Welcome Here!!

NATURAL LANGUAGE PROCESSING (NLP)

AN INTRODUCTION TO AI, ML, DEEP LEARNING & NLP

DSN LEKKI-AJAH

BY ABEREJO HABEEBLAH O.

X (TWITTER): @HABEREJO

Day 1

Tuesday, 4th March, 2025

Why are we here

To equip learners with the skills to understand how Artificial Intelligence models are built, explore Machine Learning and Deep Learning, and focus on Natural Language Processing (NLP).

We should build some project right?

Yes!!

If you are in with me,

We should build some project.

Classes Structure

Every Tuesday throughout the month of March.

4 classes in total.

Each class for about 1:45 minutes

Days are NOT likely to change, although, my schedule can be OUCH, I wil reachout earlier before. Understand me bikoooo

Who should be here

- You're curious and ready to learn something new
- You dream of becoming a Data Scientist.
- You're passionate about building the future as a Machine Learning Engineer.
- You have a basic understanding of Python programming.
- You've worked with data and want to take your skills to the next level.
- You're a researcher exploring the exciting world of AI.

which of these are you?

Who should be here

- You're familiar with the fundamentals of Machine Learning and Deep Learning.
- You love problem solving and are excited to tackle real world problems with AI.
- You want to understand the technology that is shaping the future.
- You're driven by a desire to learn and grow in the field of AI.
- You enjoy collaborating and learning from others.

Which of these are you?????
You ticked any of these boxes?
Let's Goooooooo

Where Do We Begin

- Know my audience
- Understand why they join this program
- Discuss previous projects and skills in AI

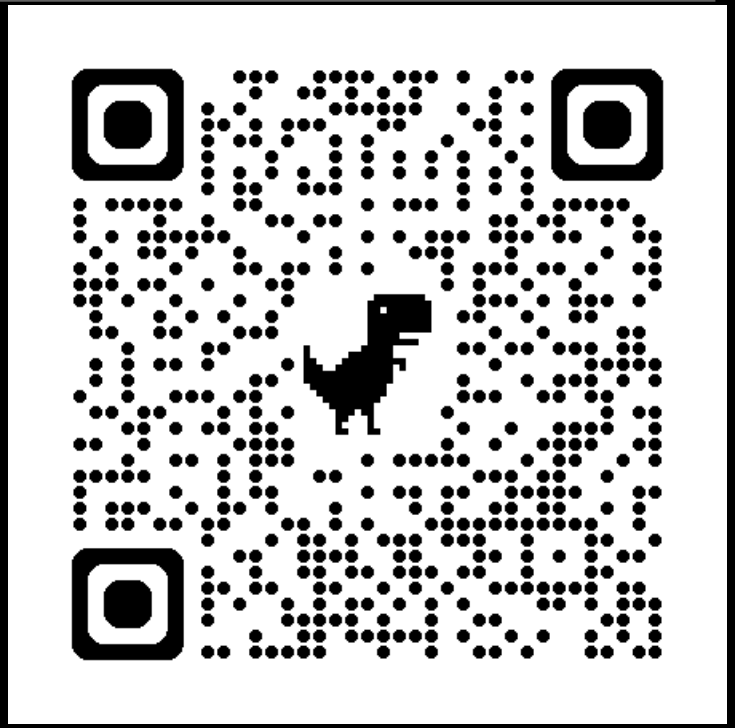
Let's have some discussion here

Before we begin

Who I am?

- A Machine Learning Engineer
- A young MAN eager to master AI

FUN FACT:
I've taught over 100 people Python, but
I'm still learning new things every day.
(Especially how to avoid typos when
coding late at night!)



Scan to visit my portfolio
or
https://bheez.netlify.app

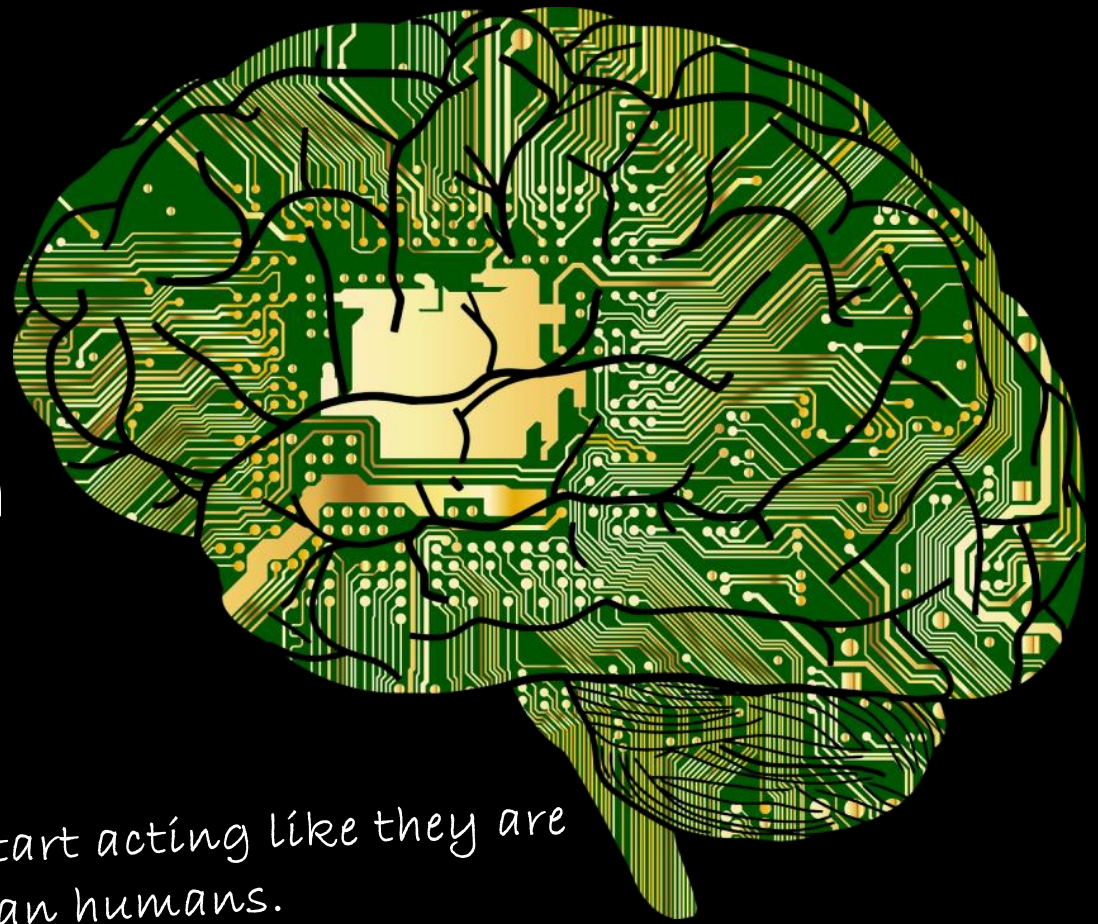
Course Structure

1. Introduction to AI, ML, DL, and NLP
2. Understanding AI vs ML vs DL
3. Introduction to NLP
4. NLP Use Cases
5. Text Preprocessing Techniques
6. Hands-on NLP Tasks
7. Brain Teasers & Discussion

And lots more

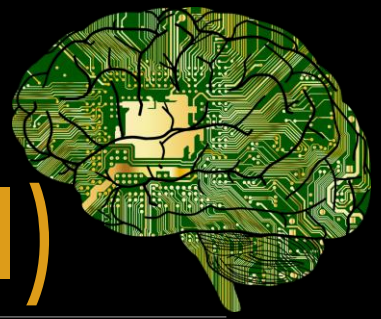
What is Artificial Intelligence (AI)

Artificial Intelligence is the ability of machines to mimic human intelligence. It enables computers and systems to learn from data, solve problem, and make decisions just like humans do.



AI is when machines start acting like they are smarter than humans.

What is Artificial Intelligence (AI)



Let's relate like this again:

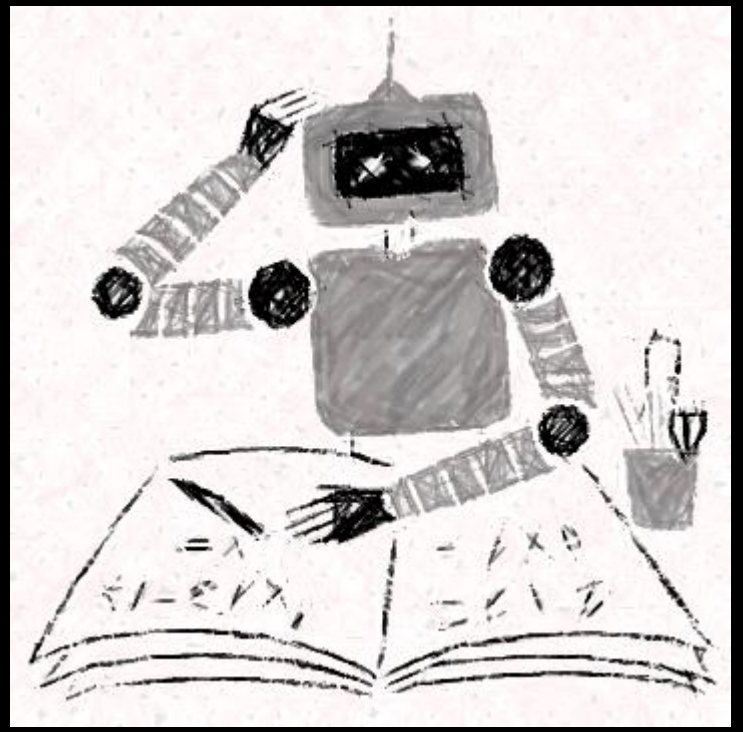
AI is like teaching your computer to do the thinking for you, only that it doesn't demand coffee breaks or complain about Mondays!!

hahaha😀😀😀😀

AI is when machines start acting like they are smarter than humans.

What is Machine Learning (ML)

Machine learning is turning things (data) into numbers and finding patterns in those numbers.



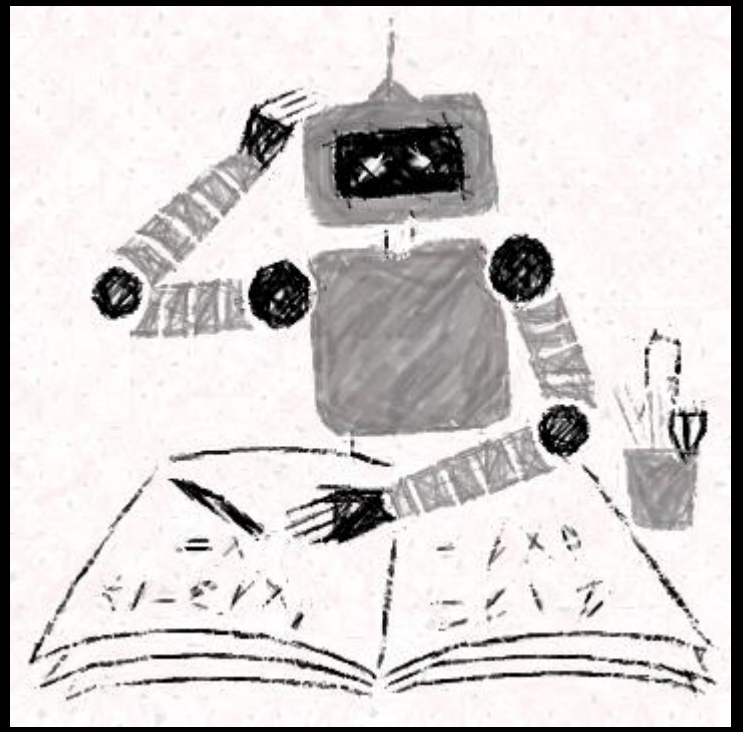
Got It???

Let's Goooooo!!!

What is Machine Learning (ML)

Think of it as teaching a computer how to recognize patterns without explicitly programming it.

It uses some complex algorithms to find patterns.



Got It???

Let's Goooooo!!!

What is Machine Learning (ML)

Find pattern? How?

Imagine teaching a child to identify fruits. You show them a banana, an apple, and an orange, telling them the names. After seeing many examples, they can recognize a new banana without needing you to tell them. That's what ML does—it learns from past examples and generalizes.

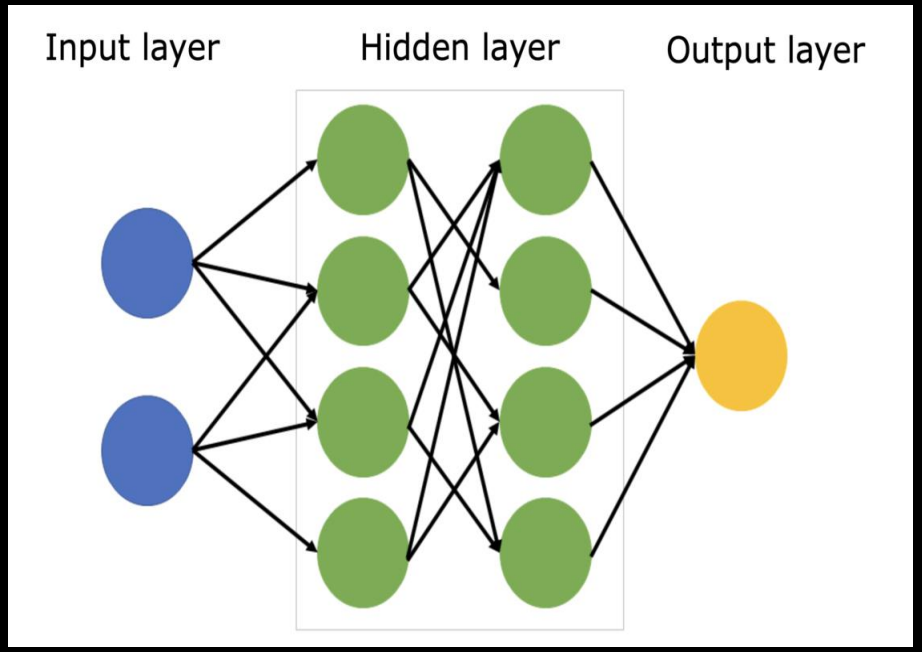
Got It??? Let's Goooooo!!!

What is Deep Learning (DL)

A deeper version of ML.

It works like our brain with many layers of learning.

Instead of just learning basic rules, it finds complex patterns automatically by using lots of examples.



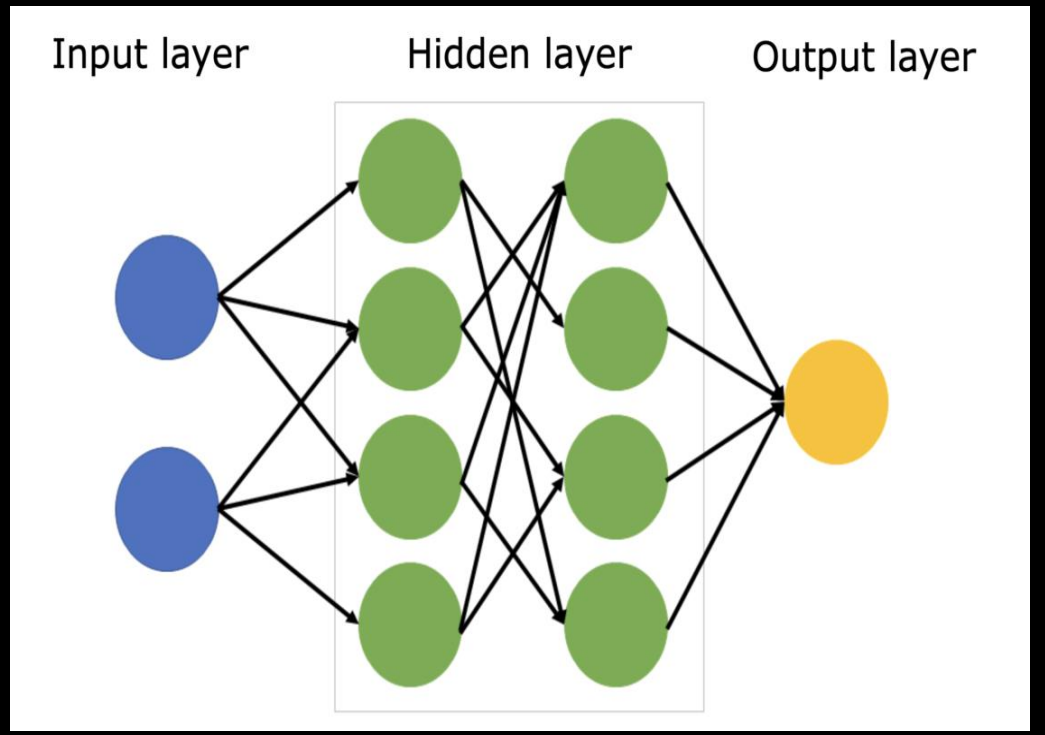
Got It???

Let's Goooooo!!!

What is Deep Learning (DL)

It uses multi-layered neural networks (i.e. Deep Neural Networks) to automatically extract features from data.

It is especially effective for tasks like image recognition, speech processing, and NLP.



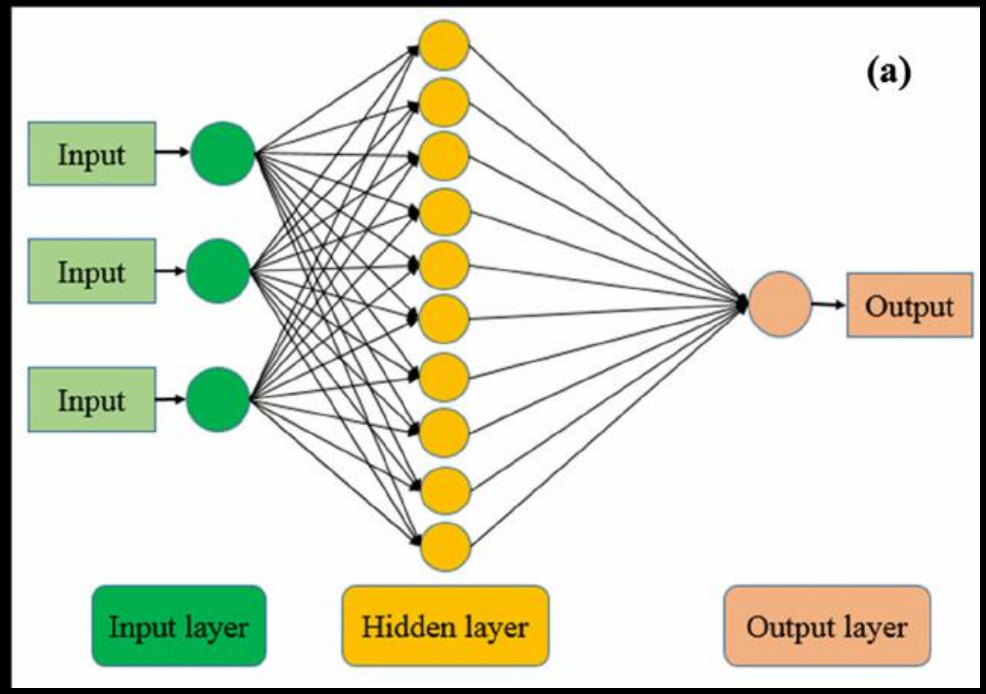
Got It???

Let's Goooooo!!!

What is Deep Learning (DL)

Neural networks? How?

Imagine teaching a robot to recognize handwriting. Traditional ML would require manually programming rules for different letters. But Deep Learning would just look at thousands of handwritten examples and figure out the differences itself—like how humans learn handwriting by seeing and practicing.

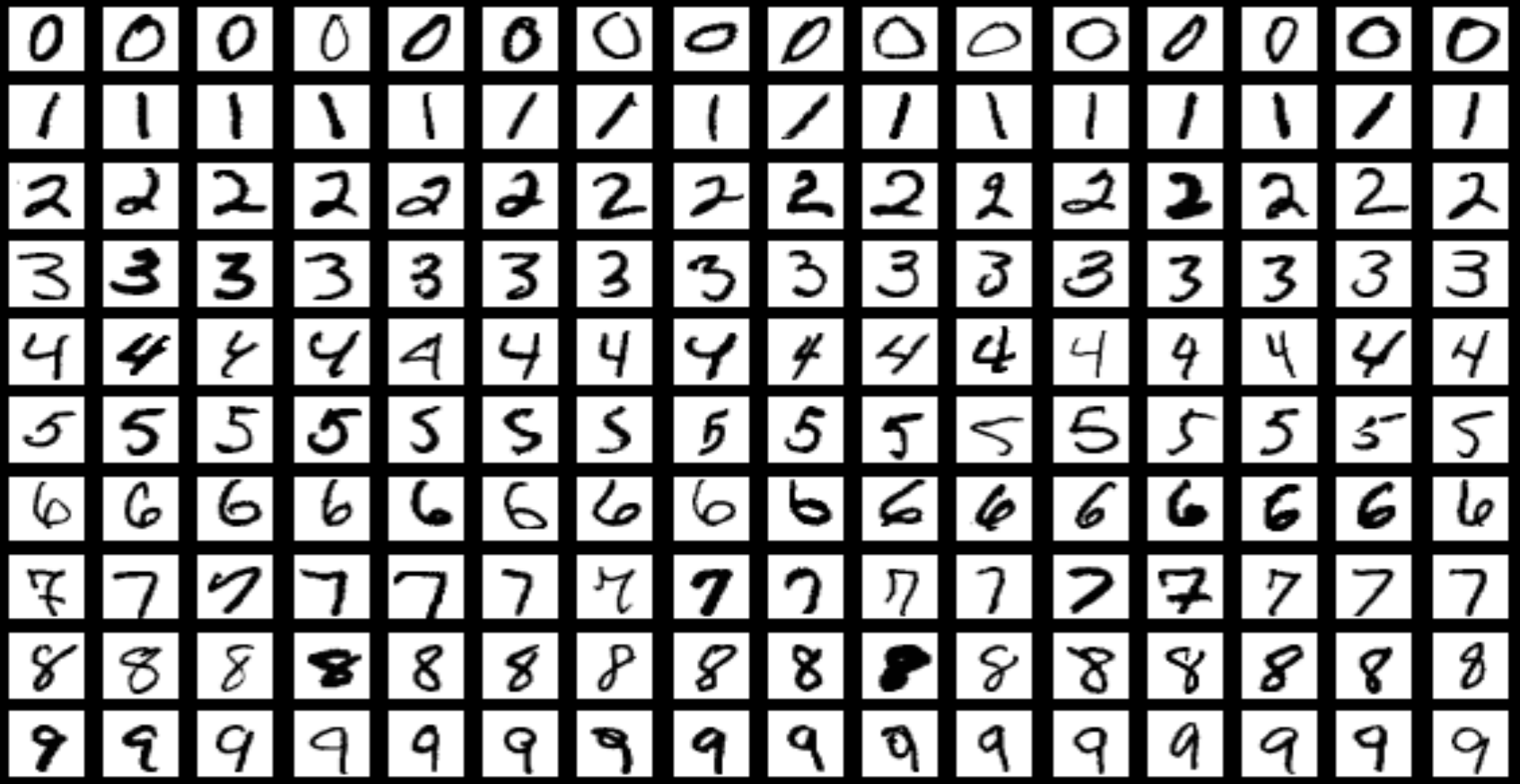


Got It???

Let's Goooooo!!!

What is Deep Learning (DL)

Neural networks? How?

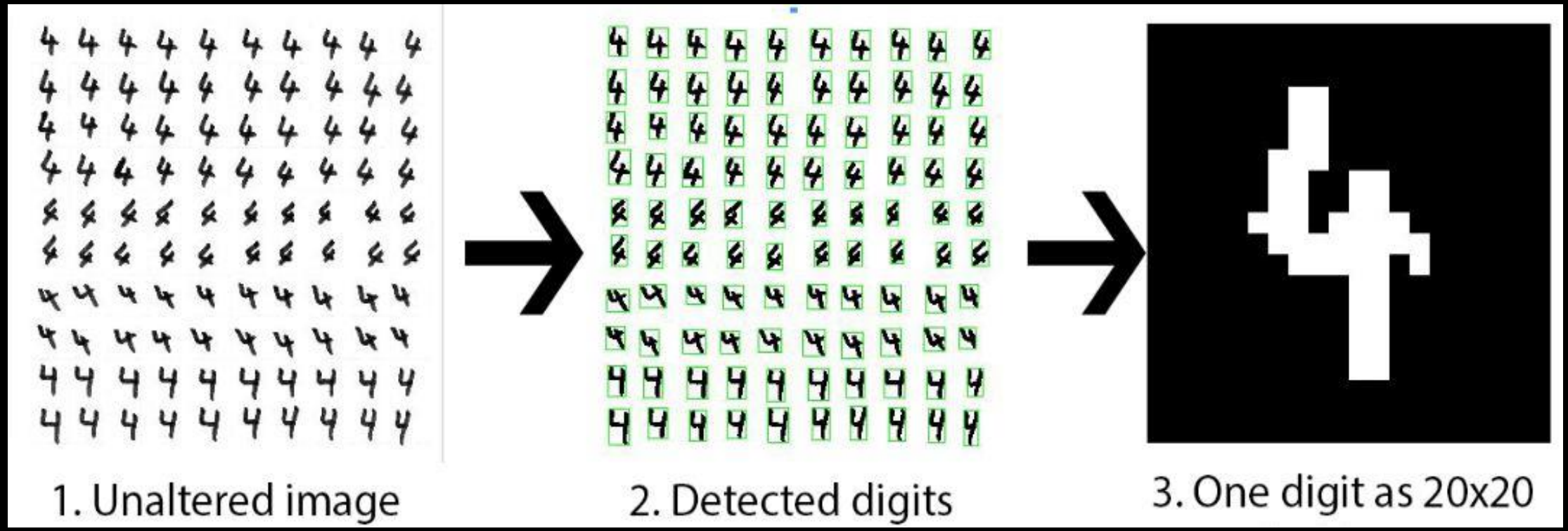


Got It???

Let's Goooooo!!!

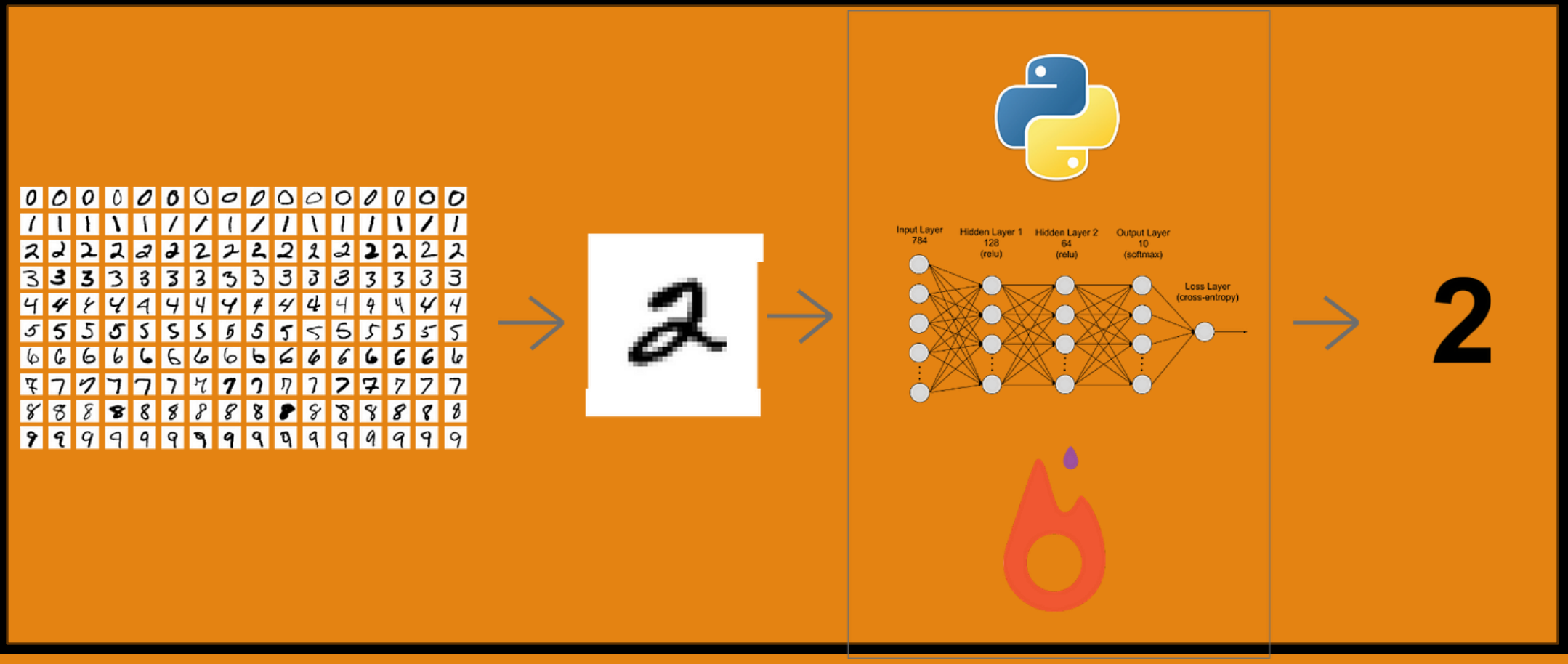
What is Deep Learning (DL)

Neural networks? How?



What is Deep Learning (DL)

Neural networks? How?

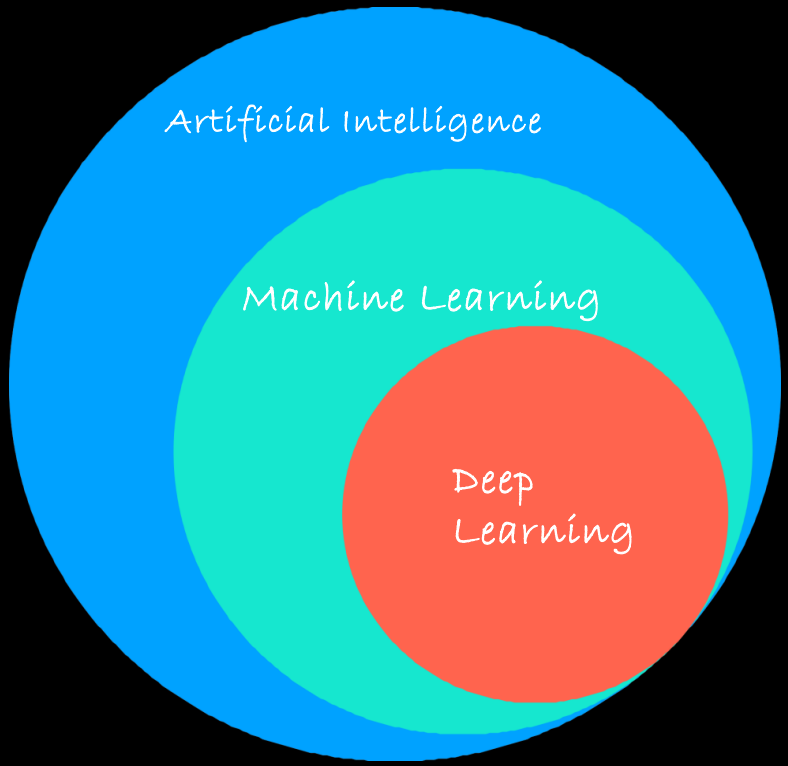


AI vs ML vs Deep Learning (DL)

AI is like a **robotic assistant** that can follow instructions,

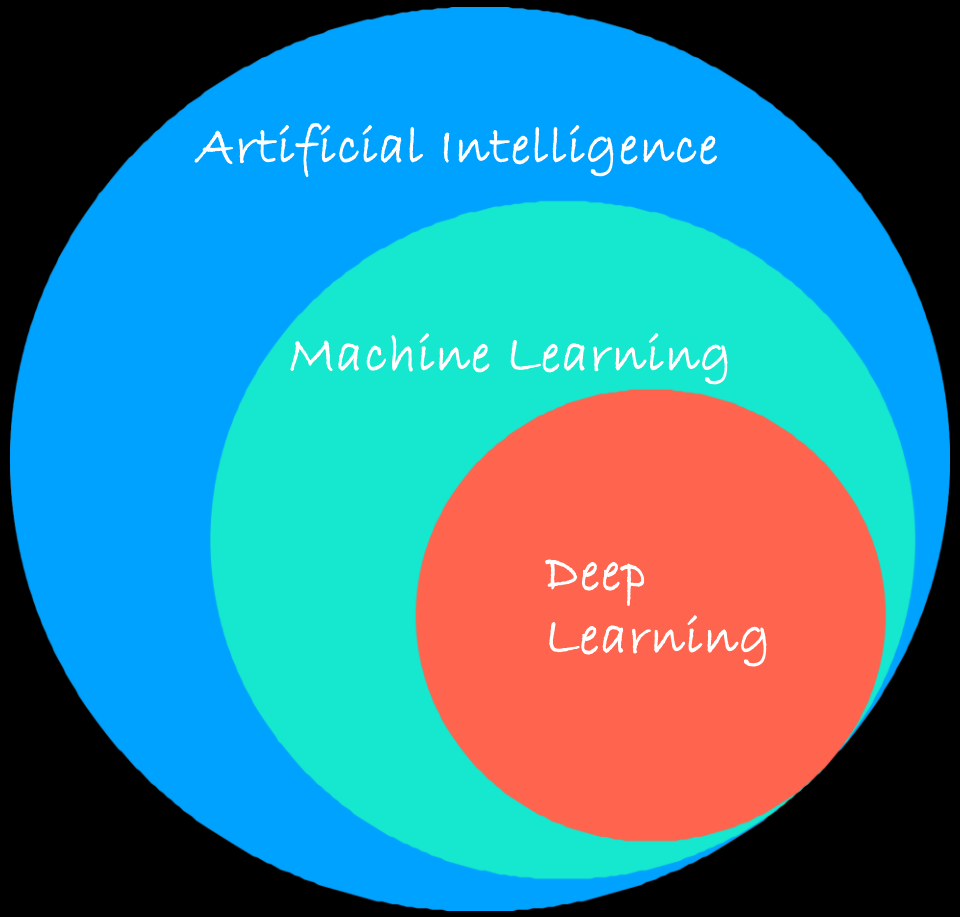
ML is when it learns from experience,

DL is when it learns so deeply that it can recognize faces, voices, and even create art! 🎨🤖



Where do we group NLP??

AI vs ML vs Deep Learning (DL)

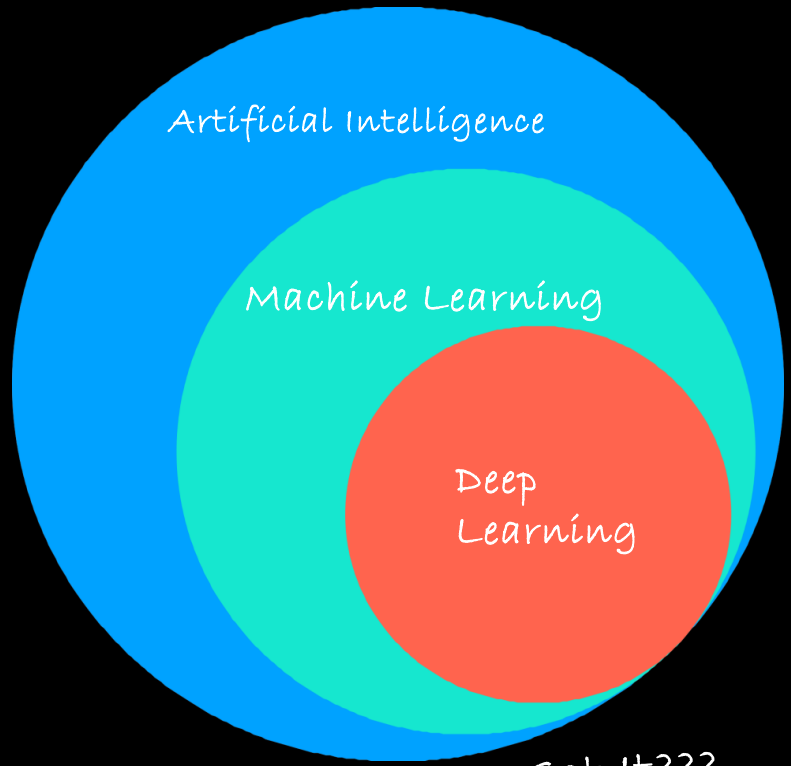


NLP is a sub of Deep Learning..

AI vs ML vs Deep Learning (DL)

Let's relate with this:

AI is like a kitchen assistant following recipes, ML is like a cook who tweaks dishes based on feedback, and DL is like a top chef who invents new recipes just by understanding flavors! 🍳🤖🔥

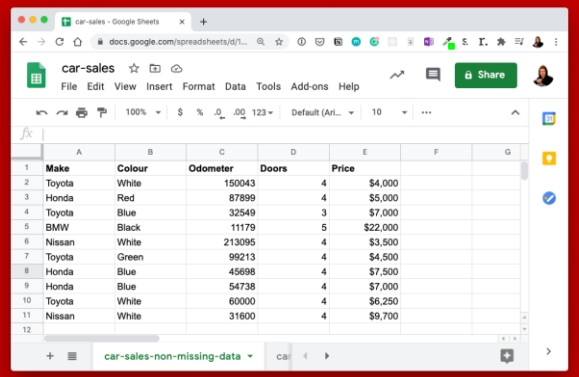


Got It???

Let's Gooooooo!!!

ML vs DL

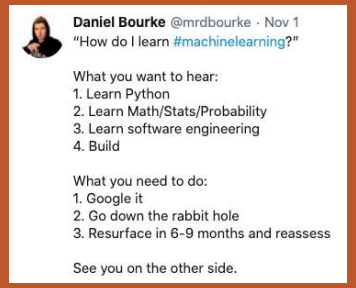
Machine Learning

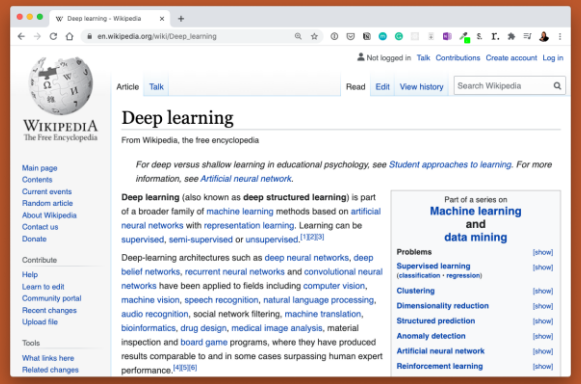


Make	Colour	Odometer	Doors	Price
Toyota	White	150043	4	$4,000
Honda	Red	87899	4	$5,000
Toyota	Blue	32549	3	$7,000
BMW	Black	11179	5	$22,000
Nissan	White	213095	4	$3,500
Toyota	Green	99213	4	$4,500
Honda	Blue	45698	4	$7,500
Honda	Blue	54738	4	$7,000
Toyota	White	60000	4	$6,250
Nissan	White	31600	4	$9,700

Structured data

Deep Leanring

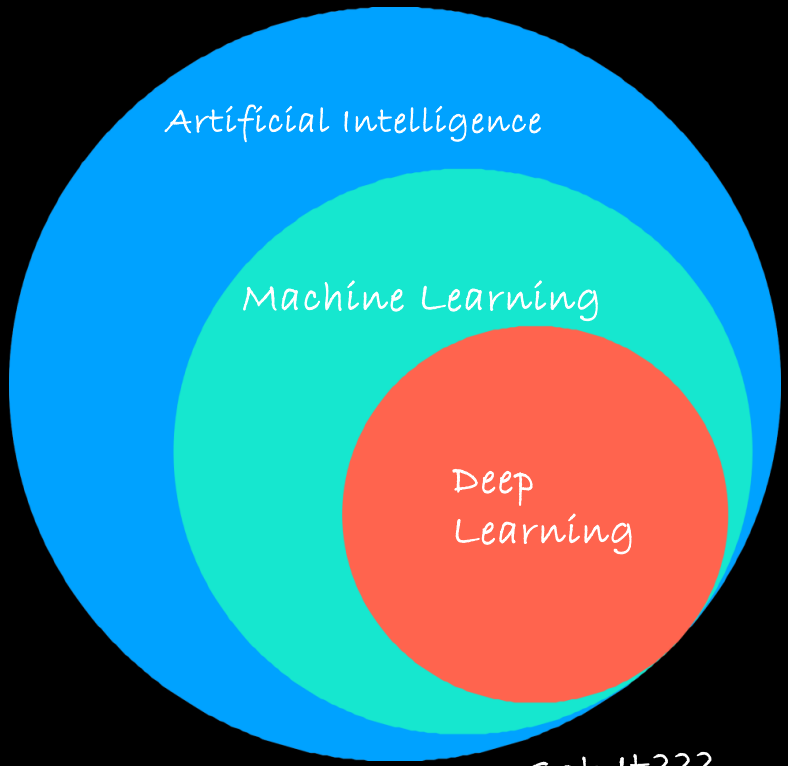




Unstructured Data

AI vs ML vs Deep Learning (DL)

Can AI exist without Machine Learning or Deep Learning???



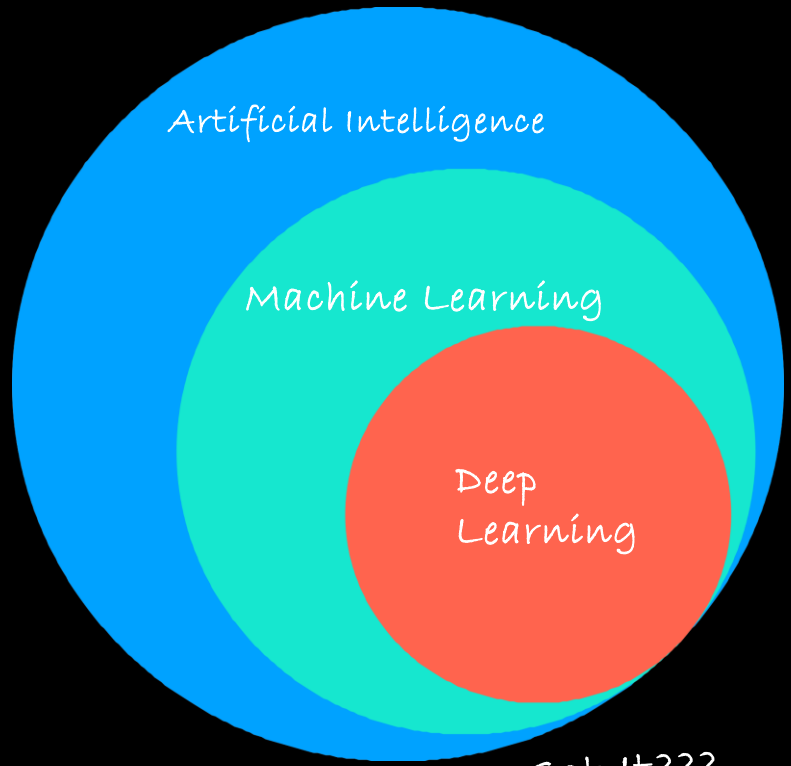
Got It???

Let's Goooooo!!!

AI vs ML vs Deep Learning (DL)

Can AI exist without Machine Learning or Deep Learning???

The answer is yes, AI can exist without ML or DL, although it might not be the kind of AI we often see in movies or science fiction.



Got It???

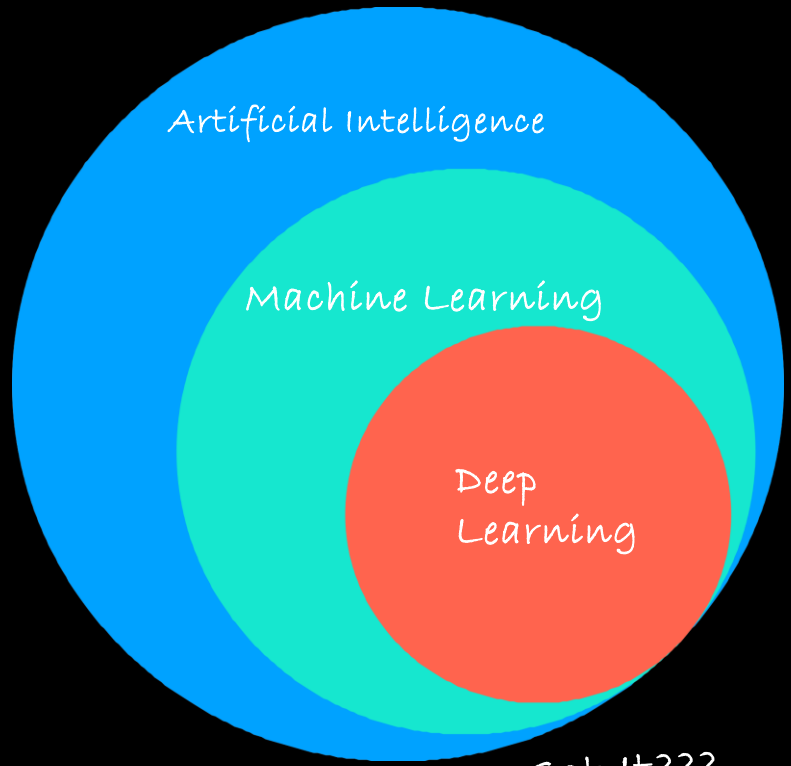
Let's Goooooo!!!

AI vs ML vs Deep Learning (DL)

AI encompasses any technique that enables computers to mimic human intelligence.

Rule-based systems were an early form of AI that relied on predefined rules rather than learning from data.

Machine Learning and Deep Learning, although dominant today, are subsets of AI that use algorithms to learn from data.



Got It???

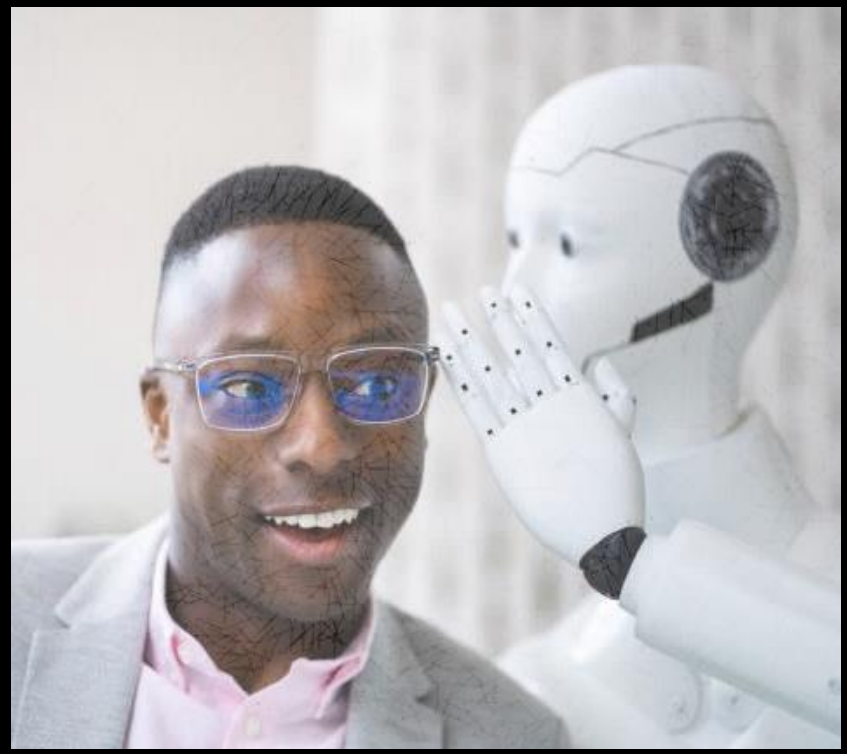
Let's Goooooo!!!

What is Natural Language Processing (NLP)

NATURAL LANGUGAE PROCESSING (NLP) helps computers understand and talk like humans.

It enables computers to read, interpret, and generate human-like text.

It allows your phone to predict your next word, understand voice commands, and translate languages.

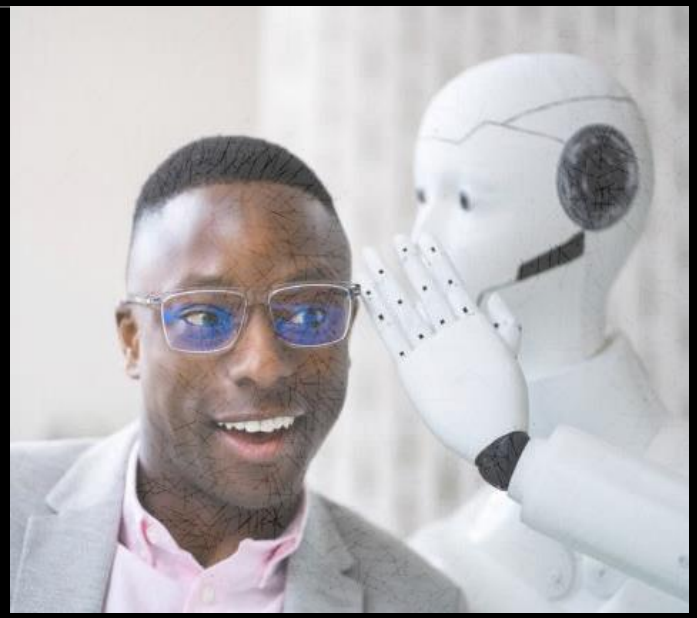


What is Natural Language Processing (NLP)

NLP combines ML and Deep Learning to process and understand human language.

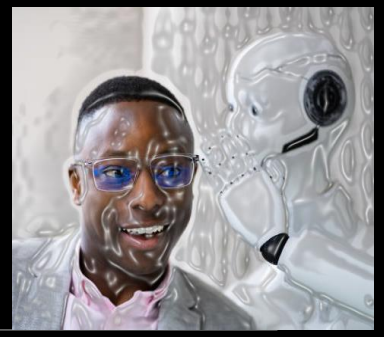
Includes tasks like stemming, tokenization, NER, sentiment analysis.

NER – named entity recognition



Here is a scenario; Imagine chatting with a smart assistant (like ChatGPT or Siri). You say, "What's the weather today?", and it understands your words, fetches present weather data, and responds. That's NLP in action.

Natural Language Processing (NLP) in Industry



Voice Assistants 🎙️ **:** Siri, Alexa, and Google Assistant understand and respond to your voice commands.

Chatbots 💬**:** Websites use NLP-powered bots to answer your questions instantly.

Autocorrect & Predictive Text 📱 **:** Your phone suggests next words while typing to speed things up, Also can correct spellings when you typed wrongly.

All this industry examples uses NLP in operation.

Natural Language Processing (NLP) in Industry

Translation Apps 🌍: Google Translate helps you understand different languages.

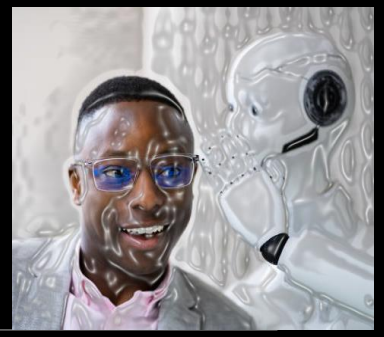
Spam Filters 📧: Your email sorting out spam messages automatically.

Search Engines 🔍: To understands what you mean, even if you type a messy query.

Sentiment Analysis 😀😡: Companies analyze tweets and reviews to see what people feel about their brand.

All this industry examples uses NLP in operation. Can I get from you??

NLP use cases



"What's the topic of this text?" (text classification)

"Does this text contain abuse?" (content filtering)

"Does this text sound positive or negative?" (sentiment analysis)

"What should be the next word in this incomplete sentence?" (language modeling)

"How would you say this in German?" (translation)

"How would you summarize this article in one paragraph?" (summarization)

Got The Point Now????

Brief History of Natural Language Processing (NLP)

2011: George Dahl and Microsoft Research trained a deep neural network to recognize words from speech recordings—a major deep learning breakthrough in NLP.

2012: AlexNet revolutionized deep learning in image recognition, leading to increased interest in machine vision.

2015: Deep learning models started competing in NLP tasks like machine translation, matching traditional ML accuracy with less effort and lower computational cost.

2016-2017: Deep learning models in NLP became more efficient and began outperforming traditional ML approaches in accuracy.

Impact: Enabled real-time translation on mobile devices without requiring an internet connection, making NLP more accessible and practical.

What's Next 😀 😀

Maybe I don't know also

After all of the theoretical explanation, can we begin some hands-on or we stop here today?

What's Next

Maybe I don't know also

If Yes, 

Let's get our Machine setup for hands-on.

Setting up our PC

We need install some tools if we are to work locally (i.e without internet)

Since we are online, we may not delve into this page in details

If you will be working locally, make sure to install the following

1. Python (ensure to add to path when installing)
2. Jupyter Lab (I like to use command - pip install jupyterlab – in my CMD/Terminal)
3. Tensorflow (use the command - pip install tensorflow – into your CMD/Terminal)
4. Some frameworks and libraries we will use as we move

If working online,
visit google colab and let's start

What's Next 😀 😀

Maybe I don't know also

After setup,
Let's begin proper proper

The END

DSN LEKKI-AJAH

BY ABEREJO HABEEBLAH O.

X: @HABEREJO

Day 1

NATURAL LANGUAGE PROCESSING (NLP)

TEXT PREPROCESSING AND TOKENIZATION

DSN LEKKI-AJAH

BY ABEREJO HABEEBLAH O.

X: @HABEREJO

Day 2

How computer sees language

Computers see text as a sequence of character,

Computers don't understand words or meaning like humans do

They see all English words as binary language (i.e. 1s and 0s)

So, now, we need to help computer understand our language to what it can fully understand

What do we do to help computers understand the text? Let's GO

Binary Language == 1 and 0

Understand Human Language

There are arguably 3 steps to help computer understand human texts,

1. Text Preprocessing
2. Tokenization
3. Numerical Representation (Word Embedding)

Now, Lets discuss each further!!

Binary Language == 1 and 0

Step 1: Text Preprocessing

Why Preprocess Text?

Imagine you have a room full of clothes, toys, and papers scattered everywhere. Finding what you need is hard, right?

Raw text data is similar. It's full of "**noise**" – things that make it difficult for computers to understand the meaning.

Got It???

Let's Goooooo!!!

Step 1: Text Preprocessing

At this stage, lets understand preprocessing of text,

Proper Text preprocessing is important to make sure we are only feeding the right processed information into our algorithm and irrelevant text are gone.

e.g: The Quick! bRoWn foX juMPS.

Got It???

Let's Goooooo!!!

Step 1: Text Preprocessing

Text preprocessing cleans and prepares text data for NLP tasks. It addresses issues like inconsistencies (uppercase/lowercase), irrelevant information (stop words), and variations in word forms (running vs. run). This makes it easier for NLP models to learn.

e.g: The Quick! bRoWn foX juMPS.

Got It???

Let's Goooooo!!!

Step 1: Text Preprocessing

We prepare text by:

1. Lowercasing
2. Removing Punctuation Marks
3. Removing Stopwords (i.e words like A, of, in, etc)
4. Stemming / Lemmatization

e.g: The Quick! bRoWn foX juMPS.

NB:
There is no specific way to prepare text, it is much dependent on what type of NLP project we are building!

Step 1: Text Preprocessing

1. Lowercasing: To humans, A capitalized word at the beginning of a sentence (e.g., She) has the same meaning as when it's used later in a sentence (she).

By converting all characters in a corpus to lowercase, we disregard any use of capitalization.

Computers treat "The" and "the" as different words. Lowercasing makes everything consistent.

The Quick! bRoWn foX juMPS.

the quick! brown fox jumps.

NB:
Corpus in this regard means collection of writings.

Step 1: Text Preprocessing

Though, in a larger corpus that has many more examples of individual uses of words, the word, **general (an adjective meaning "widespread")** have different meaning to the word **General (a noun meaning the commander of an army).** One option is not to lowercase everything, or try an option called **Part-Of-Speech (POS) Tagging**

The Quick! bRoWn foX juMPS.

the quick! brown fox jumps.

NB:
Corpus in this regard means collection of writings.

Step 1: Text Preprocessing

2. Removing Punctuation Marks: Punctuation marks generally don't add much value to a natural language model and so are **often** removed.

Removing punctuation would not be an advantage in all cases. Consider, for example, if you were building a question-answering algorithm, which could use question marks to help it identify questions.

The Quick! bRoWn foX juMPS.

the quick brown fox jumps

Depending on the NLP project, some steps are compulsory

Step 1: Text Preprocessing

3. Removing Stopwords: Stopwords are frequently occurring words that tend to contain relatively little distinctive meaning, such as **the**, **at**, **which**, **and**, **of** etc.

There is no universal consensus on the precise list of stop words, but depending on your application it may be sensible to ensure that certain words are (or aren't!) considered to be stop words.

The Quick! bRoWn foX juMPS.

quick brown fox jumps

Depending on the NLP project, some steps are compulsory

Step 1: Text Preprocessing

For example, when building a model to **classify movie reviews** as **positive** or **negative**. Some lists of stop words include negations like **didn't, isn't, and wouldn't** that might be critical for our model to identify the sentiment of a movie review, so these words probably shouldn't be removed.

Negations may be helpful as stop words for some classifiers but probably not for a sentiment classifier

So be careful with this one. In many instances, it will be best to remove only a limited number of stop words.

The Quick! bRoWn foX juMPS.

quick brown fox jumps

Depending on the NLP project, some steps are compulsory

Step 1: Text Preprocessing

4. Stemming / Lemmatization:

Stemming: Stemming is the truncation of words down to their stem, **i.e their root form.**

For example, the words **house**, **housed** and **housing** both have the stem **hous**.

With smaller datasets in particular, stemming can be productive because it pools words with similar meanings into a single token.

The Quick! bRoWn foX juMPS.

quick brown fox jumps

Depending on the NLP project, some steps are compulsory