CSE523: Machine Learning

Project Report Week- 5

➢ Team name.: **Tech_mak**

➢ Name & Roll no.:

Meet Pedhadiya	AU1841056
Ayush Kaneria	AU1841062
Mihir Chauhan	AU1841065
Krunal Pagdar	AU1841066

➢ Tasks Performed in the week

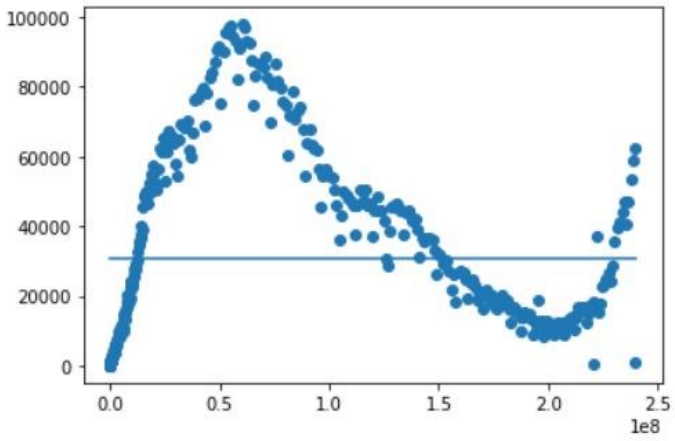
- Data filters on datasets
- In the dataset fields of deaths, recovered and confirm cases of covid-19 contains total data since that date
- Add columns from the respective fields' total value
    - Daily recovered
    - Daily Deaths
    - Daily confirm cases
- Perform **Linear Regression** using in-built library(**sk-learn**) on the datasets of different states and for India as well
- Perform **Polynomial Regression** using in-built library(**sk-learn**) library on the datasets of different states and for India

➢ Outcomes of the tasks performed

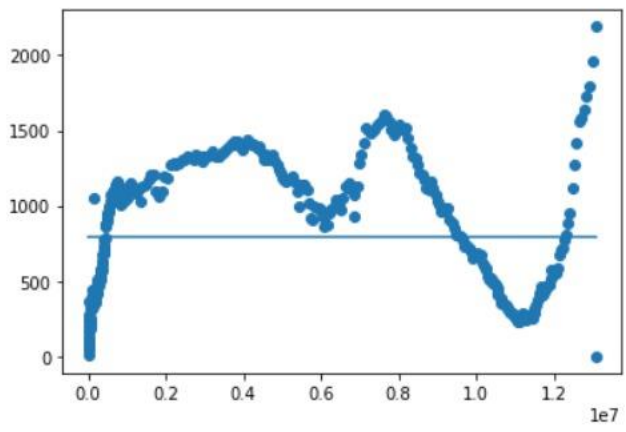
- Datasets having more useful columns like
    - Daily recovered
    - Daily Deaths
    - Daily confirm cases

Date	State	Confirmed	Recovered	Deceased	Other	Tested	Daily confirmed	Daily_recovered	Daily_deaths
13-03-2020	India	91	3	1	0	6500	10	3	1
14-03-2020	India	102	3	1	3	6500	11	0	0
15-03-2020	India	112	3	1	3	6500	10	0	0
16-03-2020	India	126	3	1	3	6500	14	0	0
17-03-2020	India	146	3	1	3	6500	20	0	0
18-03-2020	India	171	3	1	3	13125	25	0	0
19-03-2020	India	198	3	1	3	14175	27	0	0
20-03-2020	India	256	4	1	3	15404	58	1	0
21-03-2020	India	334	4	1	3	16911	78	0	0
22-03-2020	India	403	4	1	3	18127	69	0	0
23-03-2020	India	497	4	1	3	20707	94	0	0
24-03-2020	India	571	6	1	3	22694	74	2	0
25-03-2020	India	657	6	1	3	25144	86	0	0
26-03-2020	India	730	6	2	3	27688	73	0	1
27-03-2020	India	883	8	3	3	27688	153	2	1
28-03-2020	India	1019	8	3	3	27688	136	0	0
29-03-2020	India	1139	99	28	3	27688	120	91	25
30-03-2020	India	1326	141	41	3	38442	187	42	13
31-03-2020	India	1635	160	47	3	42788	309	19	6

- **Linear Regression** on the data using sk-learn library
  - Independent variable → Daily tested
  - Dependent variable → Daily confirm cases
    - India

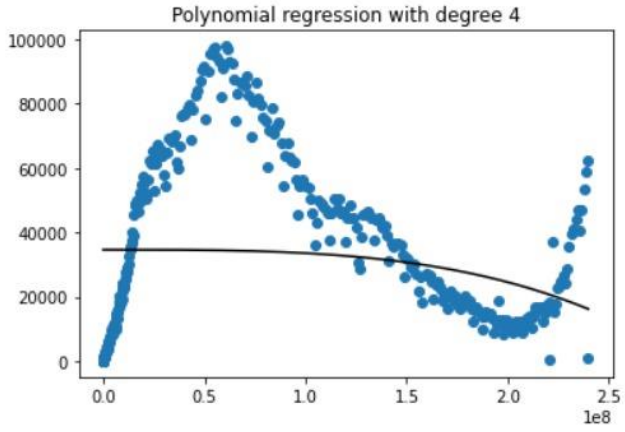


    - Gujarat

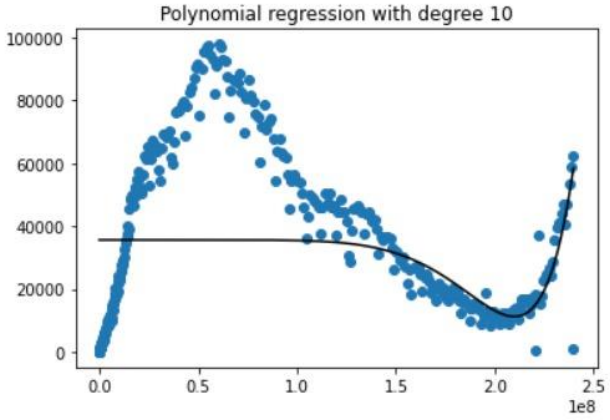


- Polynomial Regression on the data using sk-learn library
  - Independent variable → Daily tested
  - Dependent variable → Daily confirm cases

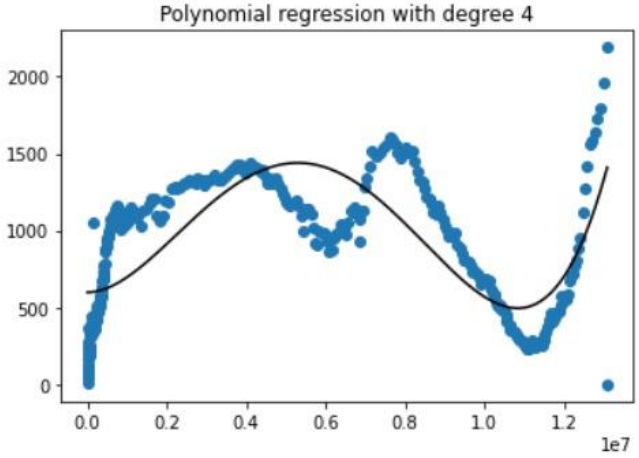
  - Data of India - Degree-4



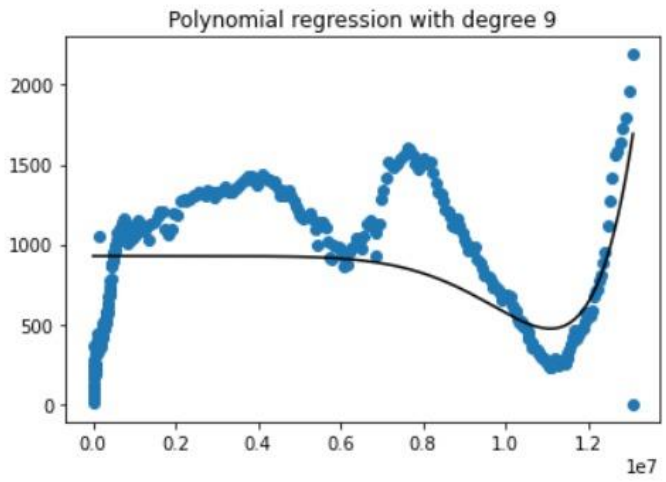
  - Data of India - Degree-10



  - Degree-4



  - Data of Gujarat – Degree – 9



- We have performed linear and polynomial regression but we don't get efficient output because of the lack of the independent features on that covid-19 daily cases.

➢ Tasks to be performed in the upcoming week

- We didn't get efficient initial output from the linear and polynomial regression.
- So, we will try to use autoregressive and time series analysis.
- Find algorithms for autoregression and time analysis
- Add one more column to the dataset – Active cases