

Artificial Intelligence

Shreeja Das & Dana Tafliovich

Possibilities With AI



ARTIFICIAL INTELLIGENCE VS MACHINE LEARNING VS DEEP LEARNING

1 Artificial Intelligence

Development of smart systems and machines that can carry out tasks that typically require human intelligence

2 Machine Learning

Creates algorithms that can learn from data and make decisions based on patterns observed
Require human intervention when decision is incorrect

3 Deep Learning

Uses an artificial neural network to reach accurate conclusions without human intervention

The background is a dark blue field decorated with various geometric elements. There are numerous small squares in solid colors (pink, orange, teal) and as thin white outlines. Some of these squares are connected to the top or bottom edges by thin, vertical white lines, creating a sense of depth or suspension. The overall aesthetic is modern and minimalist.

ARTIFICIAL INTELLIGENCE.

WHAT IS AI?

- Simulation of human intelligence by machines
 - Ex: NLP, Computer vision, speech recognition, etc
- Machine learning is a branch of artificial intelligence (AI)
- ML systems ingest large amounts of training data, analyzing the data for correlations and patterns
- These patterns are used to make predictions about future states
 - Eg: Chatbot is fed examples of text through which it can learn to generate similar text



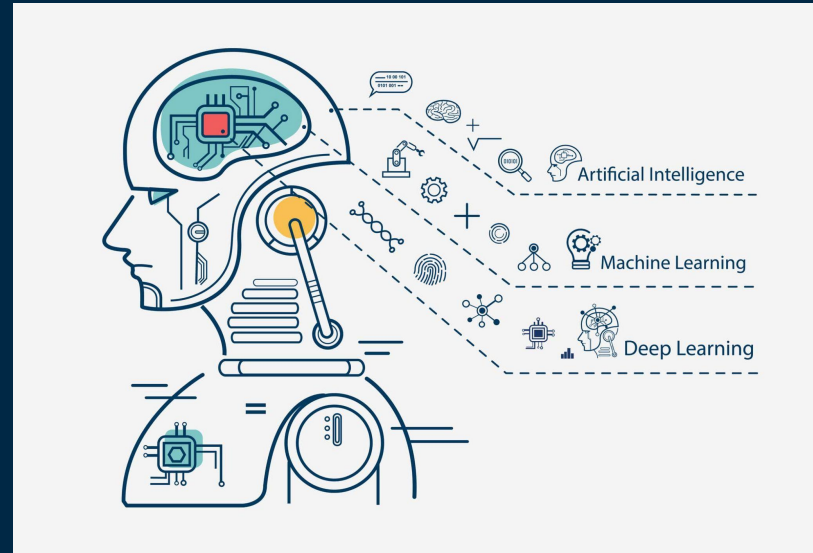
The background is a dark navy blue. It is decorated with various geometric elements: thin white vertical lines of varying lengths, small squares in teal, orange, and pink, and larger squares in teal and orange. The text 'MACHINE LEARNING' is centered in a large, white, sans-serif font.

MACHINE LEARNING

What is Machine Learning?

Machine learning is an important component of the growing field of data science. Through the use of statistical methods, algorithms are trained to make classifications or predictions, and to uncover key insights in data mining projects

- Machine learning algorithms are typically created using frameworks that accelerate solution development, such as TensorFlow and PyTorch



Timeline

Golden Age of
Science Fiction

1938-1946

1949-1950

"Can Machines Think" -
Alan Turing

The Enigma machine was
decoded using AI in World
War

"Logic Theorists" - the first
true AI program is invented

John McCarthy coins the
term: AI

1955

1961-1995

Invention of the first
industrial robot,
chatbot, and general
purpose mobile
robot

Timeline

Deep blue AI
algorithm beats
chess legend

1997

2008-2014

Voice Recognition on
iPhone and the birth of Siri
+ development of Alexa

Sophia the Robot becomes
the first robot to receive
citizenship

2016

2023

GPT-3 is introduced
and continues to
make huge strides in
cyber security, and
automation

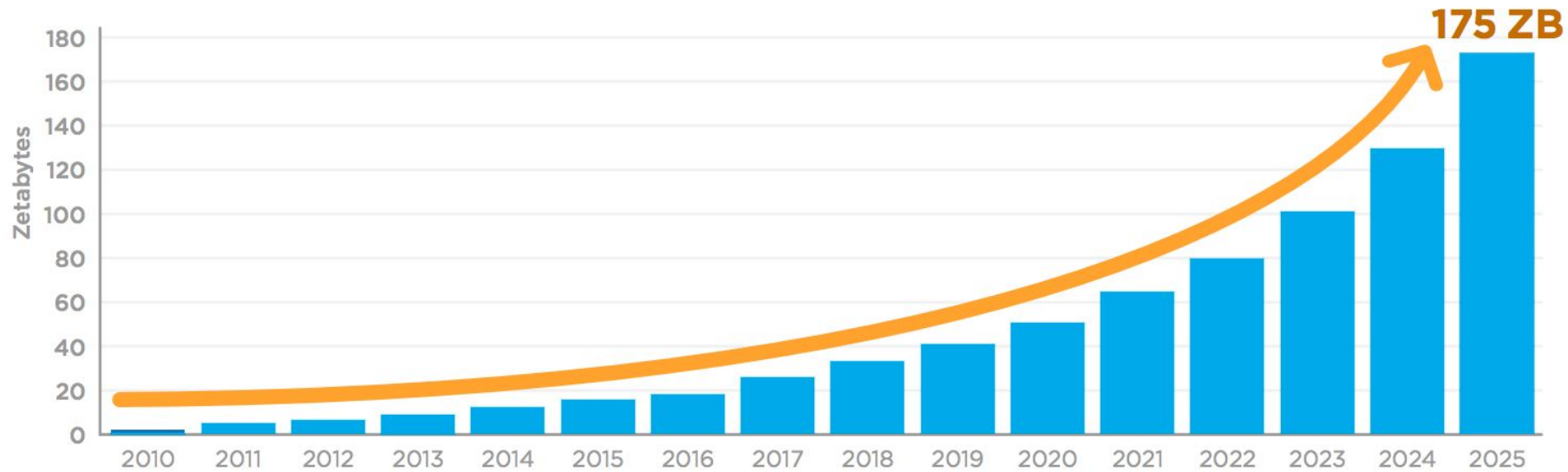
Generative AI & Predictive AI

- Type of AI that can create a wide variety of data, such as images, videos, audio, text, and 3D models
 - Learns patterns from existing data
 - GPT (Generative Pre-trained Transformer)
- Predictive artificial intelligence (AI) refers to the use of machine learning to identify patterns in past events and make predictions about future events

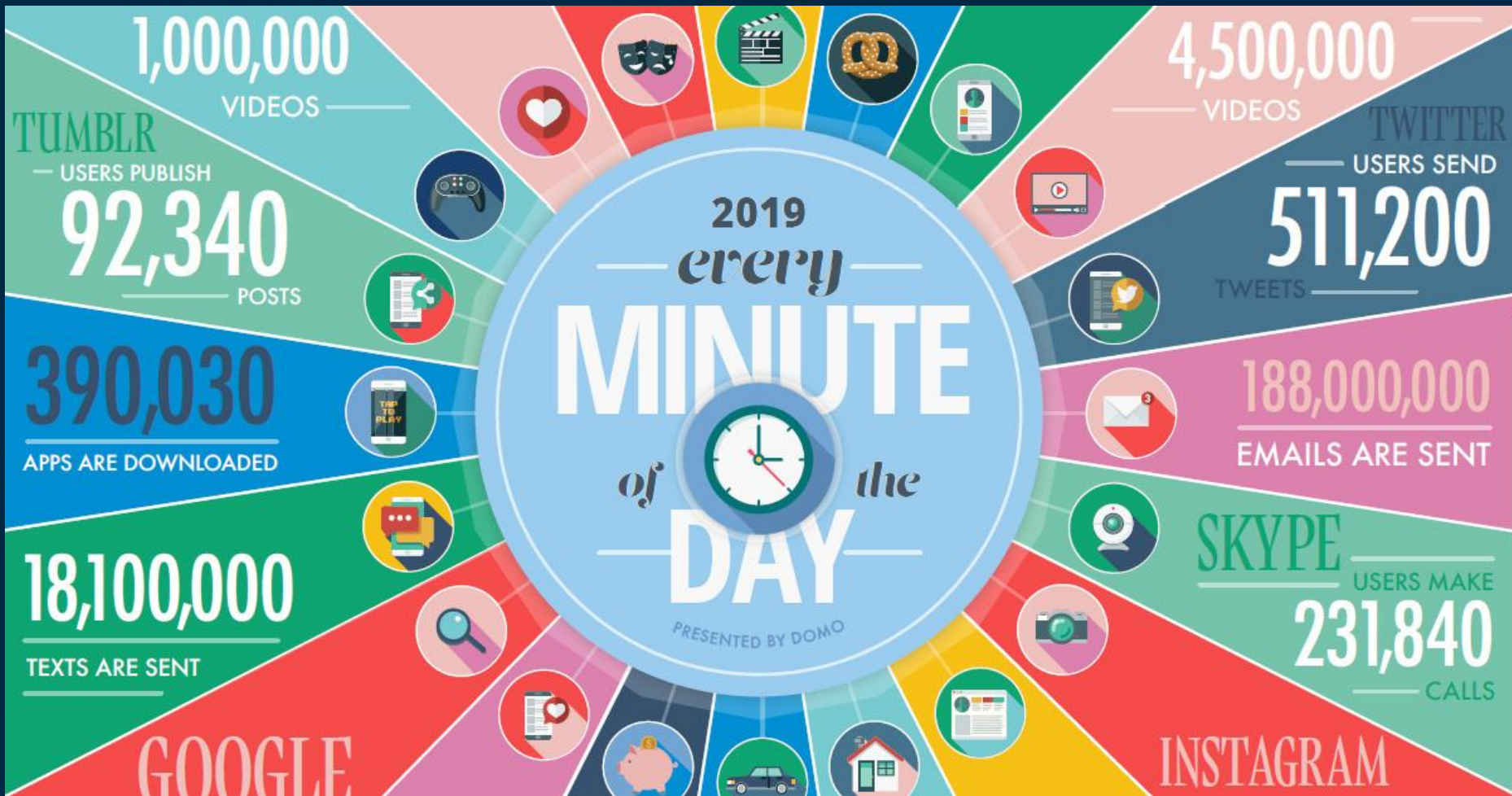
“Big Data”

Velocity, Veracity (variation), & Volume

In statistics more data generally results in more accurate analysis. For example, an opinion poll must have a minimum number of respondents to be considered reliable, and scientific studies need to be repeated several times to be considered statistically significant.



How big is “Big Data”





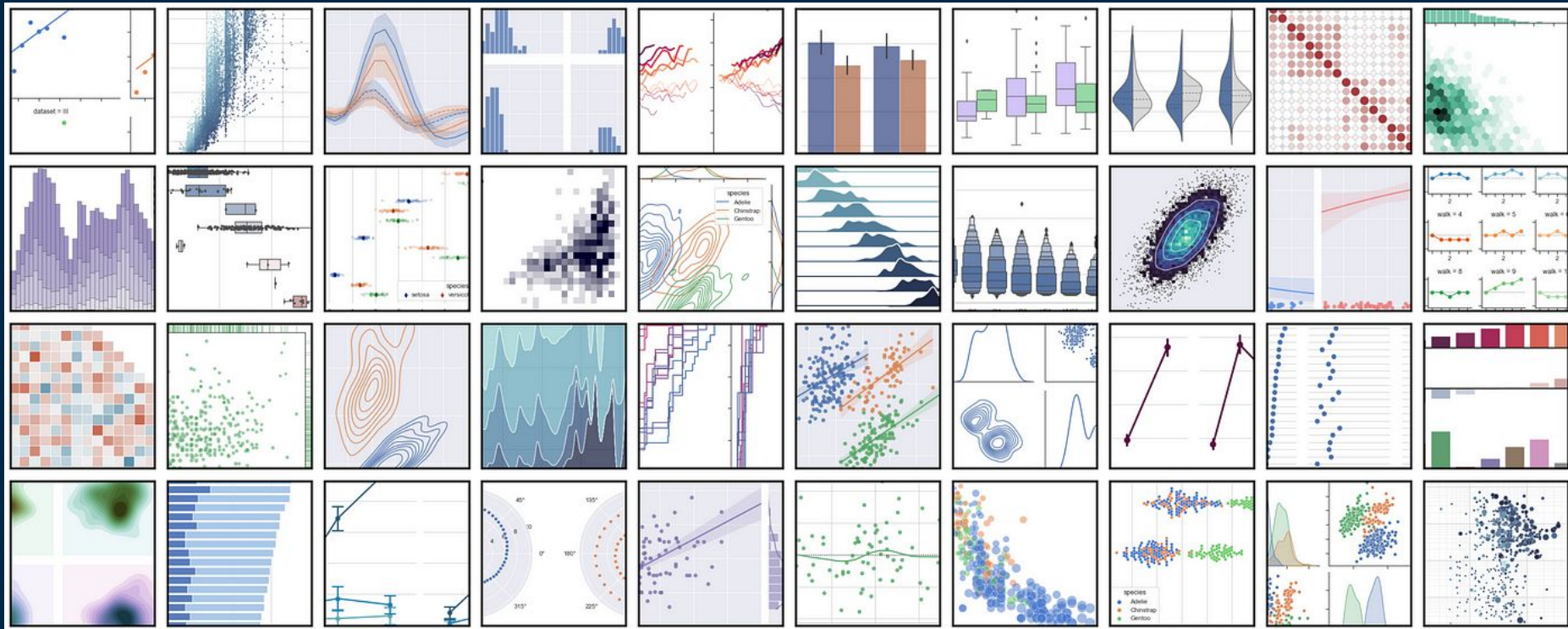
GPT 4 was trained on...

**13 TRILLION
TOKENS**

This is why it's known as a LLM

STEPS OF ML

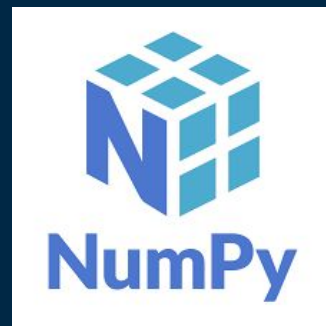
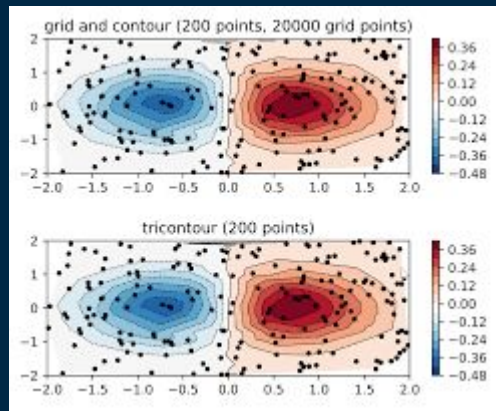
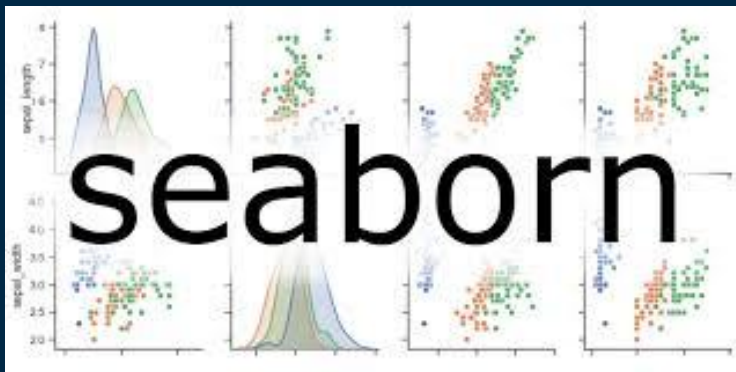
- Writing, training, and testing machine learning algorithms → AI end products
- **Learning**: acquiring data and creating rules for how to turn it into actionable information, algorithms provide computing devices with step-by-step instructions for how to complete a specific task
- **Reasoning**: choosing the right algorithm to reach a desired outcome.
- **Self-correction**: continually fine-tuning they provide the most accurate results possible.
- **Creativity**: using neural networks, rules-based systems, statistical methods and other AI techniques to generate new images, new text, new music and new ideas



DATA VISUALIZATION & ANALYSIS

Further Breakdown

- Data Cleaning
- Exploratory Data Analysis



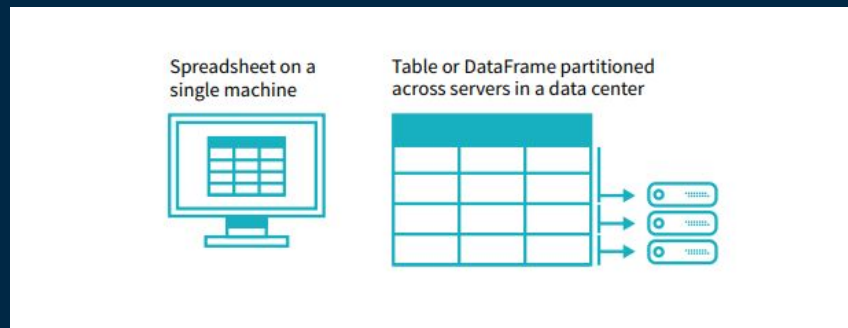
How is Data Processed in ML?

A DataFrame is a data structure that organizes data into a 2-dimensional table of rows and columns, much like a spreadsheet

DataFrame vs. Table:

They are similar.

Data tables are data frames but not all data frames are not necessarily data tables (data tables are an inheritance of dataframes)

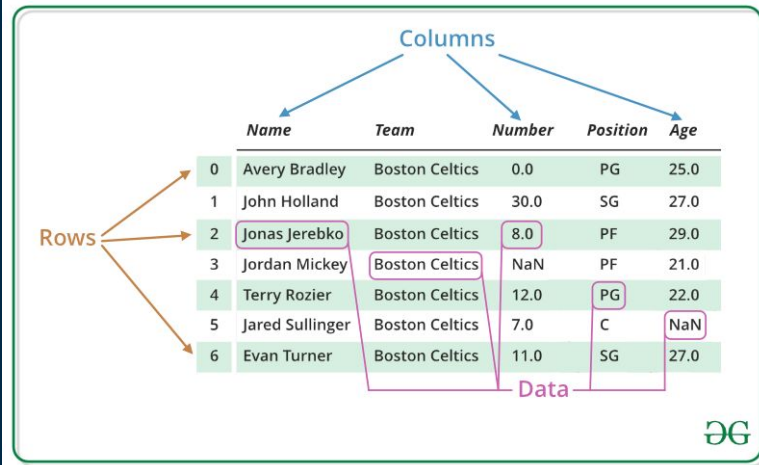


Series			Series			DataFrame	
	apples			oranges			
0	3	+	0	0	=	0	3
1	2		1	3		1	2
2	0		2	7		2	0
3	1		3	2		3	1
							oranges
							0
							3
							7
							2

Pandas + Dataframe

How to process these dataframes?

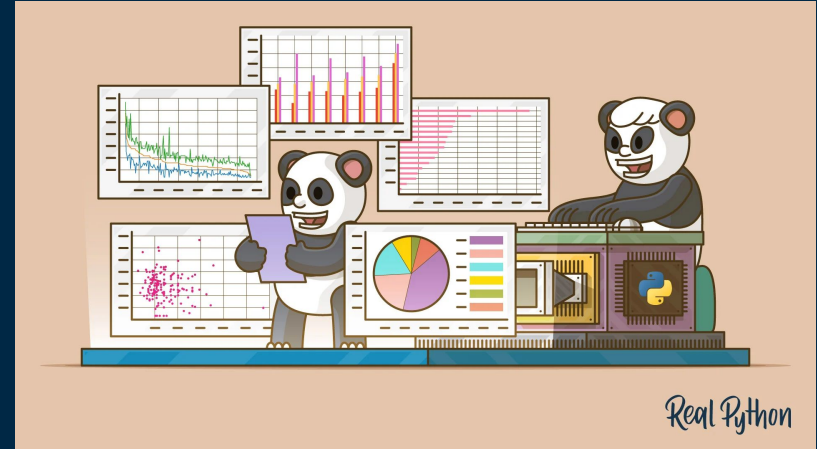
https://images.datacamp.com/image/upload/v1676302204/Marketing/Blog/Pandas_Cheat_Sheet.pdf



The diagram illustrates the structure of a Pandas DataFrame. It shows a table with 7 rows and 6 columns. The columns are labeled 'Name', 'Team', 'Number', 'Position', and 'Age'. The rows are indexed from 0 to 6. A label 'Columns' points to the column headers. A label 'Rows' points to the row indices. A label 'Data' points to the data cells. The table contains the following data:

	Name	Team	Number	Position	Age
0	Avery Bradley	Boston Celtics	0.0	PG	25.0
1	John Holland	Boston Celtics	30.0	SG	27.0
2	Jonas Jerebko	Boston Celtics	8.0	PF	29.0
3	Jordan Mickey	Boston Celtics	NaN	PF	21.0
4	Terry Rozier	Boston Celtics	12.0	PG	22.0
5	Jared Sullinger	Boston Celtics	7.0	C	NaN
6	Evan Turner	Boston Celtics	11.0	SG	27.0

The diagram also includes a small logo in the bottom right corner.



Machine Learning Approaches

Supervised Learning

- Labeled datasets
- Classify and predict outcomes
- Can measure it's accuracy over time
- These datasets are designed to train and "supervise" algorithms to make accurate predictions

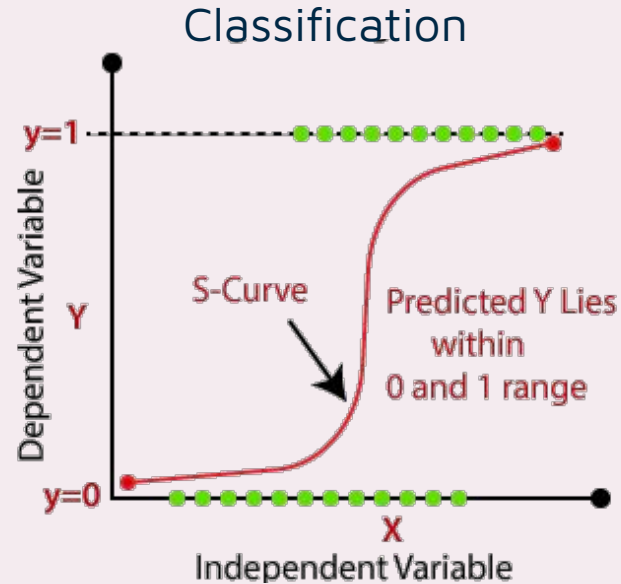
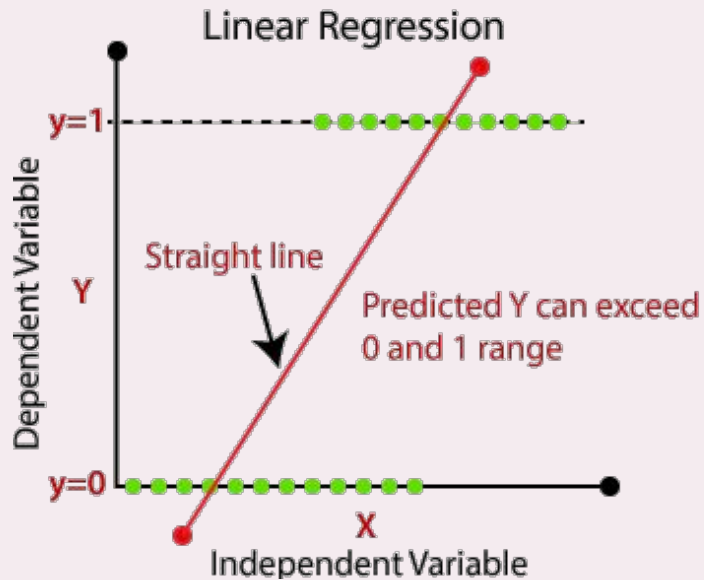
Unsupervised Learning

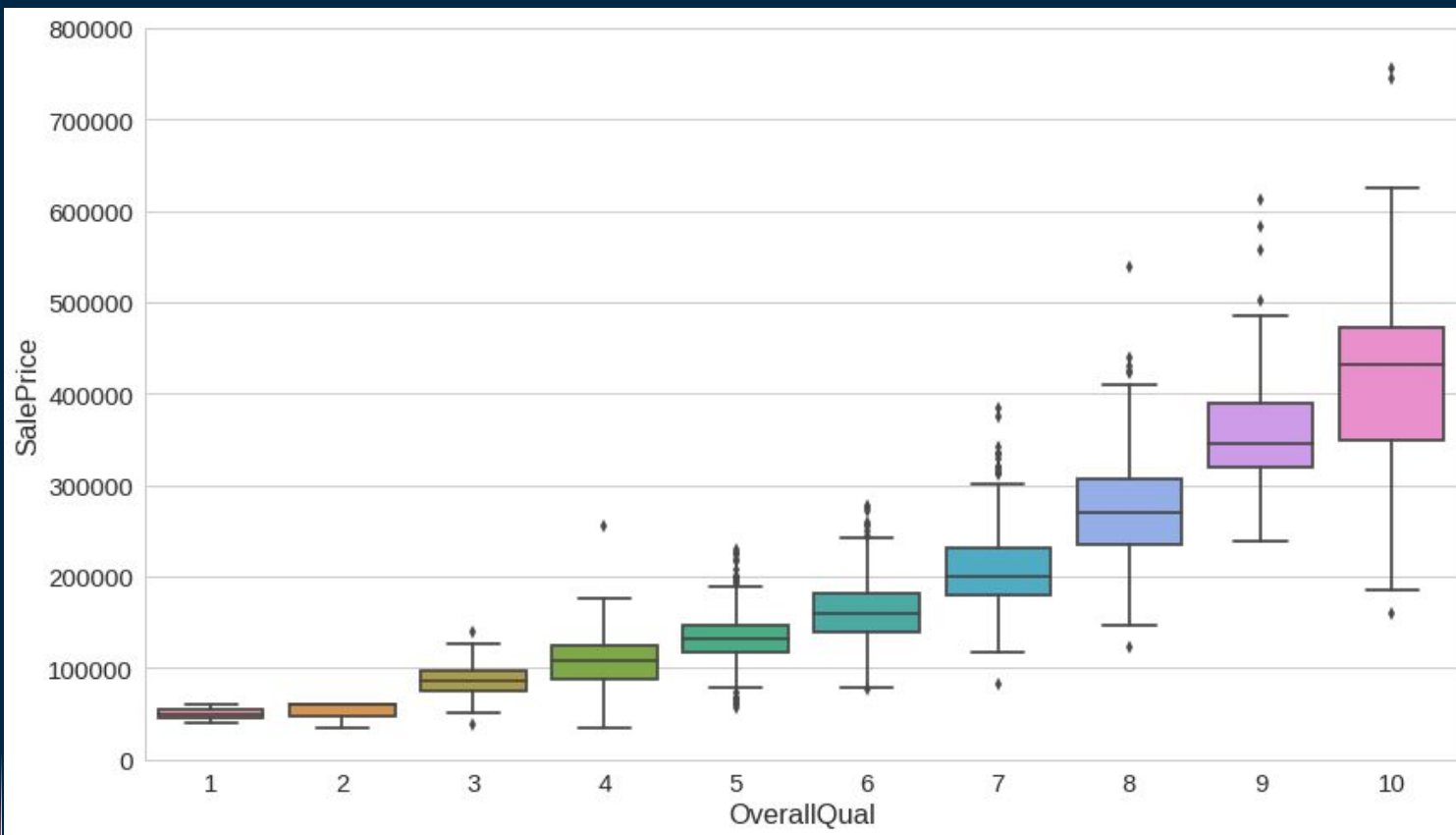
- Unlabeled datasets
- Used to analyze and classify unlabeled datasets
- Algorithms discover patterns in data without human intervention

Supervised Learning

- Classification
 - Accurately assign test data to categories (eg: separating apples from oranges)
 - Eg: Algorithm to classify spam emails from regular emails
 - Common Algorithms: Linear classifiers, SVM, Decision Trees, Random Forest
- Regression
 - Algorithm to understand the relationship between dependent and independent variables
 - Good for predicting numerical values based on different data points
 - Eg: Sales revenue projections for next quarter

Regression Vs. Classification



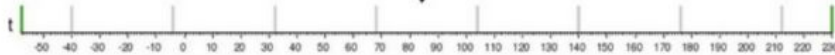


Eg: House Prices using Linear Regression



What will be the temperature
tomorrow?

84°



Fahrenheit



Will it be hot or cold
tomorrow?

COLD

HOT



Fahrenheit

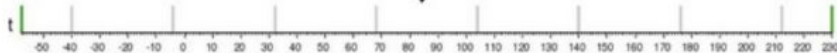
POP QUIZ!

REGRESSION



What will be the temperature tomorrow?

84°



Fahrenheit

CLASSIFICATION



Will it be hot or cold tomorrow?

COLD

HOT



Fahrenheit

POP QUIZ!

Unsupervised Learning

Clustering

Grouping unlabeled data based on their similarities or differences

Eg: K-means clustering algorithms assign similar data points into groups, where the K value represents the size of the grouping

Association

Uses different rules to find relationships between variables in a dataset

Eg: used for recommendation engines, along the lines of "Customers Who Bought This Item Also Bought" recommendations.

<https://www.naftaliharris.com/blog/visualizing-dbscan-clustering/>

Even more types of ML!

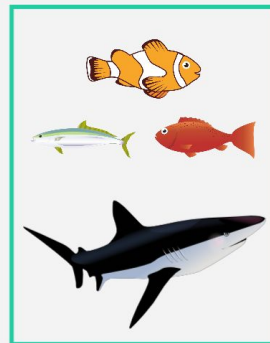
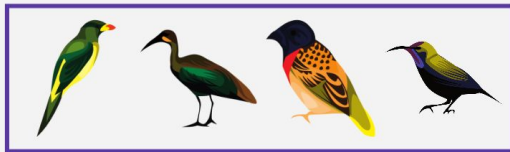
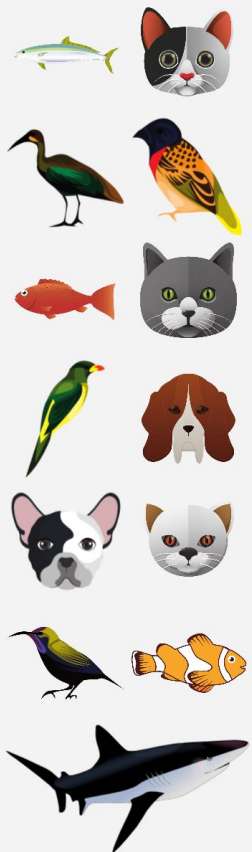
Semi Supervised Learning

- A medium between supervised and unsupervised learning
-
- During training, it uses a smaller labeled data set to guide classification and feature extraction from a larger, unlabeled data set.
 - Semi-supervised learning can solve the problem of not having enough labeled data for a supervised learning algorithm

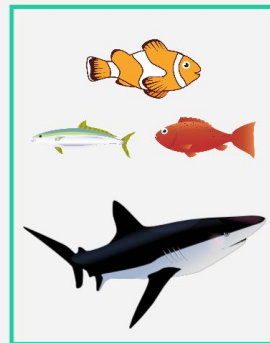
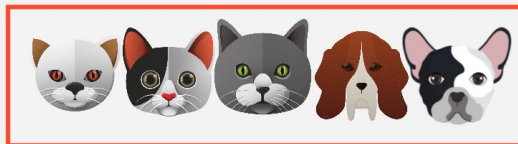
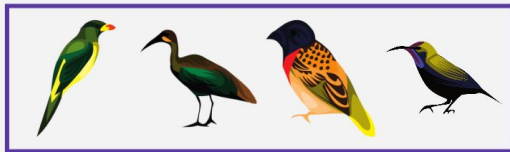
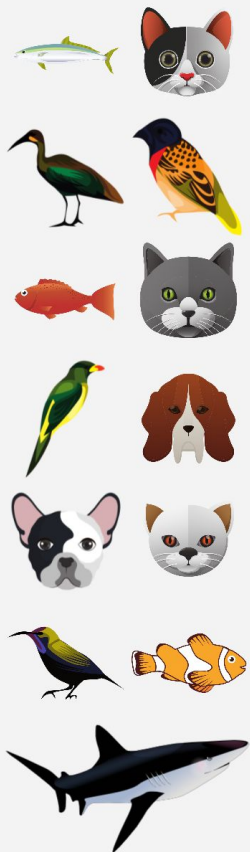
Reinforced Learning

Reinforcement machine learning is a machine learning model that is similar to supervised learning, but the algorithm isn't trained using sample data. This model learns as it goes by using trial and error. A sequence of successful outcomes will be reinforced to develop the best recommendation or policy for a given problem

No labels



No labels



UNSUPERVISED
LEARNING

DEEP LEARNING

The background is a dark blue gradient. It is decorated with various geometric elements: small squares in teal, orange, and pink, and thin white vertical lines of varying lengths. These elements are scattered across the frame, creating a modern, tech-inspired aesthetic.

What is Deep Learning?

ML Algorithms:

- Leverage structured, labeled data for predictions
- Specific features are defined and organized into tables
- May involve pre-processing of unstructured data to structure it

Deep Learning Advantages:

- Eliminates some data pre-processing in comparison to traditional ML
- Can ingest and process unstructured data (e.g., text, images)
- Automated feature extraction, reducing dependency on human experts

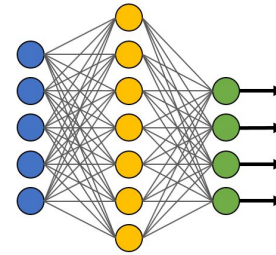
Example: Image Categorization:

- Set of pet photos categorized into "cat," "dog," "hamster," etc.
- Deep learning identifies crucial features (e.g., ears) for classification
- In machine learning, human experts manually establish feature hierarchies

Deep Learning

- Allows for the creation of complex neural networks

Simple Neural Network

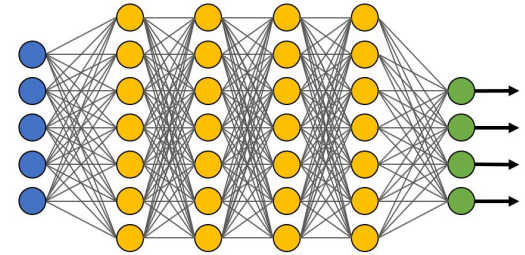


● Input Layer

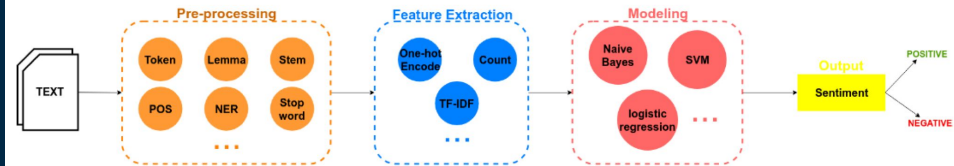
● Hidden Layer

● Output Layer

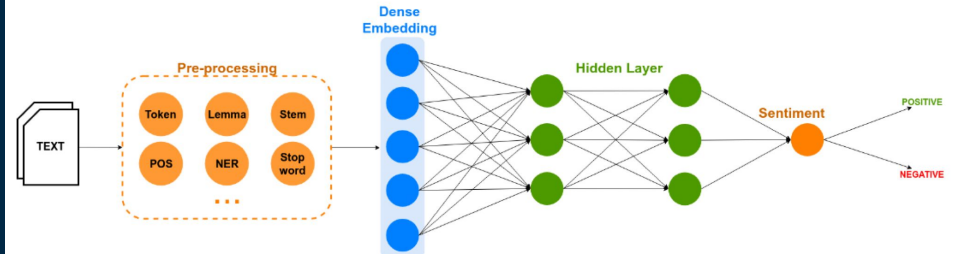
Deep Learning Neural Network



Machine Learning



Deep Learning



What is a Neural Network?

- Inspired by the human brain, mimics biological neurons' signaling.
- Comprised of node layers: input layer, one or more hidden layers, and an output layer.
- Nodes (artificial neurons) connect, each with weight and threshold.
- Need large amounts of training data & high GPU power
- Rapid clustering/classification (good at speech (RNN) and image recognition (CNN))

a. Eg: Google's search algorithm is a prominent example of a neural network.

Popular Machine Learning Algorithms

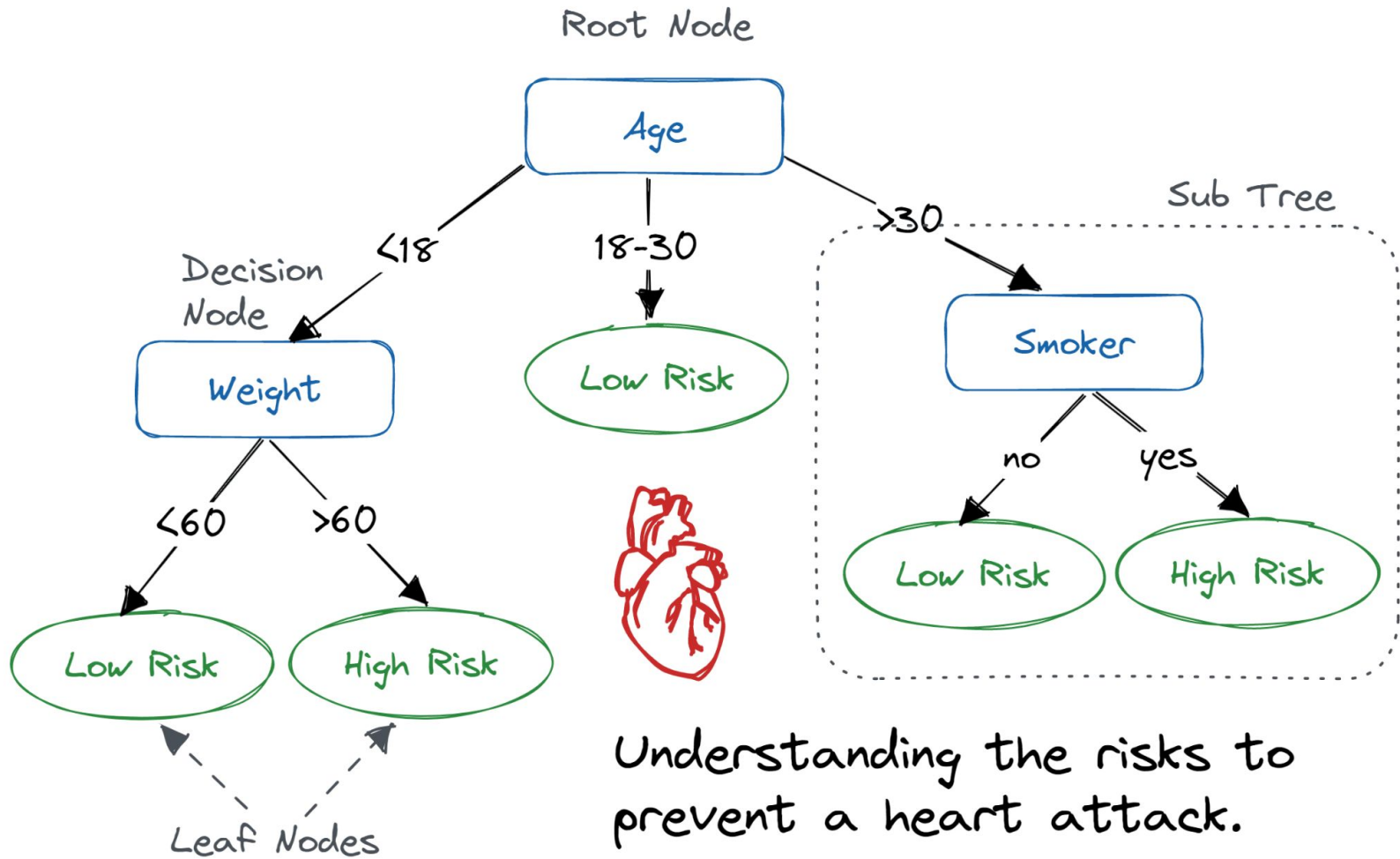
Decision Trees:

- Predicts numerical values (regression) and classifies data.
- Represents decisions in a branching tree diagram.

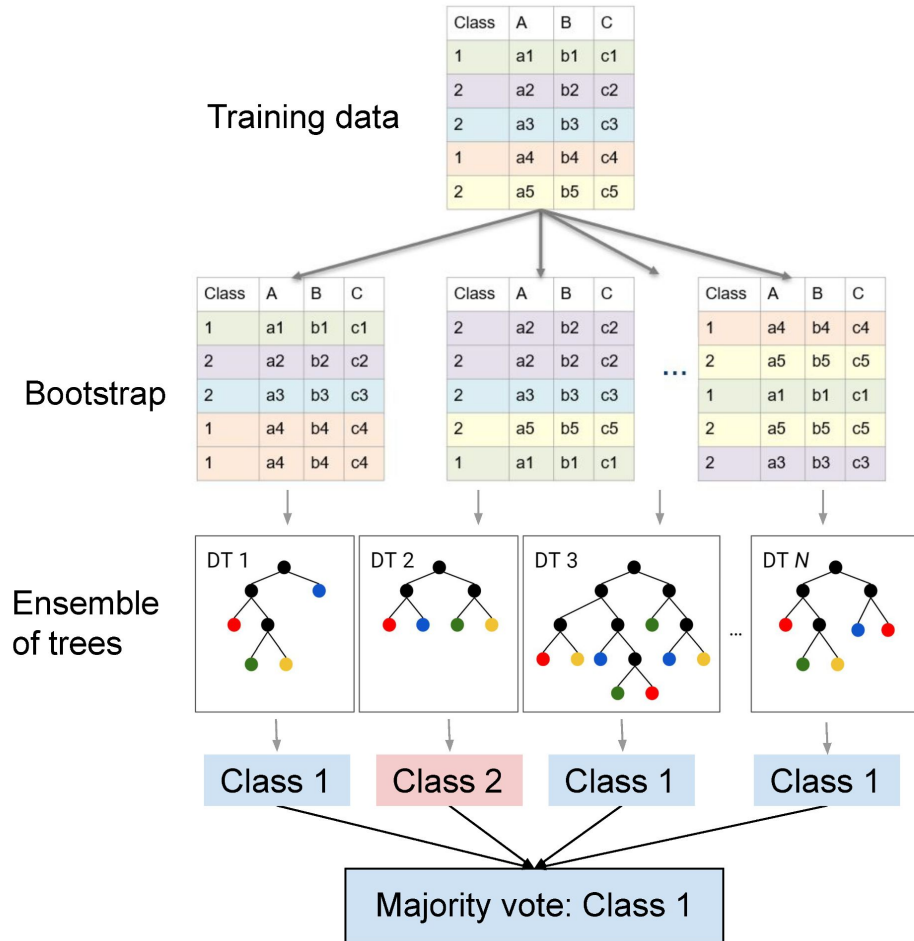
Random Forests:

- Machine learning algorithm combining results from multiple decision trees.
- Used for predicting values or categories.





Random Forest



Machine Learning

Python

A high-level, general-purpose programming language, that is easy and fast to learn (known for its readability)

Has many libraries like Tensorflow, Keras, PyTorch, NLTK for beginners and experts alike to train ML algorithms



PYTHON Review

```
print("Hello, World!")
```

Variables:

Store data

Ex:

```
x = 5
```

```
print(x)
```

```
y = "Hello World!"
```

Data Types

String, int, float, boolean, list

```
X = 5
```

```
print(type(x))
```



PYTHON (cont.)

String

```
str.len()
```

```
a = "Hello, World!"
```

```
print(a[1])
```

Boolean

```
print(10 > 9)
```

```
print(10 == 9)
```

Lists

```
thislist = ["apple", "banana", "cherry"]
```

```
print(thislist)
```

```
print(len(thislist))
```

```
print(thislist[0])
```

```
thislist[1] = "blackcurrant"
```

```
thislist.append("orange")
```

```
thislist.insert(1, "orange")
```

```
thislist.remove("banana") - remove occurrence
```

```
thislist.clear()
```



PYTHON (cont)

If/elif/else

```
a = 200
```

```
b = 33
```

```
if b > a:
```

```
    print("b is greater than a")
```

```
elif a == b:
```

```
    print("a and b are equal")
```

```
else:
```

```
    print("a is greater than b")
```

```
if not a > b:
```

```
    print("a not greater than b")
```

While loop

```
i = 0
```

```
while i < 6:
```

```
    print("i = " + i)
```

```
    i += 1
```

For loop:

```
for x in thislist:
```

```
    print(x)
```

```
for x in range(1,3)
```

```
    print(x)
```



PYTHON (cont)

Functions

```
def my_function():  
    print("Hello from a function")
```

```
my_function()
```

```
def func(a, b):  
    print(a>b)
```



CODE ALONG

01

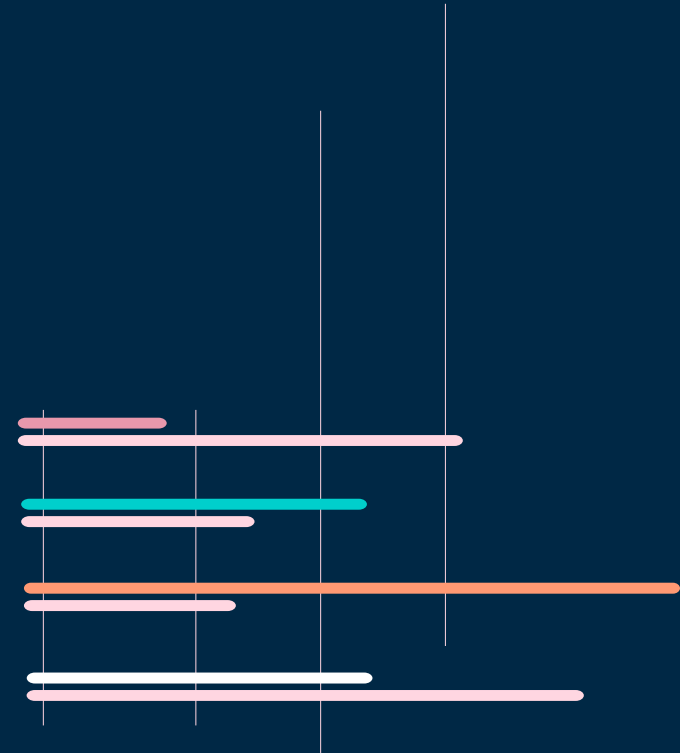
<https://www.kaggle.com/shreelearn/codealong>

Other Considerations

- Yay! We built our first ML model!!

Next steps?

- Data analysis
- Hypertuning
- Scaling
- Testing with different models
 - Overfitting vs. Underfitting?



Beware....

Overfitting

- Model fits training data too well.
- Learns noise and details, negatively impacting new data.
- Example: Decision trees prone to overfitting, addressed by pruning

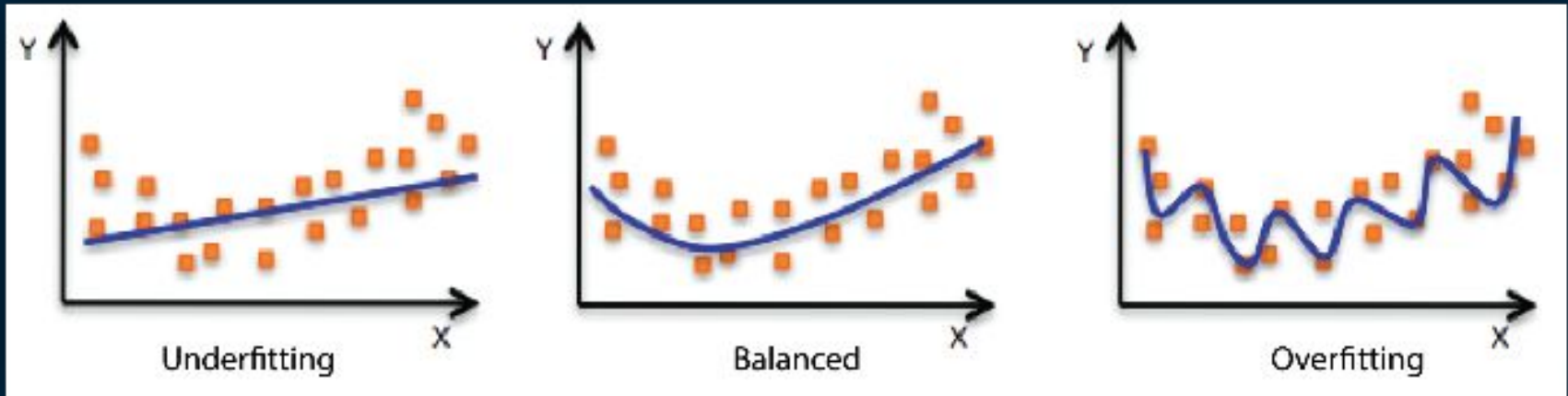
Perfect!

- Goal: Find a model between underfitting and overfitting.
- Difficult in practice to achieve a perfect balance.

Underfitting

- Model can't model training data or generalize to new data.
- Poor performance on training data.
- Easily detectable with a good performance metric.
- Solution is to try alternate machine learning algorithms

Oh NO!





THANK YOU

CREDITS: This presentation template was created by [Slidesgo](#),
including icons by [Flaticon](#), and infographics & images by [Freepik](#)
Please keep this slide for attribution

EXTRA

How ChatGPT Works Technically

