Building Computer Vision models using TensorFlow

January 23, 2022

Nitin Tiwari

- Alumni | 2017-2021 | Information Technology
- Software Engineer at LTI
- Blog Publisher, TFUG Mumbai
- Contribute to the TensorFlow Community

Today’s Menu

- What is Machine Learning?
- Supervised Learning
- Unsupervised Learning
- Neural Networks and Hyperparameters
- What is TensorFlow?
- How does object detection work?
- Training a custom object detection model
- Visualize the model
- Deploy the model over a mobile application
- Some useful resources

Key Takeaways:

- Preprocessing of data.
- Train a custom object detection model.
- Inspect an ML model to get detailed insights.
- Android application development.
- Deploy ML models on mobile devices.
- Keep the object detection mobile application with you.

What is Machine Learning?

"The ability of computers to learn without being **explicitly programmed**."

Explicitly programmed? What does that mean?

Example:
Data = 2, 3
Rule = Addition
Answer = 5

So, what happens in ML?

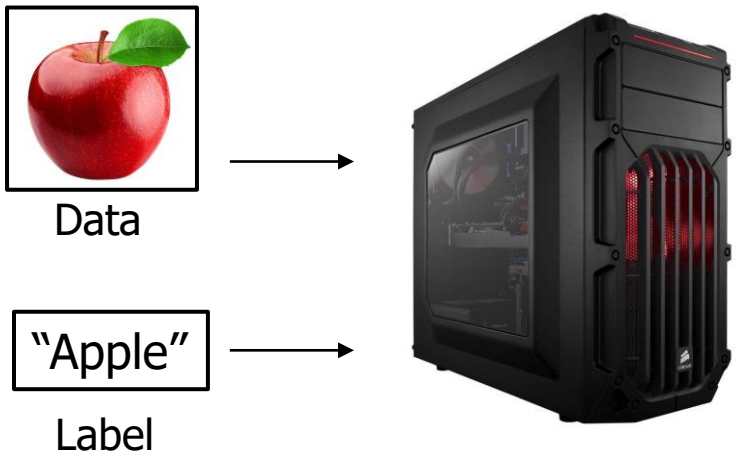
Example:
Data = 2, 3
Answer = 5
Rule = Addition

Supervised Learning

We 'supervise' the machine to learn by feeding it with both, the input data as well as correct output (label).

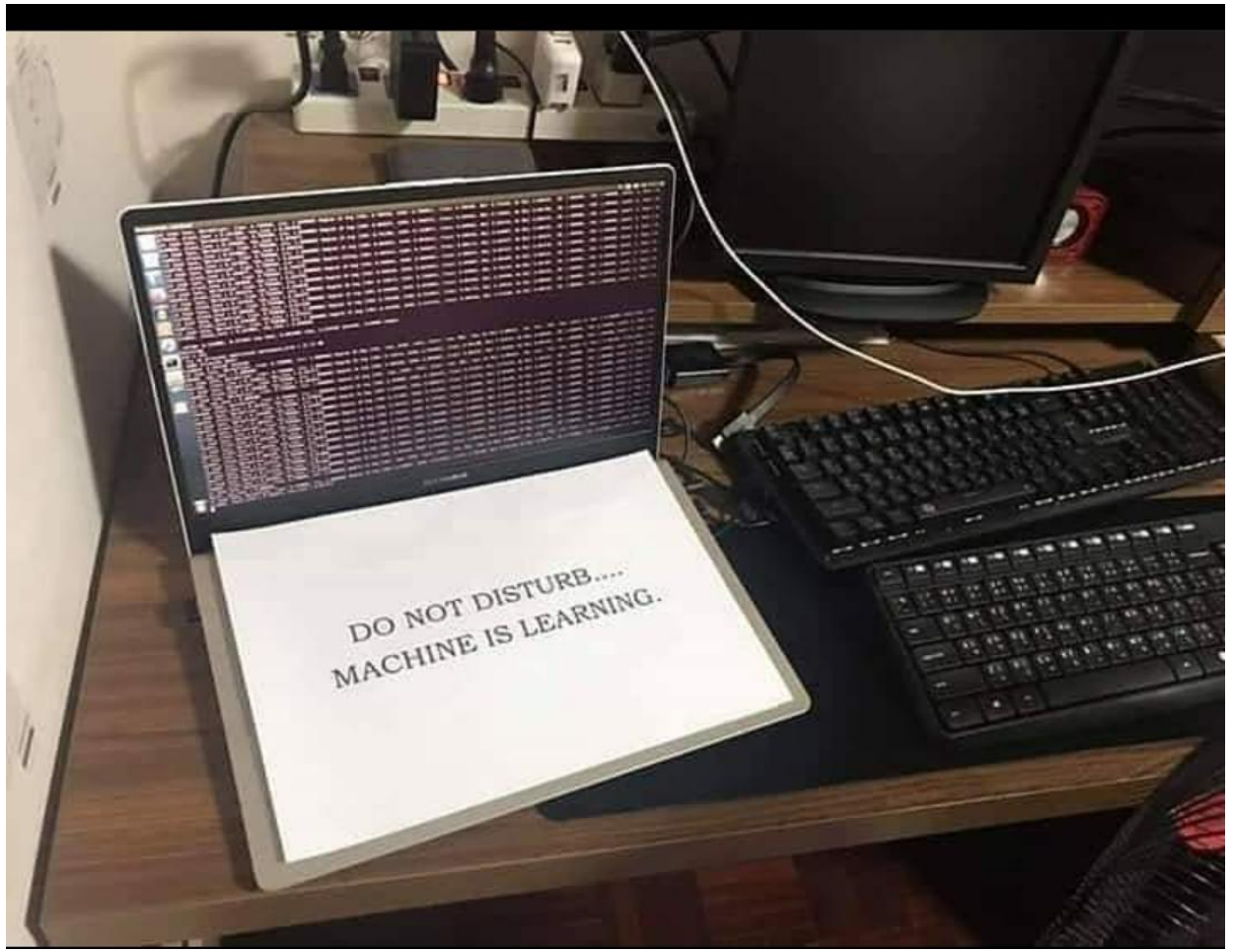
Once the machine has learned, we then feed it the data it hasn't seen previously and expect it to output the correct label.

Example:



Training a model

Testing a model


DO NOT DISTURB....
MACHINE IS LEARNING.

Task 1:

Can you list down some real-life applications of ML you use in your daily life?

Types of Data

Data can mainly be classified into:

- Structured Data
- Unstructured Data

	A	B	C	D
1	First Name	Last Name	Age	Salary
2	Jon	Smith	36	26500
3	Helen	Mirren	22	21000
4	David	Cameron	29	39000
5	Brad	Pitt	52	45000
6	Anna	Starolsky	41	22500
7	Peter	Piper	20	31500
8	David	Duck	19	15700
9	Julie	Walters	33	19000

Structured Data – Data in the form of CSV, Excel sheets, database, etc.

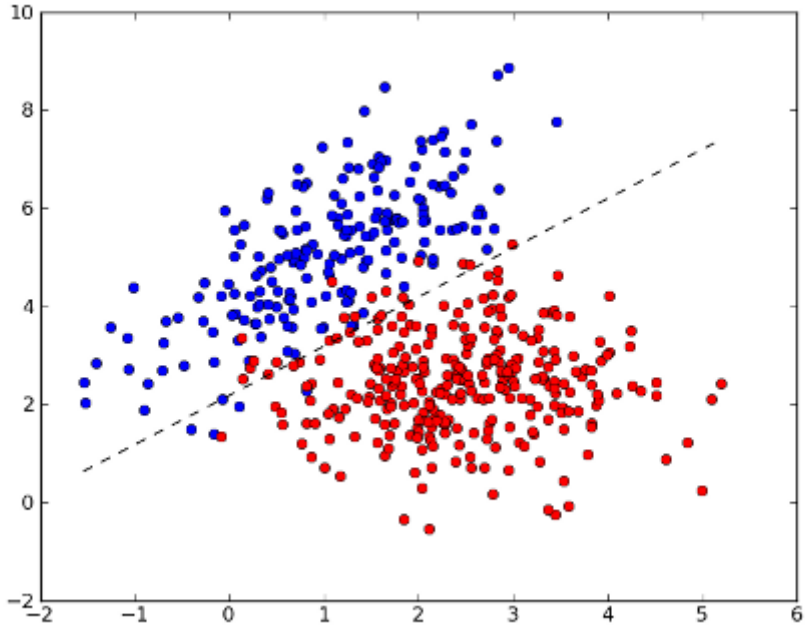
Unstructured Data – Data in the form of images, videos, emails, audio, social media posts, etc.

Supervised Learning in Structured Data

Supervised Learning in structured data mainly deals with:

- Regression
- Classification

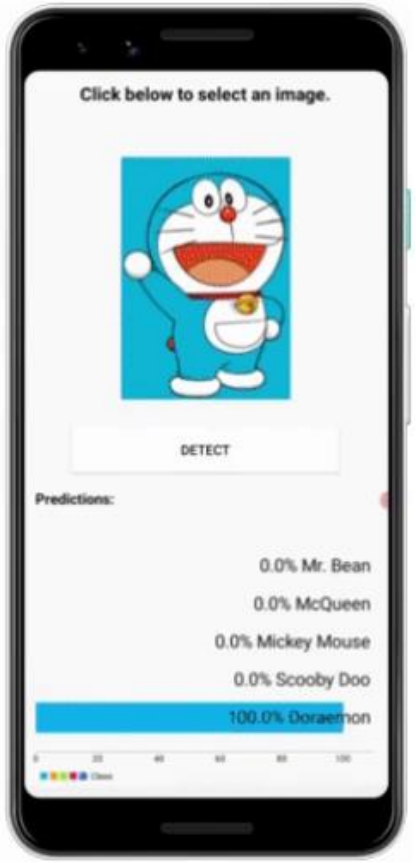
Regression – Predict a value



Classification – Categorize into two or more classes

Supervised Learning in Unstructured Data - Computer Vision

Computer Vision deals with data that is in the form of images or videos.

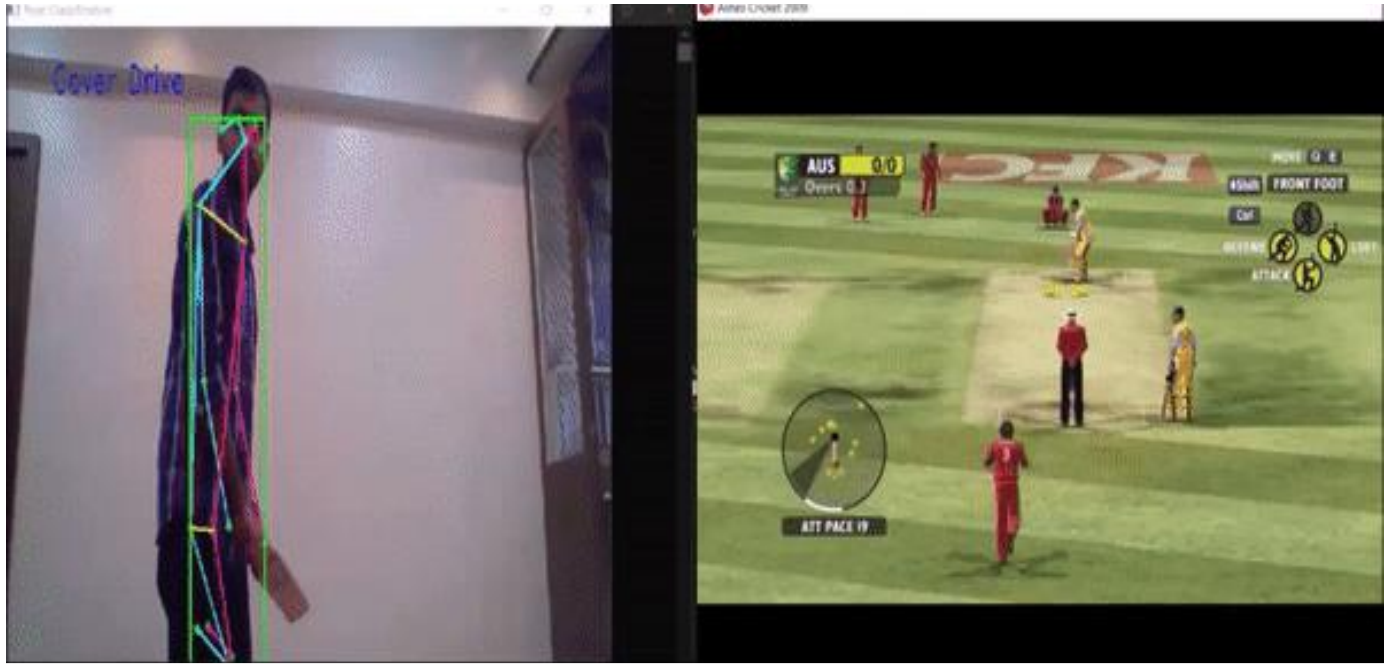


- **Image Classification**
- **Identify the content in an image**
- **Output the most relevant class**

Supervised Learning in Unstructured Data - Computer Vision

- Object Detection
- Identify the object in an image
- Find the object's region of interest
- Output object class with bounding box

Supervised Learning in Unstructured Data - Computer Vision



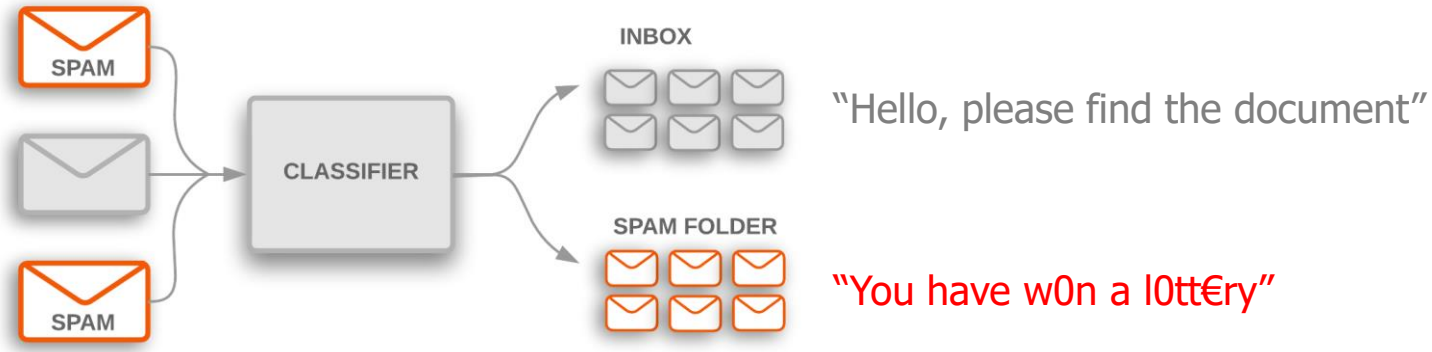
- Pose Classification
- Detect key points of human body
- Classify the pose

Supervised Learning in Unstructured Data - NLP

- NLP stands for Natural Language Processing
- The data we deal with in NLP can either be **textual** or **audio.**

Some applications of NLP are:

- Text-to-speech recognition ('Read aloud' feature in Microsoft Edge)
- Speech-to-text recognition (Digital voice assistants like Google Home, Siri, Alexa, etc.)
- Email spam classifier (Gmail automatically classifies spam)



- Hate speech recognition (tweets on Twitter that are potentially violent)

Unsupervised Learning

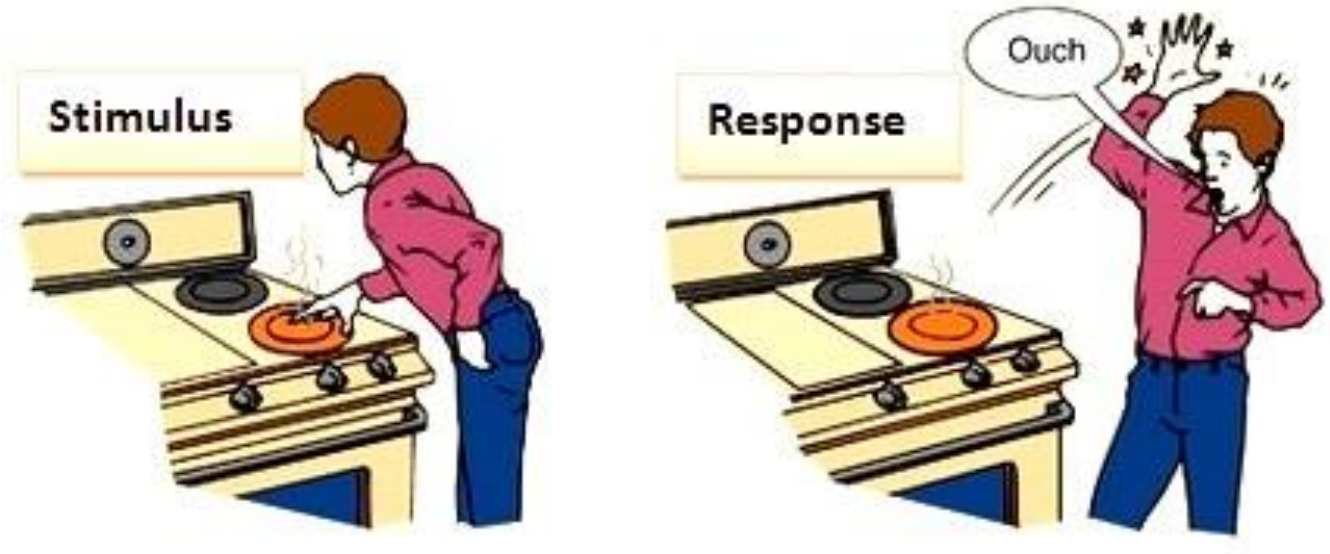
- In unsupervised learning, as the name suggests, we let the computer train the model by learning 'on its own'.
- We just feed the computer with the data and do not provide any label.

Example:

The machine 'learned' on its own and clustered the data into two categories based on the colour of the fruit.

Neural Networks

What happens when you touch a hot object?

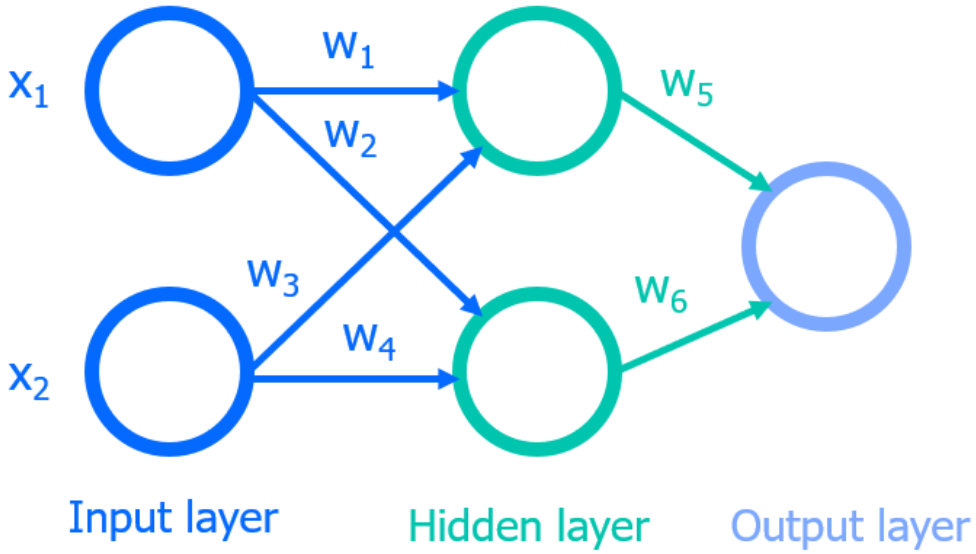


- The brain sends a signal through the neurons and activates the tissue cells of your hand so that you take your hand off the hot object immediately.
- **Input:** Signal from the brain carried by neurons.
- **Activation function:** Tell the tissue that the object is hot.
- **Output:** You take your hand off the hot object.

Neural Networks

In ML, we have a similar network of neurons that carry the data from one neuron to another.

This network is called **Artificial Neural Network** or simply, **Neural Network**.



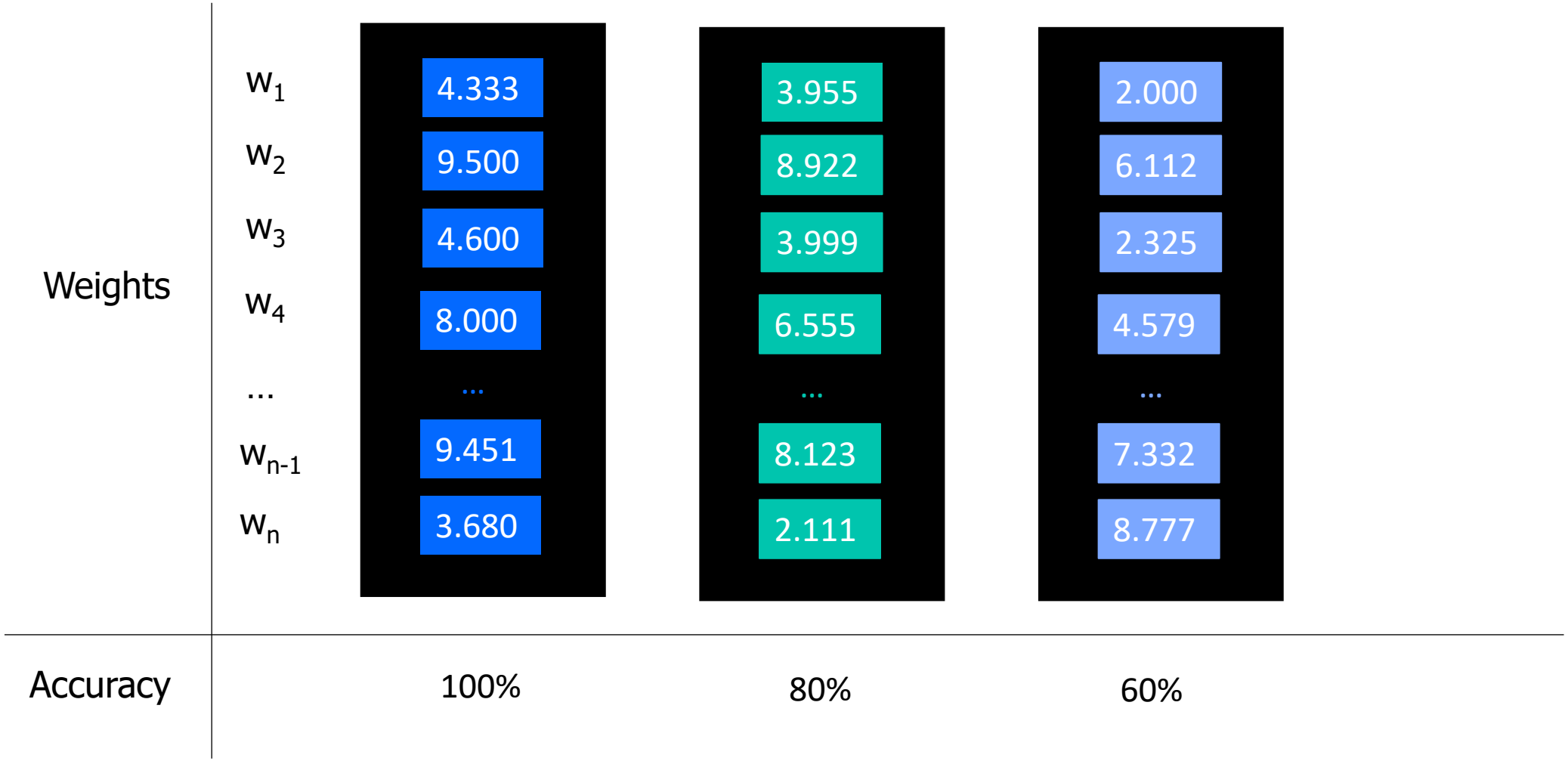
Task 2:

Find the value of ‘w’ for each of the following linear equations:

- 3w + 12 = 25 ➔ w = 4.333
- 2w - 19 = 0 ➔ w = 9.500
- 10w - 16 = 30 ➔ w = 4.600
- 5w - 17 = 23 ➔ w = 8.000

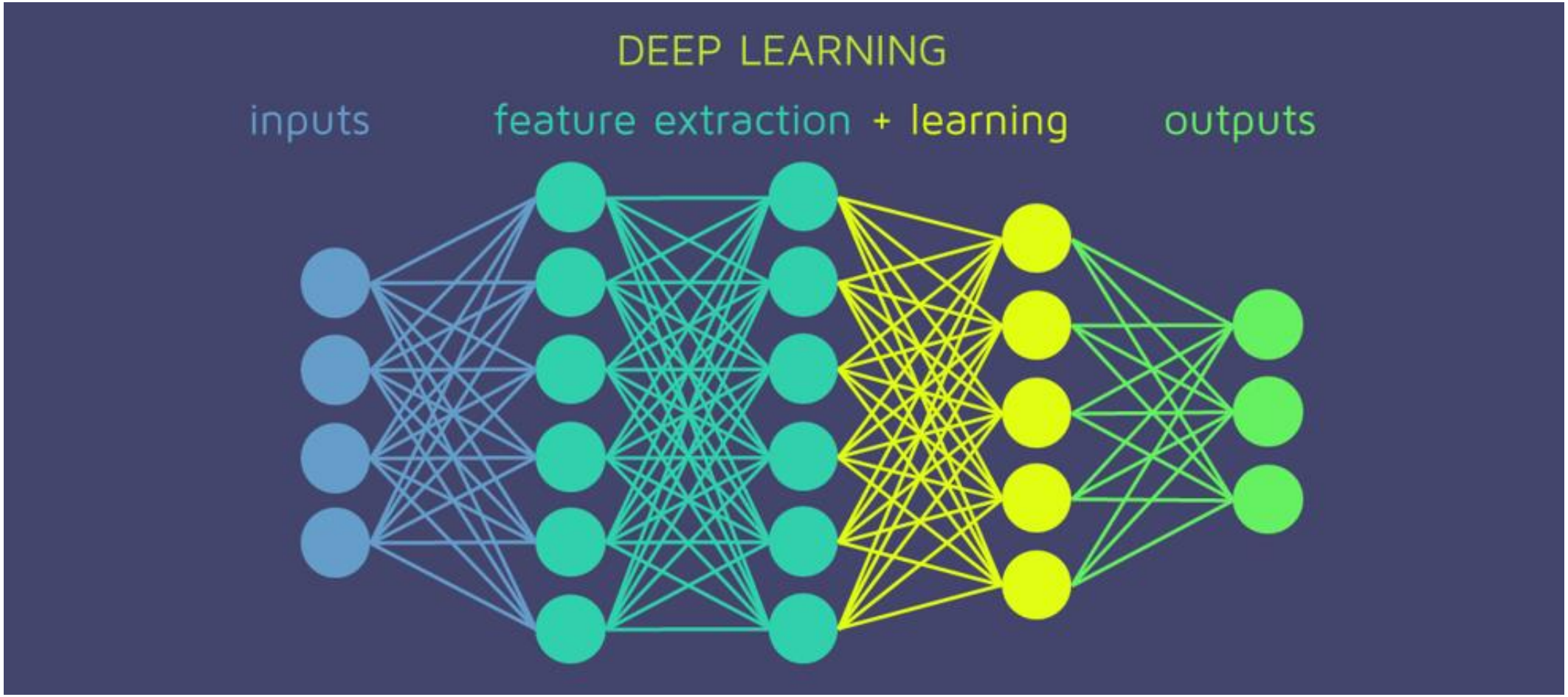
ML Model

The weights of the mathematical equations in a neural network when bundled together make an ML model.



Deep Neural Networks

Deep Neural Networks form the foundation of Deep Learning.



Task 3:

List down some different types of neural networks. Which of these is most commonly used for Computer Vision applications?

- Perceptron
- LSTM
- Recurrent Neural Network (RNN)
- Convolutional Neural Network (CNN)

Hyperparameters

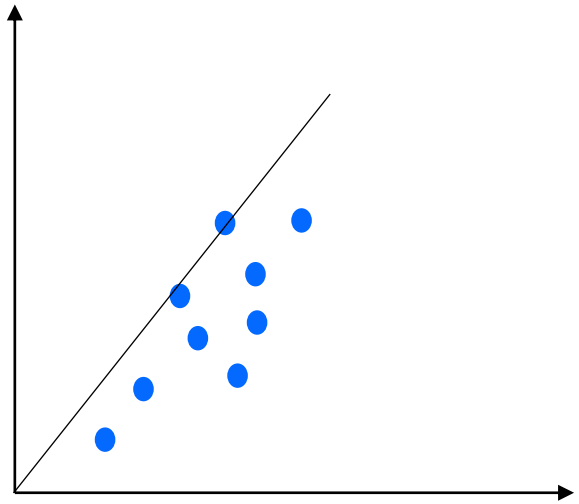
The parameters that are used to control the learning process are called hyperparameters. This process is also known as **'hyperparameter tuning'**.

- **No. of layers:** The number of layers in your neural network.
- **Activation function:** The mathematical operation that gets computed.
- **Learning rate:** The step size, or the rate, at which the algorithm learns.
- **Epochs:** The no. of times the algorithm will work thru the entire training dataset.
- **Batch size:** The no. of training samples passed to the algorithm during training for every pass.
- **Optimizer:** Helps in improving the performance of a model by adjusting the weights, learning rate, etc.
- **Train size:** The percentage of training images in your dataset.
- **Test size:** The percentage of test images in your dataset.

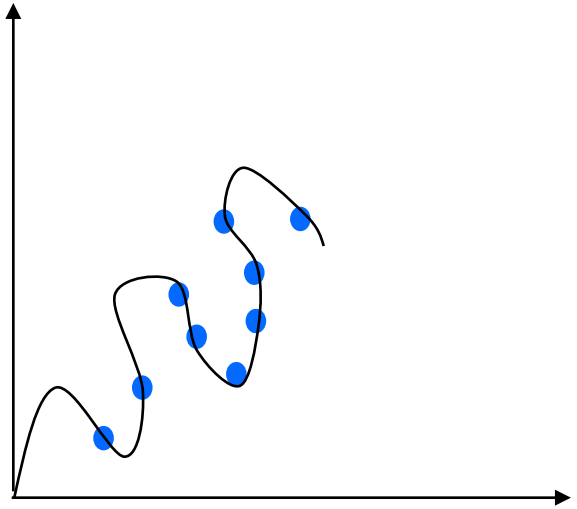
Performance Metrics and Intuitions

To evaluate how well your model is performing against the test data, the following are some metrics to be considered.

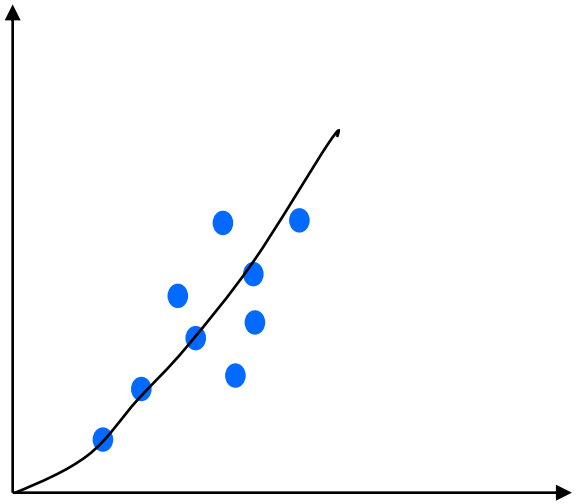
- **Accuracy:** No. of correct detections/total no. of detections.
- **Loss:** Penalty for a wrong prediction. The lower the loss, the better is the model accuracy.
- **Bias:** The learning algorithm misses out on important features in the data during training.
- **Variance:** The learning algorithm considers unwanted noise in the data which may decrease the accuracy.
- **Overfitting:** The model performs exceptionally well on training data but performs poorly on test data.
- **Underfitting:** The model doesn’t perform well even on the training data.



Underfitting



Overfitting



Optimal Fit

- High bias can cause a model to underfit.
- High variance can cause a model to overfit.
- 100% accuracy on training data is an alarm that your model is overfitting.

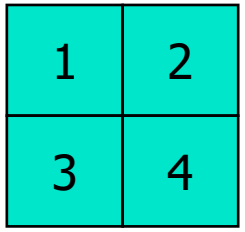
TensorFlow

- Free, open-source Machine Learning library.
- Originally developed by the Google Brain team.
- Languages supported: Python, C++, Java, JavaScript.
- Official GitHub repository: www.github.com/tensorflow
- Official website: www.tensorflow.org

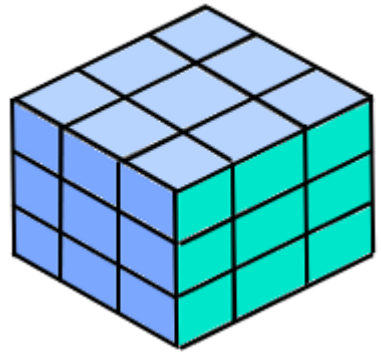
Scalar: 0D data

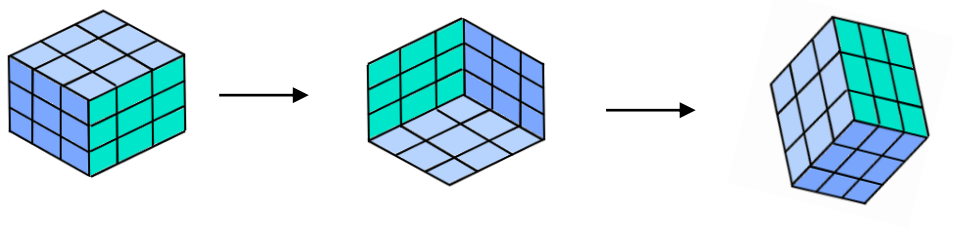
Vector: 1D data

Matrix: 2D data



Tensor: N-D data

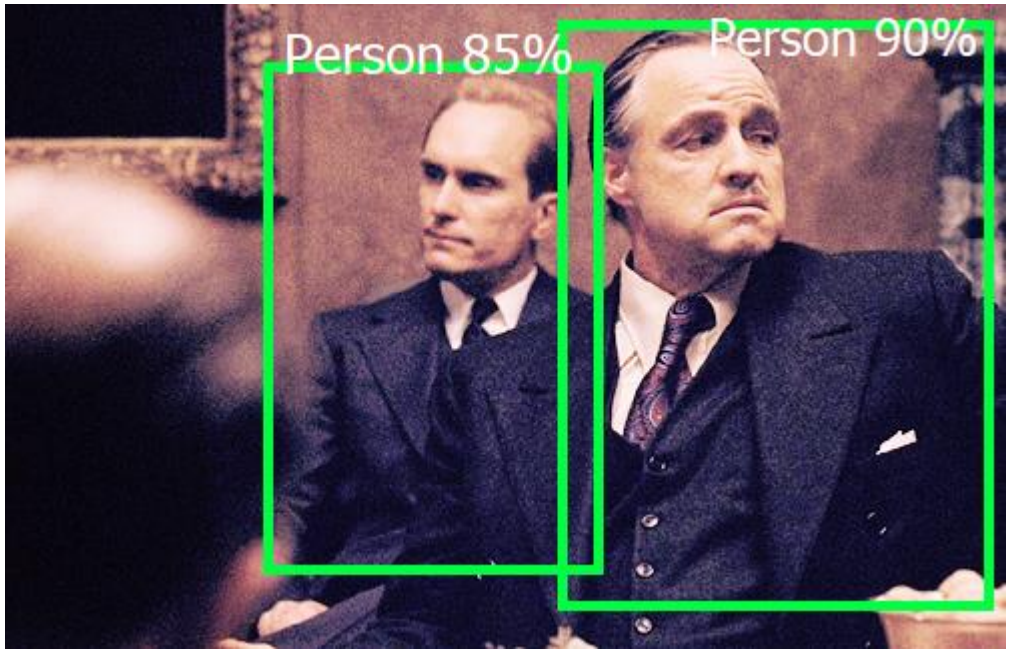




The flow of tensor data in a neural network.
Hence, the name – 'TensorFlow'.

Task 4:

Find the difference between the two pictures.

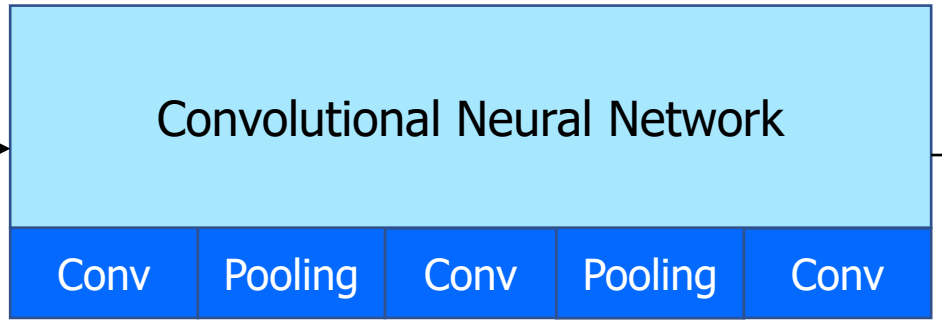


- Label
- Bounding Box
- Confidence
- No. of detections

How does object detection work?

Object Detection = Image Classification + Localization

Original Image (input)



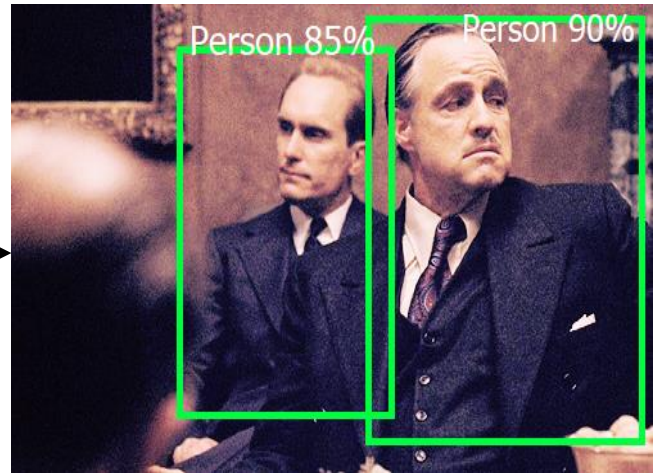


Image with detection (output)

- **Convolutional layer:** Detects the important features in the input image.
- **Pooling layer:** Reduces the size of the image while retaining the features.
- There are multiple convolutional and pooling layers in a CNN.
- Input tensor: [n x n x 3] image pixels
- Output tensor: [(labels), (bounding box coordinates), (confidence), (no. of detections)]

Lab: Training a custom object detection model

Step 1: Create a dataset of images.

Scrape images from the web or any other data source. The images should be in JPG/JPEG format only.

Step 2: Label the images.

Annotate/label the images using the LabelImg tool.

Step 3: Model Training.

Open Google Colab and let’s code along.

Step 4: Download the model.

Once the model is trained, download the TF Lite model.

Lab: Visualize the model

We can inspect the model by visualizing how the neural network looks and how every layer is connected to the other.

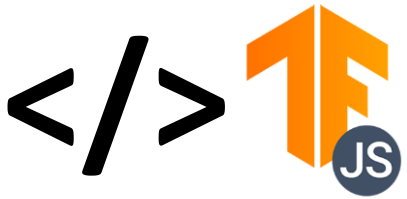
TF Lite model visualized on Netron

Lab: Deploy the model

Any ML workflow is incomplete without the deployment of the model.

TensorFlow models can be of the following formats depending upon where they need to be deployed.

Standalone Desktop Application
(TensorFlow)



Web Browser
(TensorFlow.js)

Mobile Devices
(TensorFlow Lite)

Some useful resources

Courses:

- Machine Learning | Stanford University
- Deep Learning Specialization by DeepLearning.AI
- TensorFlow Developer Professional Certificate by DeepLearning.AI
- TensorFlow: Data and Deployment

Websites/Blogs:

- Official TensorFlow website | **www.tensorflow.org**
- Official GitHub repository | **www.github.com/tensorflow**
- Medium blog | **www.medium.com/tensorflow**

www.linkedin.com/in/tiwari-nitin

tiwarinitin1999@gmail.com

www.tiwarinitin1999.medium.com

www.github.com/NSTiwari