Title :-
Linear regression by using Deep Neural Network.

Objective :-
- To implement different deep learning models.
- To illustrate the concepts of Artificial Intelligence / Machine Learning (AIIML).

Problem Statement :-
- Implement Boston Housing Price prediction problem by Linear Regression using Deep Neural Network.

Software and Hardware Requirements :-
- 64-bit open source operating system or its derivative.
- Programming languages :- Python.

Theory :-
Deep Neural Network → A Deep Neural Network (DNN) is an ANN with example hidden layers between the input and output layers.
- Similar to shallow ANNs, DNNs can model complex non-linear relationships.
- The main purpose of a neural network is to receive a set of input, perfor progressively complex calculations on them & give output to solve the real world.

problems like classification. We restrict ourselves to feed forward neural networks.

- We have an input and an output and a flow of sequential data in a deep network.

- Neural networks are widely used in supervised learning and reinforcement learning problems.

- These problems or networks are based on a set of layers connected to each other.

- In Deep Learning, the number of hidden layers, mostly non-linear, can be large! say about 1000 layers.

- DL networks or models produce much better results than normal ML networks.

- We mostly use the gradient descent method for optimizing the network and minimising the loss function.

Linear Regression :-

- Linear Regression is a machine learning algorithm based on supervised learning.

- It performs a regression task.

- Regression models a target prediction value based on independent variables.

- It is mostly used for finding out the relationship between variables and forecasting.

-Different regression models differ based on - the kind of relationship between dependent and independent variables they are considering, and the number of independent variables getting used.

-There are many names for a regression's dependent variable. It may be called an outcome variable, criterion variable, endogeous variable or regressand.

-The independent variables can be called exogenous variables, predictor variables or regressors.

-Linear regression is used in many different fields, including finance, economics and psychology to understand and predict the behaviour of a particular variable.

-Linear regression performs the task to predict a dependent variable value (y) based on a given independent variable (x).

- Hence the name is Linear Regression.

-In fig 1.1, x (input) is the work experience and y (output) is the salary of a person.

-The regression line is the best fit line for our model.

Hypothesis Function For linear Regression :-

$$y = \theta_1 + \theta_2 . X.$$

Cost Function (J):

- By achieving the best fit regression line, the model aims to predict y value such that the error , difference between predicted value and true value is minimum.

- So, it is very important to update the θ_1 and θ_2 values, to reach the best value that minimize the error between predicted y value (pred)) and true y value (y).

$$\text{minimise } \frac{1}{n} \sum_{i=1}^{n} (pred_i - y_i)^2 .$$

$$J = \frac{1}{n} \sum_{i=1}^{n} (pred_i - y_i)^2 .$$

- Cost Function (J) of linear Regression is the Root Mean Squared Error (RMSE). between predicted y value (pred) and true y value (y).
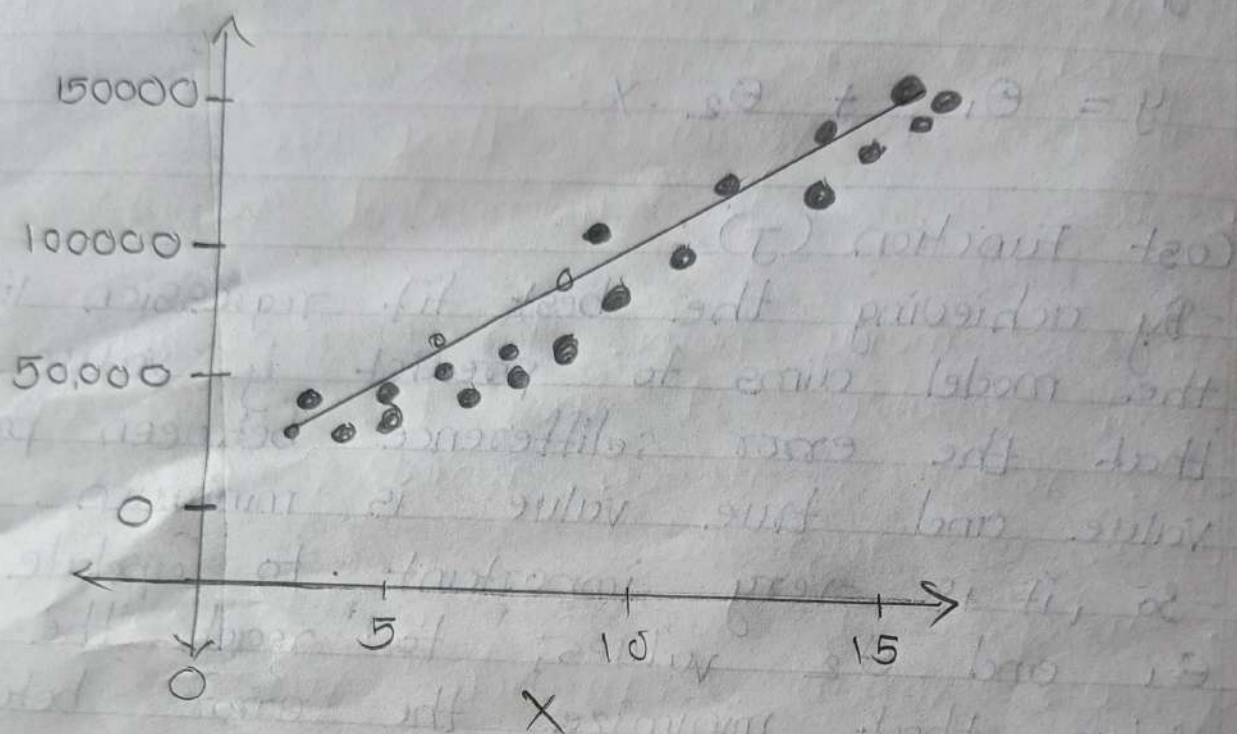
Fig → 1.1.

Gradient Descent :-
- To update θ_1 and θ_2 values in order to reduce cost function (minimizing RMSE value) and acheiving 'the best-fit line the model uses Gradient Descent.
- The idea is to start with random θ_1 and θ_2 values and then iteratively updating the values, reaching minimum cost.

Now let's understand some libraries of Python :-

D Pandas :-
- Pandas is an open-source library that is made mainly for working with rational or labeled data both easily and intuitively.
- It provides various data structures and operations for manipulating numerical data and time series.
- This library is built on top of the NumPy library.
- Pandas is fast and it has high performance & productivity for users.
- This module is generally imported as :

     import pandas as pd.
- Here, pd is referred to as an alias to the Pandas.

2) NumPy →

- NumPy stands for Numerical Python is a library consisting of multidimensional array objects and a collection of rutines for processing those arrays.
- Using NumPy, mathematical and logical operations on arrays can be performed.
- This module is generally imported as :

        import NumPy as np ;

3) matplotlib :→

- Matplotlib is an amazing visualization library in python for 2D plots of arrays.
- Matplotlib is a multi-platform data visualization library built on NumPy arrays and designed to with the broader Scipy stack.
- It was introduced by John Hunter in the year 2002.
- Matplotlib consists of several plots like line, bar, scatter, histogram, etc.
- This module is generally imported as :

        from matplotlib import pyplot as plt .
        or
        import matplotlib.pyplot as plt .

4) Seaborn →

- Seaborn is a library that uses matplotlib underneath to plot graphs.
- It will be used to visualize random distributions.
- This library is generally imported as:

import seaborn as sns.

5) %matplotlib inline →

- The %matplotlib inline command tells the IPython environment to draw the plots immediately after the current cell.
- The drawn plots are shown below the code and stored in the notebook document for the reference. (Practical File Name - D1-1.ipynb).

6) Scatter Plot →

- A Scatter Plot uses dots to represent values for two different numeric variables.
- The position of each dot on the horizontal and vertical axis indicates values for an individual data point.
- Scatter plots are used to observe relationships between variables.

7) Heatmap :-

- Heatmap is defined as a graphical representation of data using colors to visualize the value of the matrix.
- Heatmap is also defined by the name of the shading matrix.
- Heatmap in seaborn can be plotted by using the seaborn.heatmap() function.

8) Pair Plot →

- To plot multiple pairwise bivariate distributions in a dataset, you can use the .pairplot() function.

9) Dist Plot →

- It is used basically for univariant set of observations and visualizes it through a histogram i.e. only one observation and hence we choose one particular column of the dataset.

- Syntax →

    distplot (a[, bins, hist, kde, rug, ----]).

Conclusion →

- Hence, we implement the Boston housing price prediction problem successfully by using linear regression using Deep Neural Network..