architecture.

This is **Neocognition** was able to categorize handwritten digits.
The neural network architecture shown above is called LeNet 5 and is used in the postal service.
Convolutional neural networks are not perfect.
Limitations of Convolutional Neural Networks (CNNs) are evident when they are used to moderate content on social media… it doesn’t completely block or remove content.
ImageNet is publicly released dataset containing 14 million labelled images and is used in classification and object detection tasks.
CNN’s struggle with different lighting, extreme angles and parts of the object.
Machine learning is not about ‘thinking like a human’
Machine learning by default does not mimic human thinking...
Machine learning by default does not mimic human thinking...

...and we don’t want to because humans have faulty thinking.
Insight - Amazon scraps secret AI recruiting tool that showed bias against women
By Jeffrey Dastin
October 11, 2019 1:50 AM GMT+

Apple Card Investigated After Gender Discrimination Complaints
A prominent software developer said on Twitter that the credit card was “sexist” against women applying for credit.

Google’s algorithm shows prestigious job ads to men, but not to women. Here’s why that should worry you.
By Julia Carpenter
July 6, 2015 at 4:43 p.m. EDT

Predictive policing is still racist—whatever data it uses
Training algorithms on crime reports from victims rather than arrest data is said to make predictive tools less biased. It doesn’t look like it does.
By Will Douglas Heaven
February 5, 2021
Bias can be injected into a machine learning pipeline at various stages...

...including how the problem is framed.
You’re machine learning engineers, let’s predict whether England will win the Euros 2024.
How are we going to do this?
Aim = Total number of Goals
Bias can be injected into a machine learning pipeline at various stages... including how the problem is framed.
Bias can be injected into a machine learning pipeline at various stages…

…including how the problem is framed.
Nom. goals scored
Height
Nom. red cards
Goal attempts
Junk in = Junk out
Why is machine learning bias hard to fix?

1) Not obvious
The neural network was reprogrammed to ignore explicitly gendered words, like “woman”, but then the system started picking up implicitly gendered words, such as “executed” and “captured”.

An algorithm that was being tested as a recruitment tool by online giant Amazon was sexist and had to be scrapped, according to a Reuters report.

The artificial intelligence system was trained on data submitted by applicants over a 10-year period, much of which came from men, it claimed.
Why is it hard to fix?

1) Not obvious
2) Imperfect process
80% Training Data

Data

20% Test Data
Why is it hard to fix?

1) Not obvious
2) Imperfect process
3) Lack of social context
United Kingdom

India
Why is it hard to fix?

1) Not obvious
2) Imperfect process
3) Lack of social context
4) The definitions of fairness
What is ‘fairness’?

50% of men and 50% of women are considered high risk?

Everyone considered high risk?
It’s up to a machine learning engineer to be aware of bias and do everything they can to mitigate it.
How would you reduce bias in machine learning?
Maybe you could be the solution?
This is how you can help...
Welcome! We need your help to classify some audio clips. This will help us to better understand how children learn language from the world around them.

Get involved at zooniverse.org!

Help us describe and categorize more than 60,000 postcards featuring rare photographs depicting life in Michigan during the first half of the twentieth century.

As a result of your enthusiastic support and contributions, we couldn’t have reached this point without you! You can still help us by transcribing text from postcards by clicking on the project box.

Picturing Michigan’s Past

Maturity of Baby Sounds

Corresponding with Quakers

Woodpecker Cavity Cam

Dolphin Chat
Get involved at zooniverse.org!

Black Hole Hunters

Identify distant galaxies to help measure dark energy when the universe was just ~2-3 billion years old

Galaxy Zoo

Few have witnessed what you're about to see

Learn more  Get started

Planet Hunters NGTS

Discover hidden worlds with the Next-Generation Transit Survey

Learn more

Gravity Spy
We need you to help us identify the physics...

Find cosmic explosions in real-time with the Gravitational-wave Optical Transient Observer (GOTO) - new data uploaded hourly!

Learn more  Get started
...in this research a neural network is uncertain about whether it has observed an explosion and we need you to clarify!

Credit (edited): Venkatesh Tata
How are neural networks being used in Astronomy?
In 2023, a neural network (with human supervision) identified 527 new strong lensing candidates, a rare event where light bends around heavy objects like galaxies.
Neural networks have also been used to identify 301 exoplanets
Neural networks are challenging some of our theories of galaxy evolution by classifying the galaxies differently... new physics?
LearningMatch
Using neural networks to predict the ‘match’
Gravitational Waves are ripples in the curvature of spacetime produced by accelerating masses... such as two black holes or two neutron stars colliding.
Searching for gravitational waves is computationally expensive…
Searching for gravitational waves is computationally expensive…

…so we want to use neural networks to speed up the process.
What is the ‘match’?
What is the ‘match’?

* is the Complex conjugate of a function
~ is the Fourier Transform

\[ \mathcal{M}(h_1, h_2) = \max_{\phi, t_c} \langle h_1 | h_2(\phi, t_c) \rangle \]

\[ = \max_{\phi, t_c} \left[ 4\Re \int_0^\infty \tilde{h}_1^*(f) \tilde{h}_2(f) \frac{1}{S_h(f)} \right] \]

One-sided Power Spectral Density (PSD) - average noise in our detectors

\backslash The match is the weighted inner product maximized over the phase and the time of coalescence.
What is a good ‘match’? // It’s like gravitational wave online dating!

STRONG MATCH = 0.99 (or ~1)
What is a good ‘match’?

// It’s like gravitational wave online dating!

Name: IMRPhenomD
Describe yourself:
Mass1=10
Mass2=10

WEAK MATCH

= 0.01

Name: IMRPhenomD
Describe yourself:
Mass1=100
Mass2=30
The aim of the neural network is to learn the relationship between the parameters and the match.
The aim of the neural network is to learn the relationship between the parameters and the match.
What have I **learned** as a machine learning engineer?

1) **Unknown** - We need the neural network to be accurate within 1%, which hasn’t been done before.
What have I **learned** as a machine learning engineer?

1) **Unknown** - We need the neural network to be accurate within 1%, which hasn’t been done before.

2) **Misleading** - What has the neural network actually learned?
What have I **learned** as a machine learning engineer?

1) **Unknown** - We need the neural network to be accurate within 1%, which hasn’t been done before.

2) **Misleading** - What has the neural network actually learned?

3) **Frustrating** - You are looking for a ‘needle in a haystack.’
What have I **learned** as a machine learning engineer?

1) **Unknown** - We need the neural network to be accurate within 1%, which hasn’t been done before.

2) **Misleading** - What has the neural network actually learned?

3) **Frustrating** - You are looking for a ‘needle in a haystack.’

…this is why machine learning engineers are paid nice salaries!
Using neural networks in astronomy is still current research...
Using neural networks in astronomy is still current research...

...so who knows what the future looks like!
Are machines the next Einstein?
Thank you for listening and are there any questions?

Scan here to get involved with an astronomy project from your phone... we need you!

Get involved at zooniverse.org!
Convolutional layer

Neural Network

Image ➔ Feature Extraction ➔ Neural Network ➔ Output

Credit: LeCun et al. (1998)
What is a convolutional layer?

A convolutional layer extracts certain information from the image (i.e. lines at a particular orientation). The output is what the NN sees.
Feature Visualization

How neural networks build up their understanding of images

Check out ‘Colah’s blog’ for more information on convolutional layers!