Practical 2

Topics in Deep Learning

Regression and Binary Classification

❖ **Dataset**

- **User dataset -** This dataset contains information of users from companies database. It contains information about UserID, Gender, Age, EstimatedSalary, Purchased.

- **Pima Indians Diabetes Database -** This dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. The objective of the dataset is to diagnostically predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset. The datasets consists of several medical predictor variables and one target variable, Outcome. Predictor variables includes the number of pregnancies the patient has had, their BMI, insulin level, age, and so on.

- **50_Startups -** This dataset has data collected from New York, California and Florida about 50 business Startups. The variables used in the dataset are Profit, R&D spending, Administration Spending, and Marketing Spending**.**

❖ **Task to perform –**

- **Code your own functions for Gradient descent, Logistic regression, and Multiple linear regression**
- **Use python library (such as sklearn) to perform all these task and compare answers**

- **Use User database and predict that a user will purchase the company's newly launched product or not.**
- **Use Pima diabetes dataset and classify them as diabetes Or no diabetes based on features available.**
- **Predict the profit values on 50_startups dataset for each company.**