DS ASSIGNMENT-3

SUBMITTED TO:
PROF. MANDAR KARYAKARTE

VISHWAKARMA INSTITUTE OF INFORMATION TECHNOLOGY, PUNE

COMPUTER ENGINEERING DEPARTMENT

BY:
NAME: ABHISHEK MORE

G.R No.: 21810033

ROLL NO.: 323036

CLASS: T.Y COMP

BATCH: COMP-C2

Assignment-3

Description: Building a Linear Regression model in Python for the Covid19 dataset.

Linear Regression Model: In statistics, linear regression is a linear approach to modelling the relationship between a scalar response (or dependent variable) and one or more explanatory variables (or independent variables). The case of one explanatory variable is called simple linear regression. Such models are called linear regression models.

Covid19 Dataset: It contains information about the Countries, deaths, recovered, etc and the updated time when data was updated. New cases and new deaths are also there in the dataset .

Interpretation: I checked the correlations between different features (columns) of the dataset and built a basic machine learning(Linear Regression) model with this dataset. Finally I predicted my target variable. My target variable from this dataset is confirmed cases and I tried to find its relation between other variables, so that I can say that peoples will get infected this much for example, if Time on App is increased 1 minute more. I plotted graphs to analyse the correlations.

Outcomes: Patients Recovered has positive and moderate correlation with Confirmed cases. (86%) Deaths also has a correlation, more powerful than recovered. (97%).It means deaths are less

I later reached to this outcome that Deaths has really perfect linearity, so if number of covid cases increase the deaths are less . Linear Regression Score For this Model is 99.13 %. Hence, I can say that I predicted good values.