CO2 EMISSION CLUSTERING & FITTING FOR DIFFERENT COUNTRIES

INTRODUCTION

Emissions of carbon dioxide are those that come from the burning of fossil fuels. They consist of gas flaring and carbon dioxide produced during the consumption of solid, liquid, and gas fuels. Decomposition, oceanic release, and respiration are a few sources of CO2 emissions. This investigation focuses on the CO2 emissions that were produced between the years 1980 and 2030 in various nations. To meet this requirement, two clusters have been formed. The primary benefit of this activity I is that it allows us to gauge the risk posed by CO2 emissions and take the necessary precautions to keep them under control.

Fig.1 - World map for clustering

Figure 1's geographic graph classifies some countries into Cluster 1—New Zealand, China, India, Afghanistan, Albania, Austria, and Algeria—while classifying others into Cluster 2—the United States, Canada, and Australia.

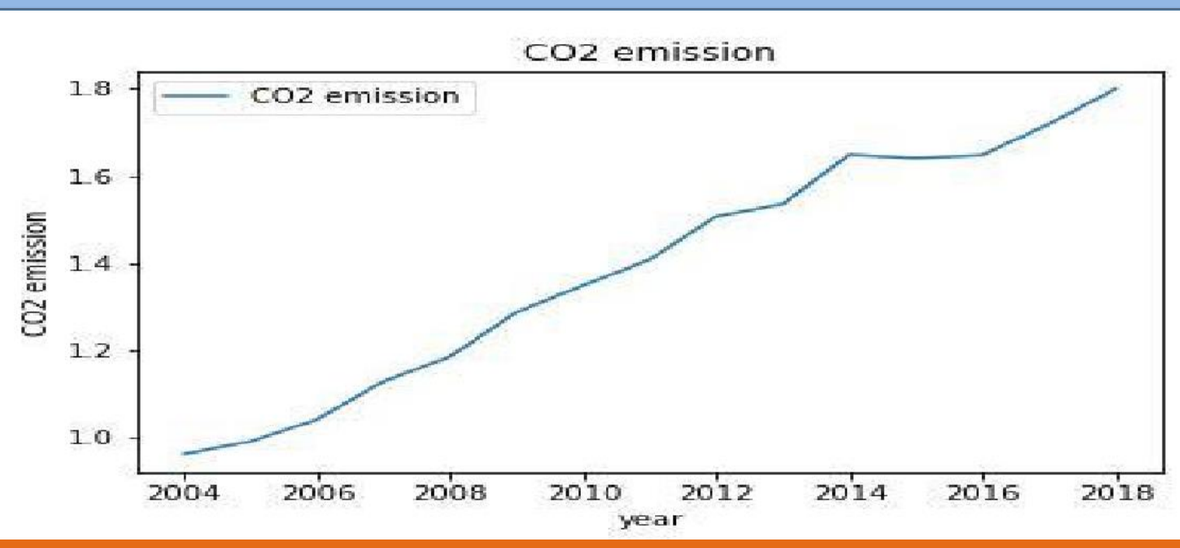


Fig.2-CO2 emission

Using the plot shown above, we can examine the trend of the range of CO2 emissions from 2004 to 2018 and see that they have been rising annually.

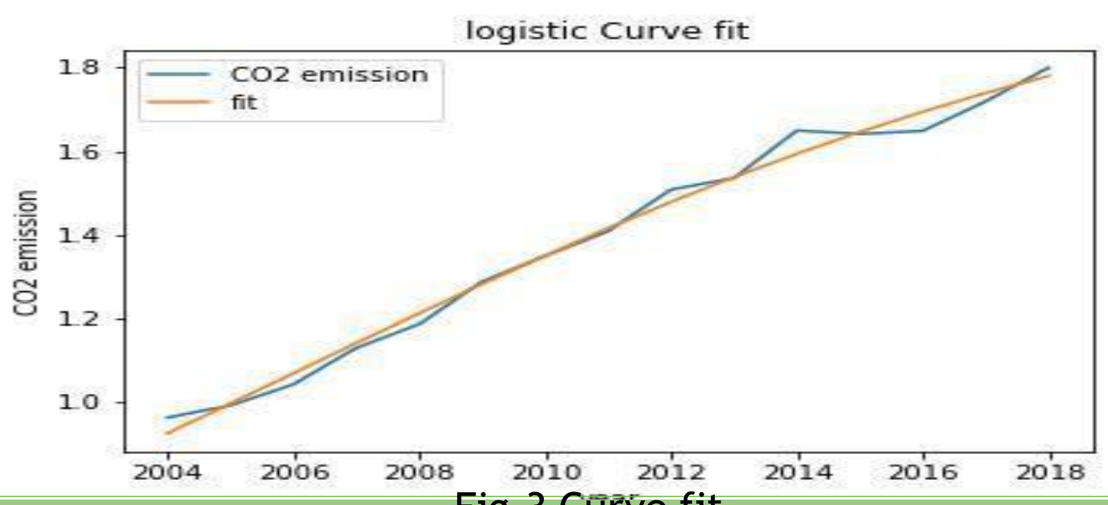


Fig.3 Curve fit

We are calculating the co2 emission from the aforementioned plot for the years 1980 - 2030, and the forecast is accurate for the entire time frame.

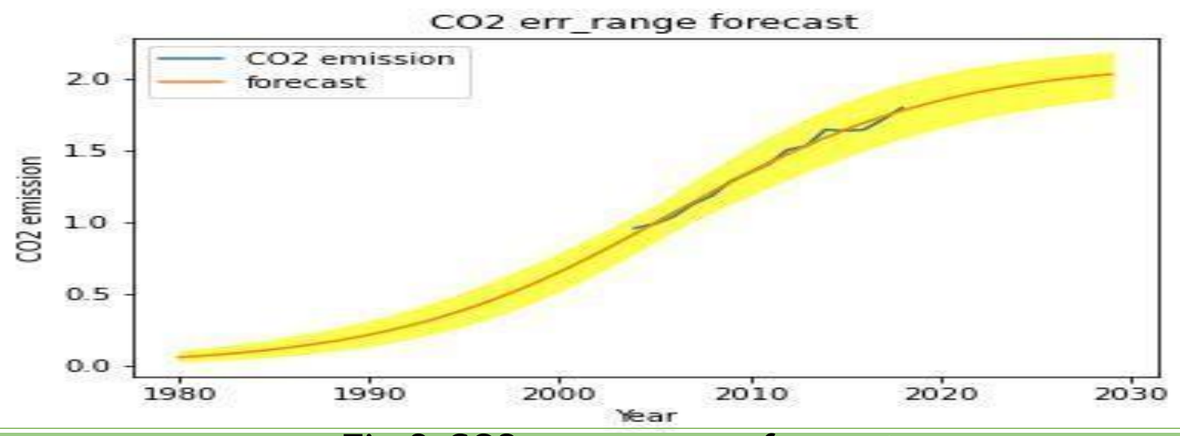
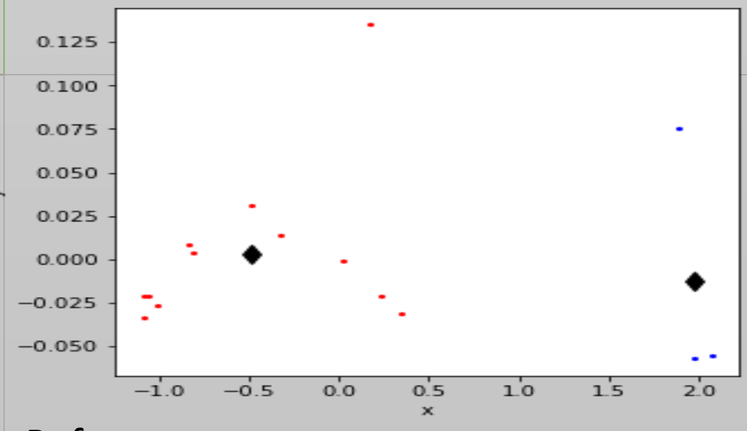


Fig.3-CO2 error ranges forecast

We are attempting to identify large intervals of data for which the proportional growth rate is linear by applying curve fit to the data in figure.3 The graph for features of CO2 emissions and year for India allows us to clearly see this phenomenon.

Conclusion:

After using K means clustering to divide the world bank data into two clusters, analyzing the clusters, and looking at various correlations between the features, we can determine the forecast of CO2 emissions for the upcoming years based on historical data evidence. This serves as a signal and aids in projecting future country-level increases in CO2 emissions. Results are identical whether normalization is used or not. After applying the curve fit, we calculated the err_ranges, and the results were 1.88919 -2.19177. From 1980 to 2030, predictions of CO2 emissions were made.



←--After K-Means clustering, extracting the estimated cluster centres

References:

[1]-Sudhir Singh, Dr. Nasib Singh Gill,Comparative Study Of Different Data Mining Techniques : A Review, www. ijltemas.in, Volume II, Issue IV, APRIL 2013 IJLTEMAS ISSN 2278 – 2540.

[2]-N.K. Lünsdorf and J.O. Lünsdorf 2016 Journal of Coal Geology 160–161 51

Github link:
https://github.com/niveditha2000/Clustering-and-Fitting.git