

# **Assignment No. 1**

- **Title:** Linear regression by using Deep Neural network
  - **Date of Completion:**
  - **Objective:** Study and understand neural network by using Linear Regression for predicting house prices.
  - **Problem Statement:** Implement Boston housing price prediction problem by Linear regression using Deep Neural network. Use Boston House price prediction dataset.
  - **Theory:**

## Linear Regression

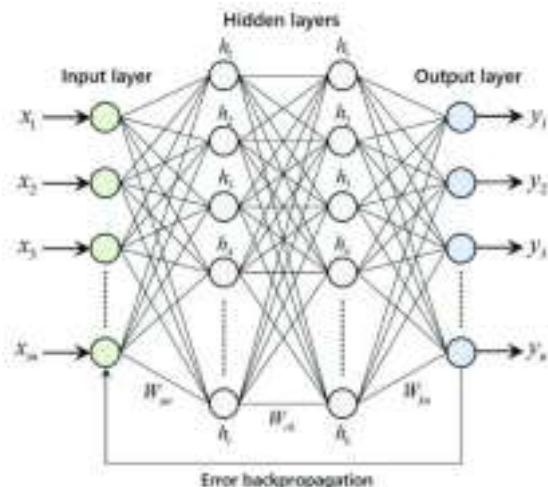
Linear Regression is a supervised learning technique that involves learning the relationship between the features and the target. The target values are continuous, which means that the values can take any values between an interval. Use-cases of regression include stock market price prediction, house price prediction, sales prediction, and etc.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

↓                      ↓                      ↓                      ↓  
 Dependent Variable   Population Y Intercept   Population Slope Coefficient   Independent Variable   Random Error term  
Y<sub>i</sub> =  $\beta_0 + \beta_1 X_i$  +  $\varepsilon_i$   
Linear component   Random Error component

## Neural network

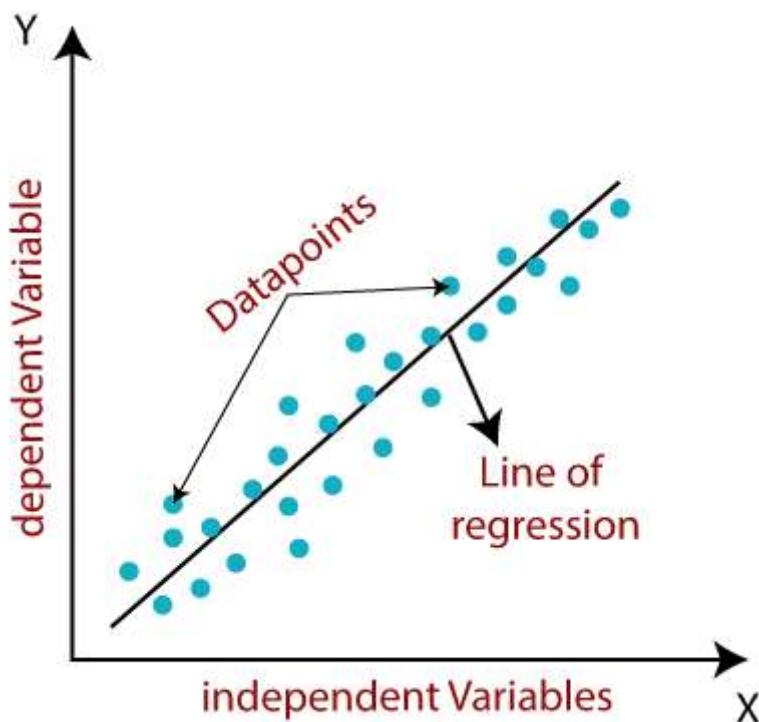
Neural networks are formed when multiple neural layers combine with each other to give out a network, or we can say that there are some layers whose outputs are inputs for other layers.



The purpose of using Artificial Neural Networks for Regression over Linear Regression is that the linear regression can only learn the linear relationship between the features and target and therefore cannot learn the complex non-linear relationship. In order to learn the complex non-linear relationship between the features and target, we are in need of other techniques. One of those techniques is to use Artificial Neural Networks. Artificial Neural Networks have the ability to learn the complex relationship between the features and target due to the presence of activation function in each layer. Artificial Neural Networks are one of the deep learning algorithms that simulate the workings of neurons in the human brain.

Linear regression algorithm shows a linear relationship between a dependent ( $y$ ) and one or more independent ( $x$ ) variables, hence called as linear regression. Since linear regression shows the linear relationship, which means it finds how the value of the dependent variable is changing according to the value of the independent variable.

The linear regression model provides a sloped straight line representing the relationship between the variables. Consider the below image:



Mathematically, we can represent a linear regression as:

$$y = a_0 + a_1 x + \varepsilon$$

Here,

$Y$  = Dependent Variable (Target Variable)

$X$  = Independent Variable (predictor Variable)

$a_0$  = intercept of the line (Gives an additional degree of freedom)

$a_1$  = Linear regression coefficient (scale factor to each input value).

$\varepsilon$  = random error

The values for  $x$  and  $y$  variables are training datasets for Linear Regression model representation.

## Types of Linear Regression

Linear regression can be further divided into two types of the algorithm:

- Simple Linear Regression:

If a single independent variable is used to predict the value of a numerical dependent variable, then such a Linear Regression algorithm is called Simple Linear Regression.

- Multiple Linear regression:

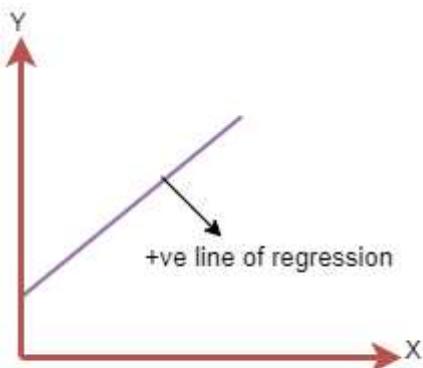
If more than one independent variable is used to predict the value of a numerical dependent variable, then such a Linear Regression algorithm is called Multiple Linear Regression.

## Linear Regression Line

A linear line showing the relationship between the dependent and independent variables is called a regression line. A regression line can show two types of relationship:

- Positive Linear Relationship:

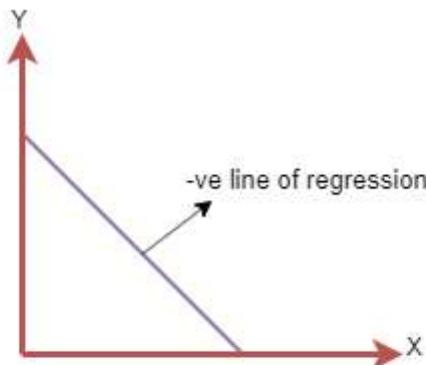
If the dependent variable increases on the Y-axis and independent variable increases on X-axis, then such a relationship is termed as a Positive linear relationship.



The line equation will be:  $Y = a_0 + a_1x$

- Negative Linear Relationship:

If the dependent variable decreases on the Y-axis and independent variable increases on the X-axis, then such a relationship is called a negative linear relationship.



The line of equation will be:  $Y = -a_0 + a_1x$

- **Dataset :** Boston Housing Dataset

- **References/ Available link :**

- 1.<https://towardsdatascience.com/linear-regression-on-boston-housing-dataset-f409b7e4a155>
- 2.<https://www.kaggle.com/code/shreayan98c/boston-house-price-prediction>

- **Questions and answer :**

- Q1: What is Linear Regression?
- Q2: How would you detect Overfitting in Linear Models?
- Q3: What are types of Linear Regression?
- Q4: How can you check if the Regression model fits the data well?
- Q5: Define Linear Regression and its structure ?
- Q6: What is the difference between Mean Absolute Error (MAE) vs Mean Squared Error (MSE)?
- Q7: What's the difference between Covariance and Correlation?
- Q8: Explain the intuition behind Gradient Descent algorithm ?
- Q9: How is the Error calculated in a Linear Regression model?
- Q10: Why use Root Mean Squared Error (RMSE) instead of Mean Absolute Error (MAE)?

- **Mcqs:**

1. The best fit line method for data in Linear Regression?
  - A) Least Square Error
  - B) Maximum Likelihood
  - C) Logarithmic Loss
  - D) Both A and B

2. Which of the following metrics can be used to evaluate a model with a continuous output variable?
- A) Precision-Recall Curve
  - B) Accuracy
  - C) Logloss
  - D) Mean-Squared-Error
3. Which statement is true about outliers in Linear regression?
- A) Linear regression model is not sensitive to outliers
  - B) Linear regression model is sensitive to outliers
  - C) Can't say
  - D) None of these

- **Conclusion :**

We have successfully solved Boston house price prediction problem by Linear Regression using deep neural networks.

## **Assignment No. 2**

---

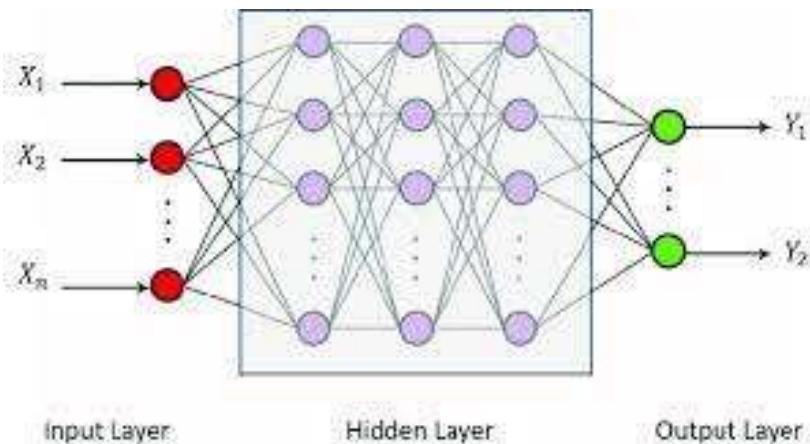
- **Title:** Classification using Deep neural network
- **Date of Completion:**
- **Objective:** To study and understand how to solve classification problem using deep neural network.
- **Problem Statement:** Binary classification using Deep Neural Networks Example:  
Classify movie reviews into positive" reviews and "negative" reviews, just based on the text content of the reviews. Use IMDB dataset.
- **Theory:**

Binary classification refers to those **classification tasks that have two class labels**. Examples include: Email spam detection (spam or not). Churn prediction (churn or not).

Binary classification is one of the most common and frequently tackled problems in the machine learning domain. In it's simplest form **the user tries to classify an entity into one of the two possible categories**. For example, give the attributes of the fruits like weight, color, peel texture, etc.

Deep learning (DL) is a subfield of machine learning based on learning multiple levels of representations by making a hierarchy of features where the higher levels are defined from the lower levels and the same lower level features can help in defining many higher level features. DL structure extends the traditional neural networks by adding more hidden layers to the network architecture between the input and output layers to model more complex and nonlinear relationships.

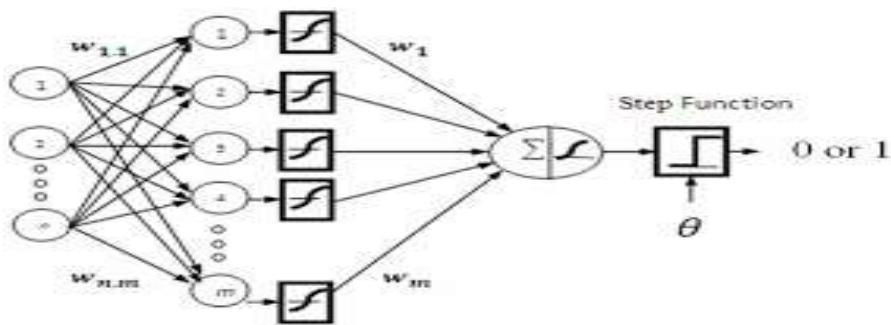
Deep Neural Network is another DL architecture that is widely used for classification or regression with success in many areas. It's a typical feedforward network which the input flows from the input layer to the output layer through number of hidden layers which are more than two layers.



The main purpose of a neural network is to receive a set of inputs, perform progressively complex calculations on them, and give output to solve real world problems like classification.

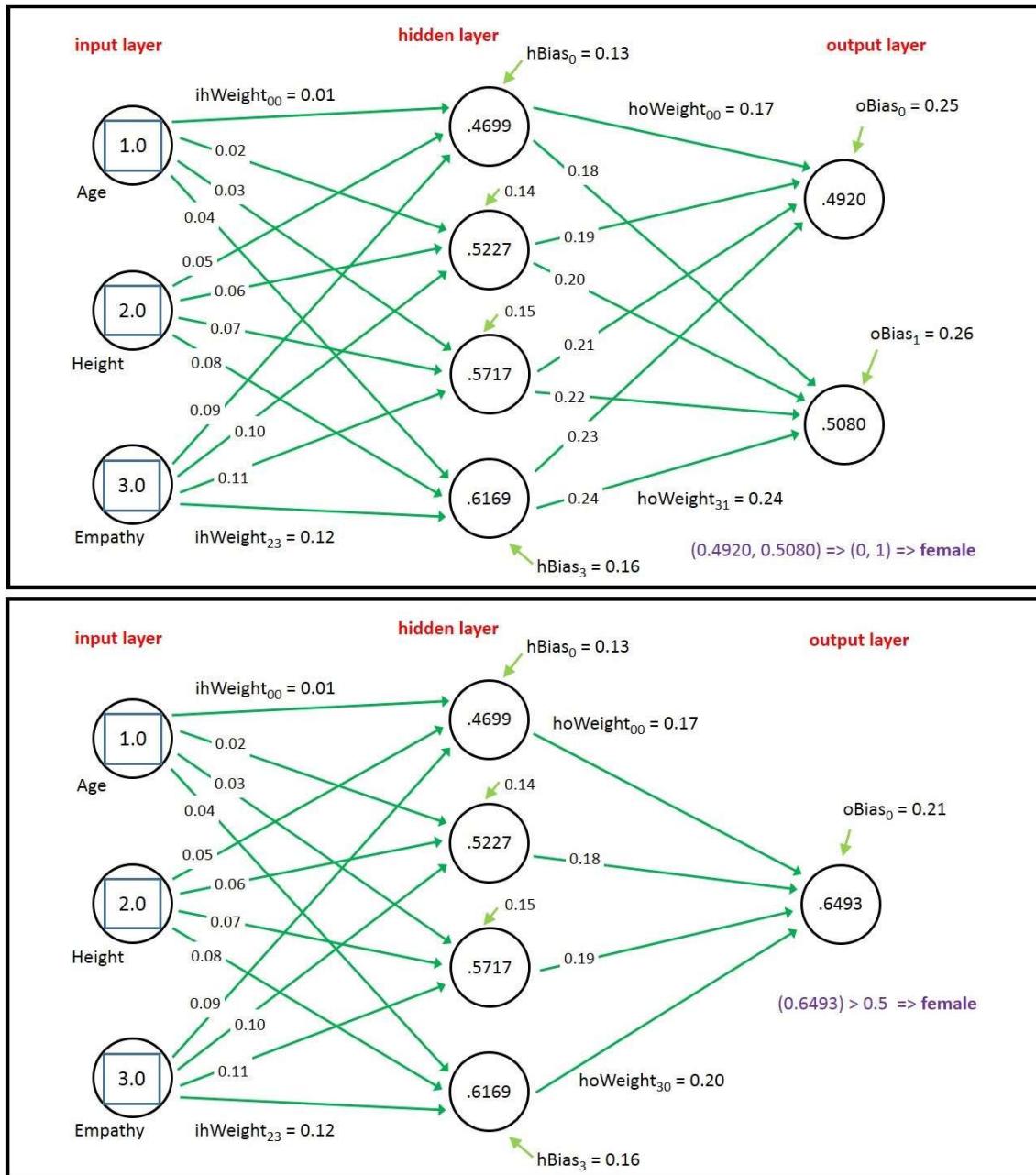
The IMDB sentiment classification dataset consists of 50,000 movie reviews from IMDB users that are labeled as either positive (1) or negative (0). The reviews are preprocessed and each one is encoded as a sequence of word indexes in the form of integers. The words within the reviews are indexed by their overall frequency within the dataset.

Building a neural network that performs binary classification involves making two simple changes:



- Add an activation function – specifically, the *sigmoid* activation function – to the output layer. Sigmoid reduces the output to a value from 0.0 to 1.0 representing a probability. For a reminder of what a sigmoid function does, see my post on binary classification.

- Change the loss function to *binary\_crossentropy*, which is purpose-built for binary classifiers. Accordingly, change *metrics* to ‘accuracy’ so accuracies computed by the loss function are captured in the history object returned by *fit*.



- **Dataset :** IMDB dataset

- **References/ Available link :**

1. <https://towardsdatascience.com/binary-classification-of-imdb-movie-reviews-648342bc70dd>

2.<https://vsokolov.org/courses/41000/notes/nlp-imdb.html>

- **Questions and answer :**

- 1.What is Deep Learning?
- 2.What is a Neural network?
- 3.What is Binary Classification?
- 4.What types of Classification Algorithms do you know?
- 5.Can you choose a classifier based on the size of the training set?
- 6.Could you convert Regression into Classification and vice versa?
- 7.How do you use a supervised Logistic Regression for Classification?
- 8.How does ROC curve and AUC value help measure how good a model is?
- 9.Name some classification metrics and when would you use each one?
- 10.What's the difference between Generative Classifiers and Discriminative Classifiers? Name some examples of each one?

- **Mcqs:**

- 1.Following are the types of supervised learning \_\_\_\_\_
  - a.regression
  - b.classification
  - c.subgroup discovery
  - d.All of above
- 2.Following are the descriptive models \_\_\_\_\_
  - a.classification
  - b.clustering
  - c.association rule
  - d.Both 1 and 2
- 3.In binary classification number of classes must be \_\_\_\_\_
  - a.equals to two
  - b.less than two
  - c.greater than two
  - d.None

- **Conclusion :**We have successfully classified movie reviews using deep neural networks.

## **Assignment No. 3**

---

- **Title:**Convolutional neural network (CNN)
- **Date of Completion:**
- **Objective:** Study and understand Convolutional Neural Network by creating a classifier.
- **Problem Statement:**Use MNIST Fashion Dataset and create a classifier to classify fashion clothing into categories.
- **Theory:**

CNNs are a class of Deep Neural Networks that can recognize and classify particular features from images and are widely used for analyzing visual images. Their applications range from image and video recognition, image classification, medical image analysis, computer vision and natural language processing.

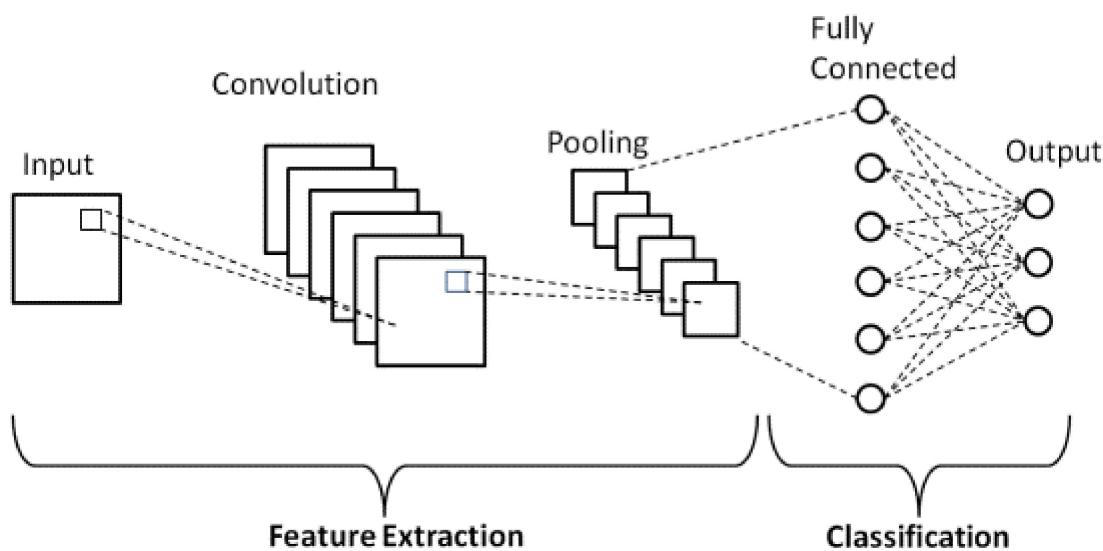
CNN has high accuracy, and because of the same, it is useful in image recognition. Image recognition has a wide range of uses in various industries such as medical image analysis, phone, security, recommendation systems, etc.

The term ‘Convolution’ in CNN denotes the mathematical function of convolution which is a special kind of linear operation wherein two functions are multiplied to produce a third function which expresses how the shape of one function is modified by the other. In simple terms, two images which can be represented as matrices are multiplied to give an output that is used to extract features from the image.

### **Basic Architecture**

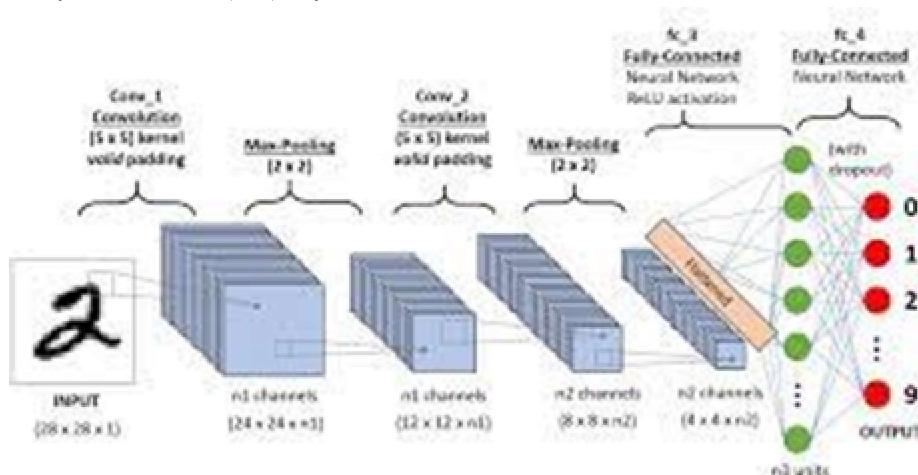
There are two main parts to a CNN architecture

- A convolution tool that separates and identifies the various features of the image for analysis in a process called Feature Extraction.
- The network of feature extraction consists of many pairs of convolutional or pooling layers.
- A fully connected layer that utilizes the output from the convolution process and predicts the class of the image based on the features extracted in previous stages.
- This CNN model of feature extraction aims to reduce the number of features present in a dataset. It creates new features which summarizes the existing features contained in an original set of features. There are many CNN layers as shown in the CNN architecture diagram.



Convolutional neural networks are distinguished from other neural networks by their superior performance with image, speech, or audio signal inputs. They have three main types of layers, which are:

- Convolutional layer
- Pooling layer
- Fully-connected (FC) layer



## 1. Convolutional Layer

This layer is the first layer that is used to extract the various features from the input images. In this layer, the mathematical operation of convolution is performed between the input image and a filter of a particular size  $M \times M$ . By sliding the filter over the input image, the dot product is taken between the filter and the parts of the input image with respect to the size of the filter ( $M \times M$ ).

The output is termed as the Feature map which gives us information about the image such as the corners and edges. Later, this feature map is fed to other layers to learn several other features of the input image.

The convolution layer in CNN passes the result to the next layer once applying the convolution operation in the input. Convolutional layers in CNN benefit a lot as they ensure the spatial relationship between the pixels is intact.

## **2. Pooling Layer**

In most cases, a Convolutional Layer is followed by a Pooling Layer. The primary aim of this layer is to decrease the size of the convolved feature map to reduce the computational costs. This is performed by decreasing the connections between layers and independently operates on each feature map. Depending upon method used, there are several types of Pooling operations. It basically summarises the features generated by a convolution layer.

In Max Pooling, the largest element is taken from feature map. Average Pooling calculates the average of the elements in a predefined sized Image section. The total sum of the elements in the predefined section is computed in Sum Pooling. The Pooling Layer usually serves as a bridge between the Convolutional Layer and the FC Layer.

This CNN model generalises the features extracted by the convolution layer, and helps the networks to recognise the features independently. With the help of this, the computations are also reduced in a network.

## **3. Fully Connected Layer**

The Fully Connected (FC) layer consists of the weights and biases along with the neurons and is used to connect the neurons between two different layers. These layers are usually placed before the output layer and form the last few layers of a CNN Architecture.

In this, the input image from the previous layers are flattened and fed to the FC layer. The flattened vector then undergoes few more FC layers where the mathematical functions operations usually take place. In this stage, the classification process begins to take place. The reason two layers are connected is that two fully connected layers will perform better than a single connected layer. These layers in CNN reduce the human supervision

There are two main types of pooling:

Max pooling: As the filter moves across the input, it selects the pixel with the maximum value to send to the output array. As an aside, this approach tends to be used more often compared to average pooling.

Average pooling: As the filter moves across the input, it calculates the average value within the receptive field to send to the output array.

The name of the full-connected layer aptly describes itself. The pixel values of the input image are not directly connected to the output layer in partially connected layers.

However, in the fully-connected layer, each node in the output layer connects directly to a node in the previous layer. This layer performs the task of classification based on the features extracted through the previous layers and their different filters. While

convolutional and pooling layers tend to use ReLu functions, FC layers usually leverage a softmax activation function to classify inputs appropriately, producing a probability from 0 to 1.

- **Dataset :** MNIST Fashion Dataset.

- **References/ Available link :**

1.<https://machinelearningmastery.com/how-to-develop-a-cnn-from-scratch-for-fashion-mnist-classification/>  
2.<https://www.kaggle.com/code/anindya2906/fashion-mnist-classification>

- **Questions and answer :**

- 1.What do you mean by Convolutional Neural Network?
- 2.Why do we prefer Convolutional Neural networks (CNN) over Artificial Neural networks (ANN) for image data as input?
- 3.Explain the different layers in CNN?
- 4.Explain the significance of the RELU Activation function in Convolution Neural Network?
- 5.Why do we use a Pooling Layer in a CNN?
- 6.What is the size of the feature map for a given input size image, Filter Size, Stride, and Padding amount?
7. Explain the terms “Valid Padding” and “Same Padding” in CNN?
- 8.What are the different types of Pooling? Explain their characteristics?
9. Does the size of the feature map always reduce upon applying the filters? Explain why or why not?
- 10.What is Stride? What is the effect of high Stride on the feature map?

- **Mcqs:**

1. Which is the most direct application of neural networks?
  - a) vector quantization
  - b) pattern mapping
  - c) pattern classification
  - d) control applications
2. What are pros of neural networks over computers?
  - a) they have ability to learn b examples
  - b) they have real time high computational rates
  - c) they have more tolerance
  - d) all of the mentioned

3. which of the following is false?

- a) neural networks are artificial copy of the human brain
- b) neural networks have high computational rates than conventional computers
- c) neural networks learn by examples
- d) none of the mentioned

● **Conclusion :**

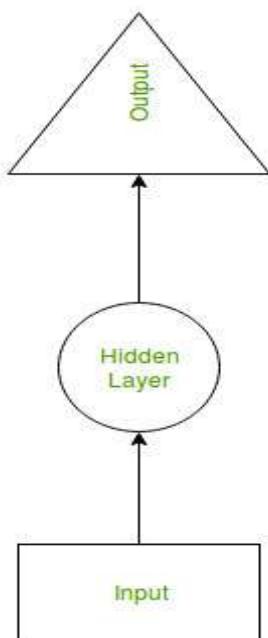
We have successfully created a classifier using Convolutional neural network.

## **Assignment No. 4**

- **Title:** Recurrent neural network (RNN) .
- **Date of Completion:**
- **Objective:** To study and understand Recurrent Neural Network by doing analysis and designing prediction systems.
- **Problem Statement:** Use the Google stock prices dataset and design a time series analysis and prediction system using RNN.

- **Theory:**

Recurrent Neural Network(RNN) is a type of Neural Network where the output from the previous step are fed as input to the current step. In traditional neural networks, all the inputs and outputs are independent of each other, but in cases like when it is required to predict the next word of a sentence, the previous words are required and hence there is a need to remember the previous words. Thus RNN came into existence, which solved this issue with the help of a Hidden Layer. The main and most important feature of RNN is Hidden state, which remembers some information about a sequence.



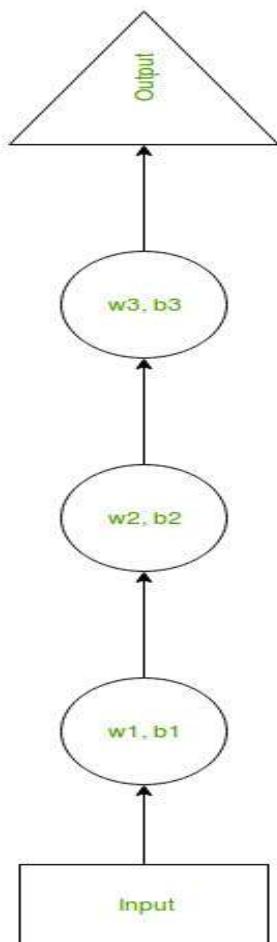
RNN have a “memory” which remembers all information about what has been calculated. It uses the same parameters for each input as it performs the same task on all the inputs or hidden

layers to produce the output. This reduces the complexity of parameters, unlike other neural networks.

#### How RNN works

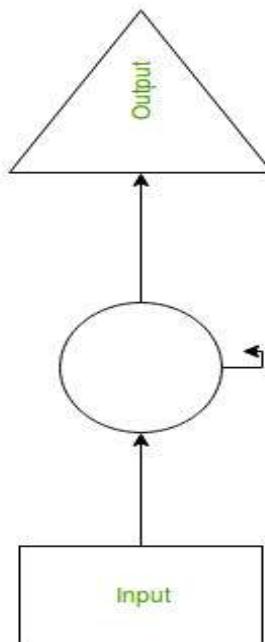
The working of an RNN can be understood with the help of the below example:

Example: Suppose there is a deeper network with one input layer, three hidden layers, and one output layer. Then like other neural networks, each hidden layer will have its own set of weights and biases, let's say, for hidden layer 1 the weights and biases are  $(w_1, b_1)$ ,  $(w_2, b_2)$  for the second hidden layer, and  $(w_3, b_3)$  for the third hidden layer. This means that each of these layers is independent of the other, i.e. they do not memorize the previous outputs.



Now the RNN will do the following:

- RNN converts the independent activations into dependent activations by providing the same weights and biases to all the layers, thus reducing the complexity of increasing parameters and memorizing each previous output by giving each output as input to the next hidden layer.
- Hence these three layers can be joined together such that the weights and bias of all the hidden layers are the same, in a single recurrent layer.



$$h_t = f(h_{t-1}, x_t)$$

The formula for calculating current state:

$h_t \rightarrow$  current state

$h_{t-1} \rightarrow$  previous state

$x_t \rightarrow$  input state

Formula for applying Activation function(tanh): where

current state: where:

here:

$whh \rightarrow$  weight at recurrent neuron

$wxh \rightarrow$  weight at input neuron

The formula for calculating output:

$$h_t = \tanh (W_{hh}h_{t-1} + W_{xh}x_t)$$

$$y_t = W_{hy}h_t$$

$y_t \rightarrow$  output

$Why \rightarrow$  weight at output layer

Training through RNN

1. A single-time step of the input is provided to the network.
2. Then calculate its current state using a set of current input and the previous state.
3. The current  $h_t$  becomes  $h_{t-1}$  for the next time step.
4. One can go as many time steps according to the problem and join the information from all the previous states.

5. Once all the time steps are completed the final current state is used to calculate the output.
6. The output is then compared to the actual output i.e the target output and the error is generated.
7. The error is then back-propagated to the network to update the weights and hence the network (RNN) is trained.

#### Advantages of Recurrent Neural Network

1. An RNN remembers each and every piece of information through time. It is useful in time series prediction only because of the feature to remember previous inputs as well. This is called Long Short Term Memory.
2. Recurrent neural networks are even used with convolutional layers to extend the effective pixel neighborhood.

#### Disadvantages of Recurrent Neural Network

1. Gradient vanishing and exploding problems.
2. Training an RNN is a very difficult task.
3. It cannot process very long sequences if using tanh or relu as an activation function.

#### Applications of Recurrent Neural Network

1. Language Modelling and Generating Text
2. Speech Recognition
3. Machine Translation
4. Image Recognition, Face detection
5. Time series Forecasting

#### Long short-term memory (LSTM)

This is a popular RNN architecture as a solution to vanishing gradient problem. It addresses the problem of long-term dependencies. That is, if the previous state that is influencing the current prediction is not in the recent past, the RNN model may not be able to accurately predict the current state. LSTMs have “cells” in the hidden layers of the neural network, which have three gates—an input gate, an output gate, and a forget gate. These gates control the flow of information which is needed to predict the output in the network.

- **Dataset :** Google stock prices dataset
- **References/ Available link**  
1.<https://towardsdatascience.com/lstm-for-google-stock-price-prediction-e35f5cc84165>  
2.<https://medium.com/linkit-intecs/developing-a-recurrent-neural-network-rnn-for-a-google-stock-price-dataset-having-continuous-data-97b6e4e724a3>
- **Questions and answer :**

- 1.What's the difference between Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) and in which cases would use each one?
- 2.How many dimensions must the inputs of an RNN layer have?
- 3.What are the main difficulties when training RNNs? How can you handle them?
- 4.What are the uses of using RNN in NLP?
- 5.What types of Recurrent Neural Networks (RNN) do you know?
- 6.What's the difference between Stateful RNN vs Stateless RNN? What are their pros and cons?
7. What's the difference between Traditional Feedforward Networks and Recurrent Neural Networks?
- 8.What's the difference between Recurrent Neural Networks and Recursive Neural Networks?
9. When would you use MLP, CNN, and RNN?
- 10.Why are RNNs (Recurrent Neural Network) better than MLPs at predicting Time Series Data?

- **Mcqs:**

- 1.RNN Stands for?
  - a) Recursive Neural Network
  - b) Recurrent Neural Network
  - c). Removable Neural Network
  - d).Recurring Neural Network
2. The main and most important feature of RNN is \_\_\_\_\_.
  - a) Visible State
  - b) Hidden State
  - c) Present State
  - d) none of above
- 3.RNN remembers each and every information through\_\_\_\_\_
  - a) Work
  - b) Time
  - c) Hours
  - d) Memory

- **Conclusion :**We have successfully implemented a recurrent neural network to create a classifier.