Introduction to Machine Learning
Part II

Mehran Behjati

UNIVERSITI KEBANGSAAN MALAYSIA
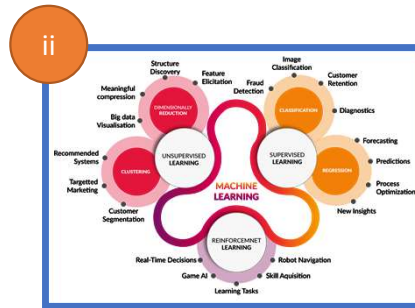*National University of Malaysia*

Layout

- Recap

- Neural Network

- Tools:
  - Colab
  - pandas
  - Numpy
  - Scikit-learn
  - TensorFlow
  - Edge Impulse

Recap ...
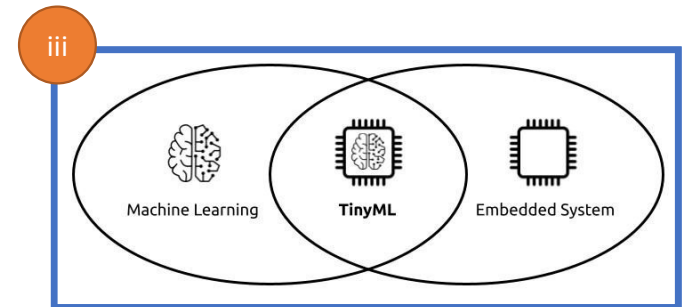
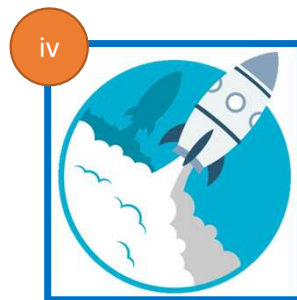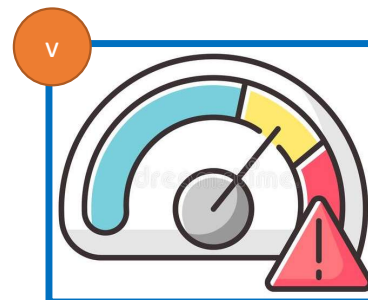**What is AI, ML, & DL**

ML categories

ML vs TinyML

ML enablers

ML Limitations

... Recap

Get Data

1

Importance of data
- Training set
- Validation set
- Test set

Issues:
- Bias
- Unbalanced
- Small

Clean, Prepare & Manipulate Data

2

Data cleaning
- Removing unwanted
- Managing outliers
- Handling missing data

Feature selection
Feature extraction

Train Model

3

Split dataset
Model selection
- Regression
- Classification
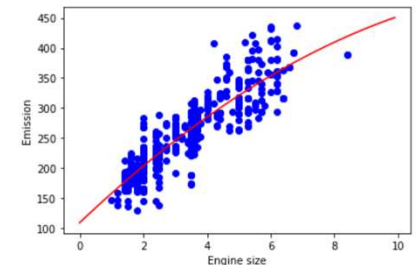- Clustering

Test Data

4

Overfitting
Underfitting
Evaluation matrices
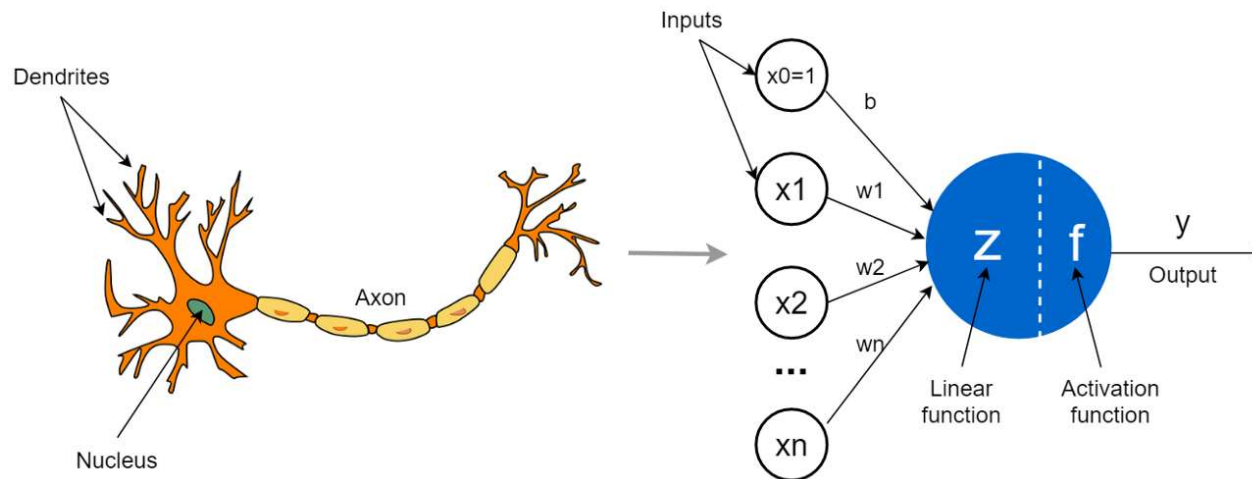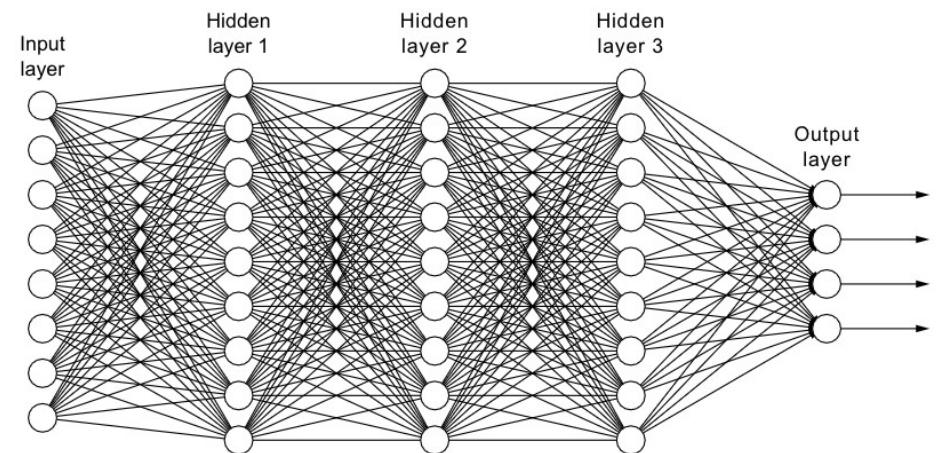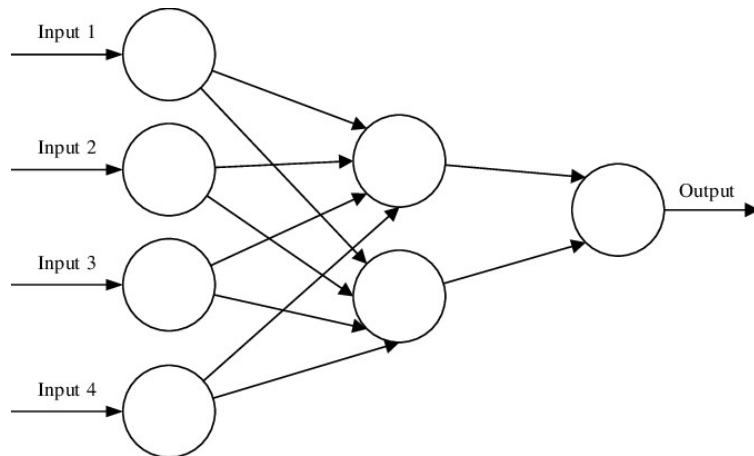Confusion matrix

jupyter CO2Emission

Neural Network

- What is Artificial Neural?
- It was inspired by the understanding around how biological neurons work and operate in the human brain.
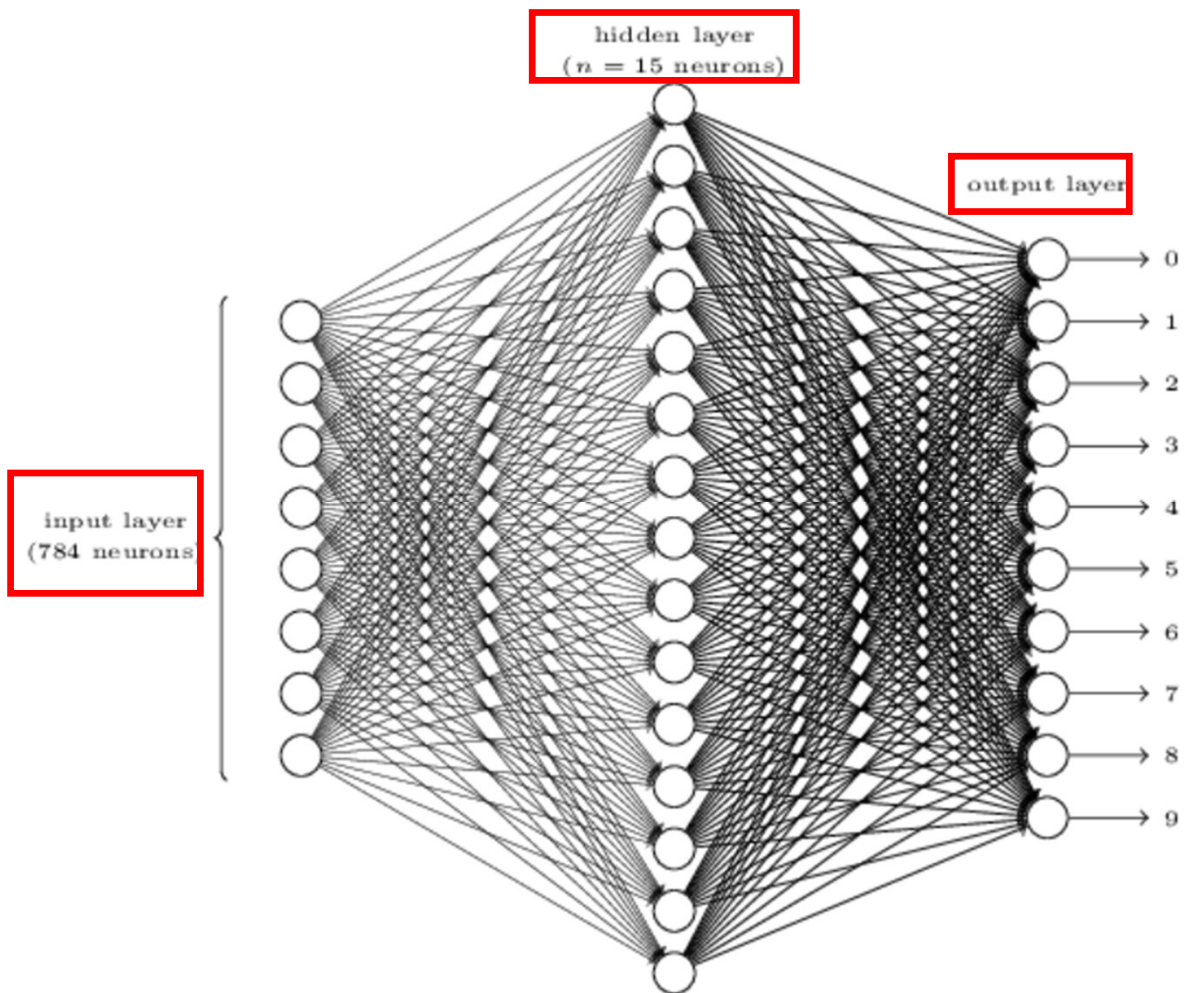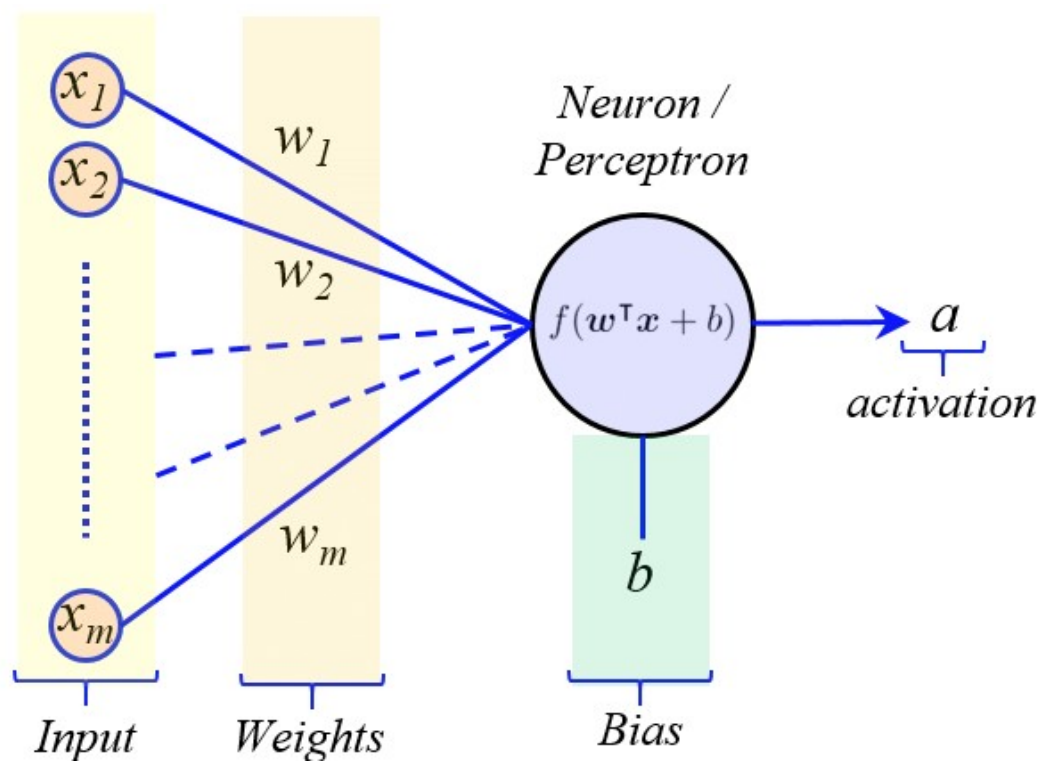
Deep Neural Network

- **Neural Network**

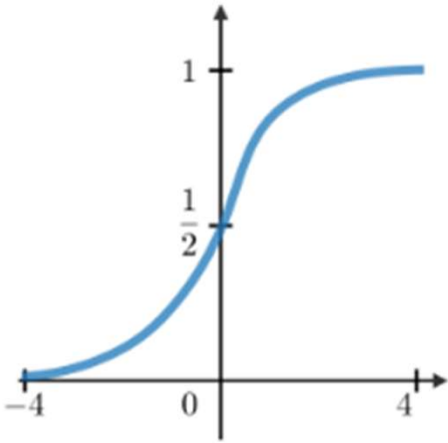- **Deep/dense NN:** NN models with hidden layers, where every layer densely connected with the previous and the next layer
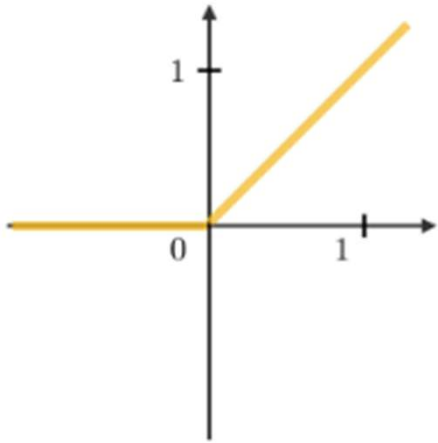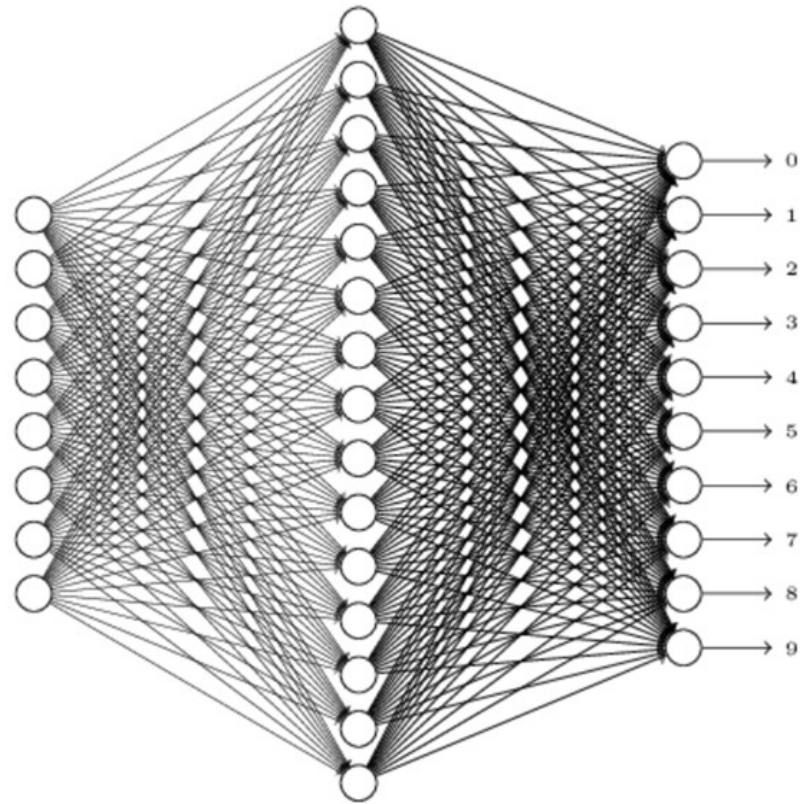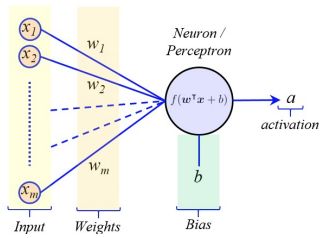
Handwritten digit recognition

28 x 28
784 pixels

```
0  0  0  0  0  0  0  0  0  0   0   0   0   0   0   0   0   0   0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0   0   0   0   0   0   0   0   0   0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0   0   0   0   0   0   0   0   0   0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0   0   0   0   0   0   0   0   0   0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0   0   0   0   0   0   0   0   0   0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0   4  62 146 182 254 254 181 176 139 15  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0  34 186 253 217 208 136 136 136 166 232 99  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  61 242 208 111  3   0   0   0   0   0  18 32 107 43  0  0  0  0  0  0
0  0  0  0  0  0  0  0 156 242 23   0   0   0   0   0   0   0  13 191 181 6  0  0  0  0  0  0
0  0  0  0  0  0  0  0 121 255 98   3   0   0   0   0   0   8 194 225 12  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0 169 253 120  3   0   0   0   0 128 247 51  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0  3 111 244 169 19   0  14 131 249 117  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0 59 241 235 72 142 229 66  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0 25 218 254 231 36  0  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0  0 133 253 221 33  0  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0 19 237 111 196 217 19  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0 174 138  0  23 193 204 18  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0 96 224   0   0   0 25 218 169  3  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0 215 138   0   0   0 86 253 99  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0 215 97    0   0   0  3 162 214 11  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0 215 97    0   0   0  0 118 253 68  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0 185 157   0   0   0 40 254 98  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0 50 244 61   0   0   0 112 244 58  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0  0 174 251 142 59 83 167 244 111  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0  0  6 133 253 253 253 169 61  3  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0   0   0   0   0   0   0   0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0   0   0   0   0   0   0   0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0   0   0  0   0   0   0   0   0   0   0  0  0  0  0  0  0  0  0  0  0
```

pixel 1 →
pixel 2 →
pixel 3 →
pixel 4 →
pixel 5 →
pixel 6 →
pixel 7 →
pixel 8 →
pixel 9 →
pixel 10 →
pixel 11 →
pixel 12 →
pixel 13 →
pixel 14 →
pixel 15 →
pixel 16 →
pixel 17 →
pixel 18 →
pixel 19 →
⋮
pixel 784 →

Input layer

hidden layer
($n = 15$ neurons)

output layer

input layer
(784 neurons)

0
1
2
3
4
5
6
7
8
9

$$f(\mathbf{w}^\mathsf{T}\mathbf{x} + b)$$

x_1

x_2

x_m

w_1

w_2

w_m

Neuron / Perceptron

a

activation

b

Input   Weights   Bias

Activation Function

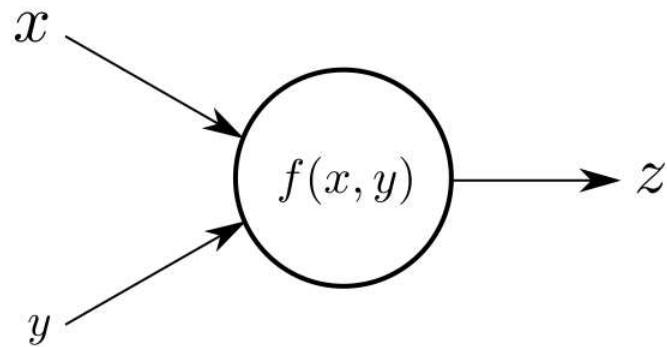| Sigmoid	RELU
$g(z) = \dfrac{1}{1 + e^{-z}}$	$g(z) = \max(0, z)$

Feedforward

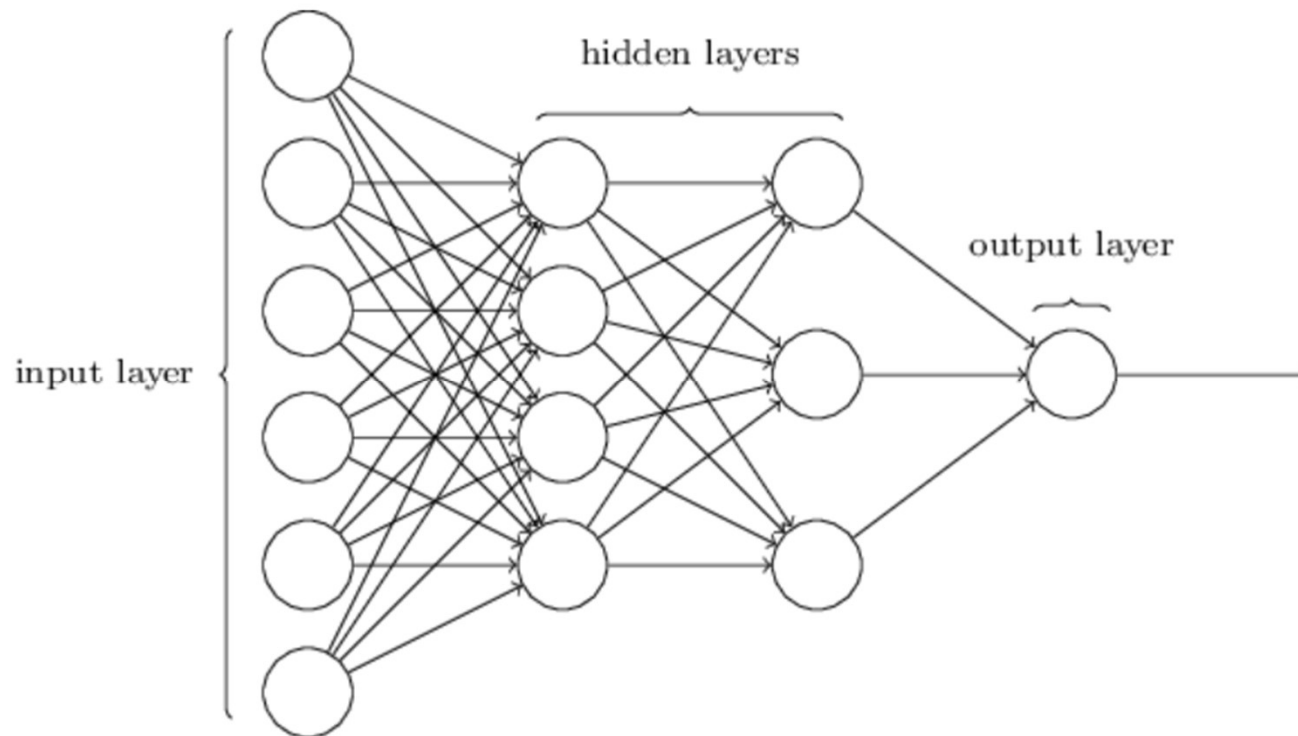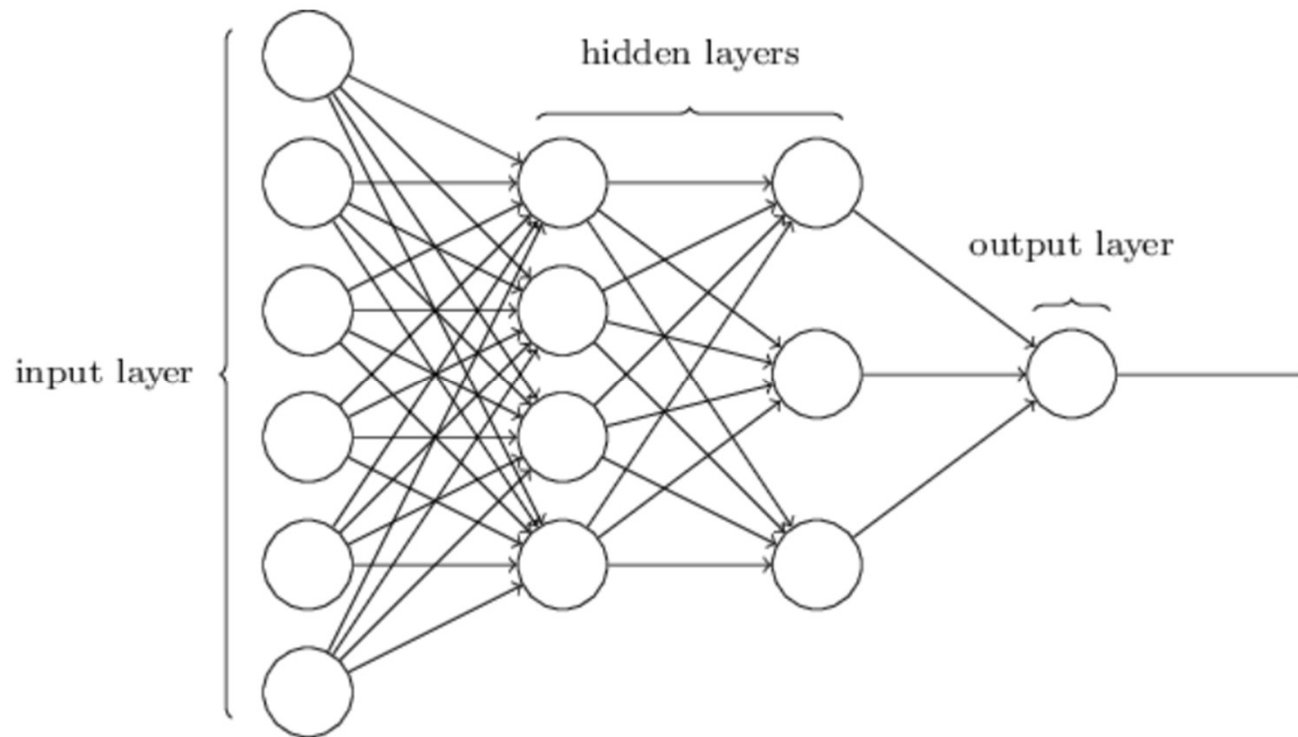# Cost/Loss/Error Function
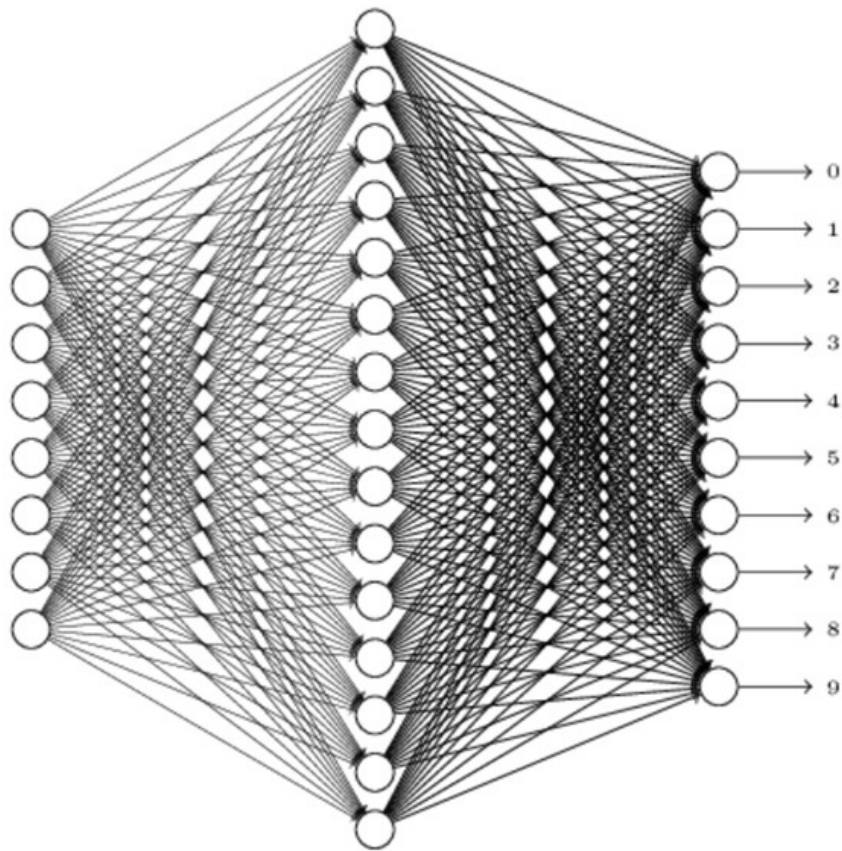
Backpropagation

Forwardpass

Backwardpass

$$\frac{dL}{dx} = \frac{dL}{dz}\frac{dz}{dx}$$

$$\frac{dL}{dy} = \frac{dL}{dz}\frac{dz}{dy}$$

$$df$$

$$\frac{dL}{dz}$$
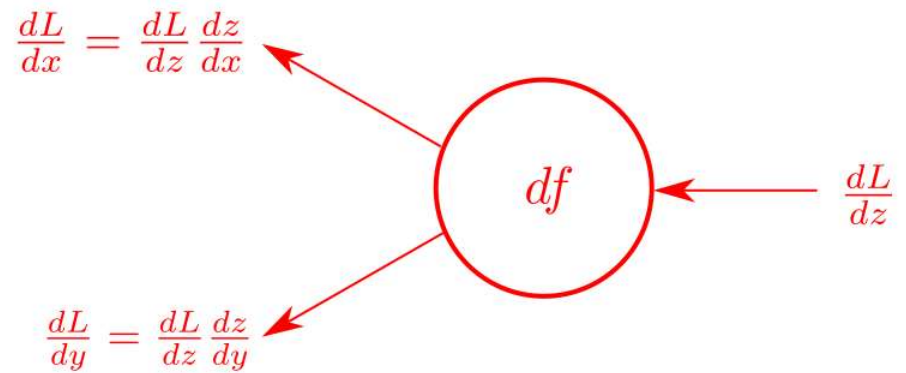
Gradient decent

Iteration
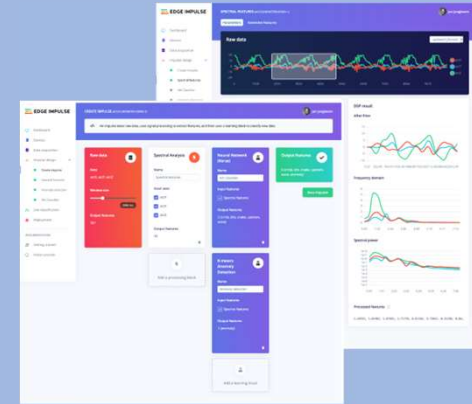

input layer
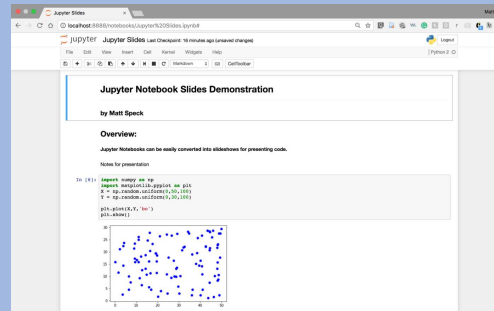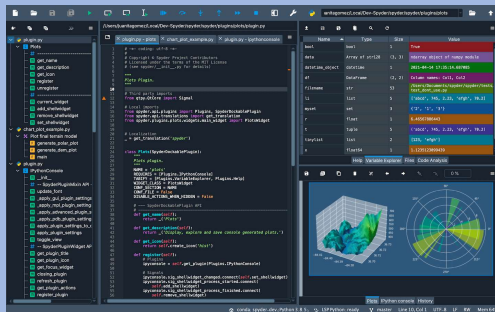
hidden layers
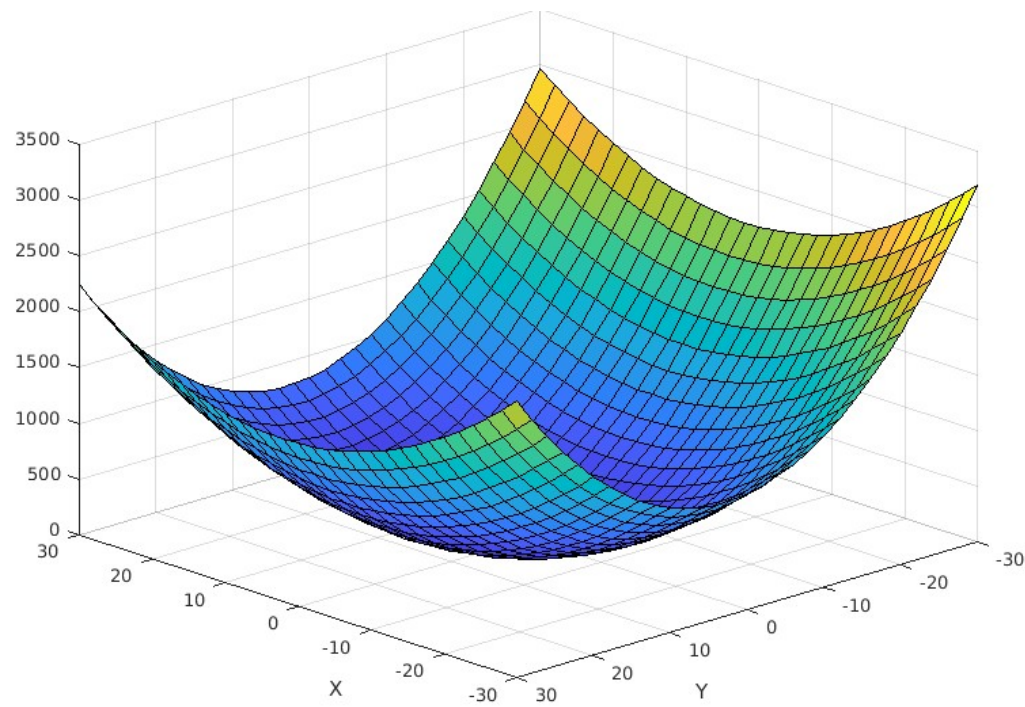
output layer

Epoch

Some of Essential Terminologies

- Neuron/node
- Activation
- Weight
- Bias
- Activation function
- Input, hidden, and output layer
- Forward propagation / Feedforward
- Backpropagation
- Cost/Loss function
- Iteration
- Epoch

How to create a ML project

colab

- An executable document lets you write and run your code on browser

- Already pre-installed

- Connects your notebook to a cloud-based runtime
  - We can use the power of cloud to train our models efficiently

- Easy to share code within Google Drive or GitHub

- Edit collaboratively

- Hardware acceleration, GPUs & TPUs

- The volume of data is rapidly growing
- The importance of processing this data is much more than decades ago
- Excel is good but we need more flexible, powerful and advanced tools
- It is an open-source library that has been developed in Python
- Almost everyone who works on the field of data science needs to know pandas
  - read data from csv
  - work with series
  - arithmetic and statistical operations
  - sort
  - work with data frames (combinations of series)
  - choosing rows and columns
  - filtering
  - optimizing memory

NumPy

- A famous library among the data scientists for scientific computing in Python.

- Something similar to MATLAB, for array and matrix operations

- Working on numbers, array, and matrices
  - Array creation
  - Indexing, slicing, and iterating
  - Shape manipulation
  - Arrays stacking
  - Other operations, including mathematical, logical, sorting, selecting, discrete Fourier transforms, basic linear algebra, basic statistical operations, random simulation and much more.

- An open source ML library that supports supervised and unsupervised learning

- Provides various tools for model fitting, data preprocessing, model selection, model evaluation, and many other utilities

- Provides dozens of built-in ML algorithms and models

- An open source end-to-end platform for ML

- Provides a comprehensive ecosystem of tools to built ML applications

- Developed and supported by Google to process and analyze data

- Supports wide range of ML algorithms and models

- Provides APIs for different languages, Python, C++, Javascript, …

- TensorFlow is originally designed for big computing systems
- Is it suitable for TinyML implementation?
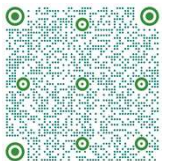
Installations

- You can install Python and other tools separately on your device
- Alternatively you can install Anaconda

- A great tool to create ML projects
- Makes model training much easier
- Gives a graphical interface & representation