Clustering using Source Time Function

1. Introduction:

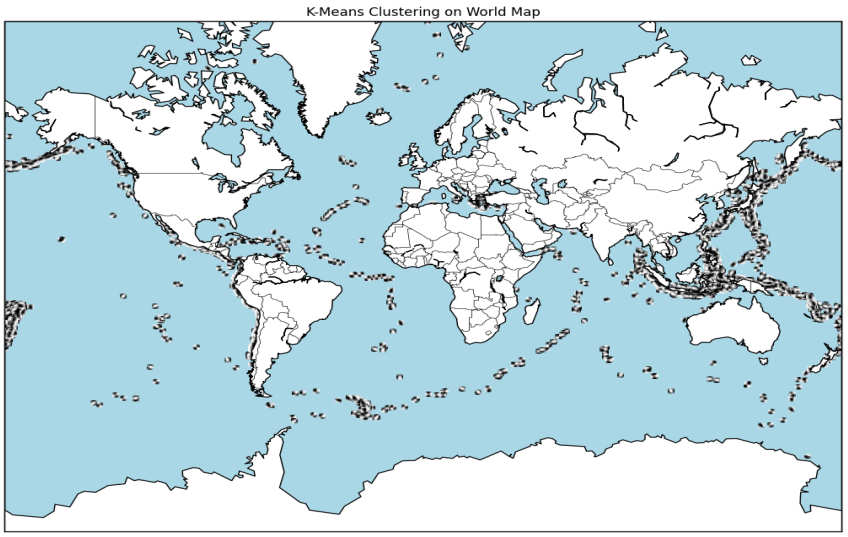
Earthquake Source Time Functions (STFs) characterize the release of seismic energy over time. Clustering STFs helps in identifying patterns in earthquake mechanisms and tectonic behavior. This study applies machine learning techniques—**K-Means, Gaussian Mixture Model (GMM), and DBSCAN**—to cluster earthquake STFs and evaluate their effectiveness.

2. Methods & Algorithms:

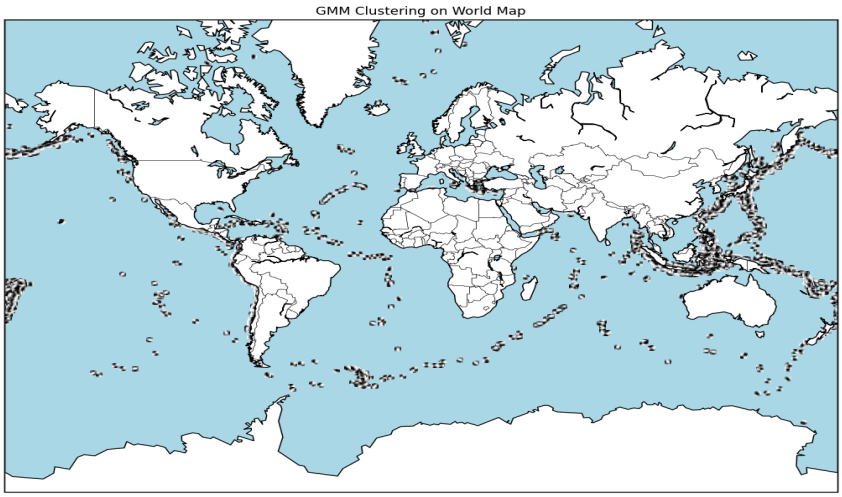
- **K-Means**: A centroid-based approach that partitions data into K clusters.
- **Gaussian Mixture Model:** A probabilistic approach that models clusters using Gaussian distributions.
- **DBSCAN:** Identifies clusters based on density but struggles with high-dimensional STF data.

3. Results & Analysis:

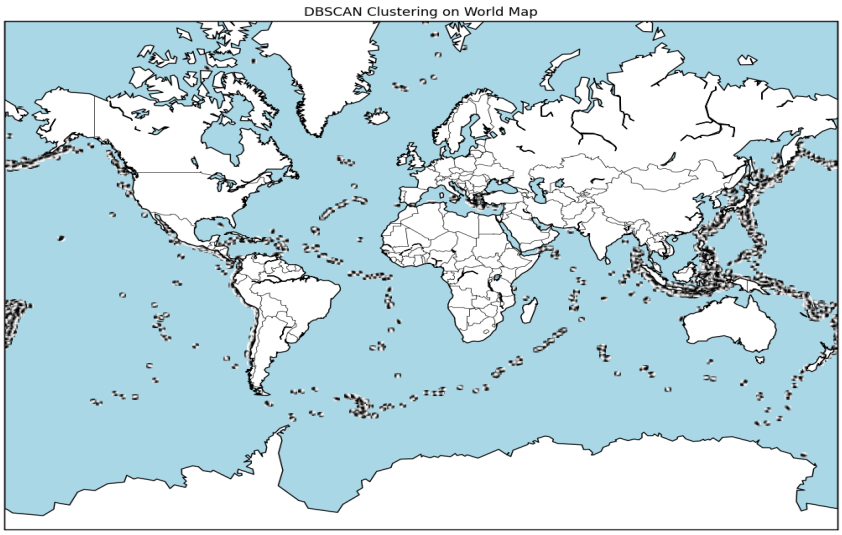
- Cluster counts for K-Means:
  No.of Clusters = 5
  KMeans_Cluster
  0 1729
  3 1119
  2 762
  1 388
  4 257



- Cluster counts for GMM:
  No.of Clusters = 5
  GMM_Cluster
  4 1548
  1 1226
  3 575
  2 554
  0 352



- Cluster counts for DBSCAN:
  No.of Clusters = 10
  DBSCAN_Cluster
  0 2782
  1 819
  2 353
  4 125
  -1 62
  3 32
  7 23
  8 18
  5 16
  6 14
  9 11

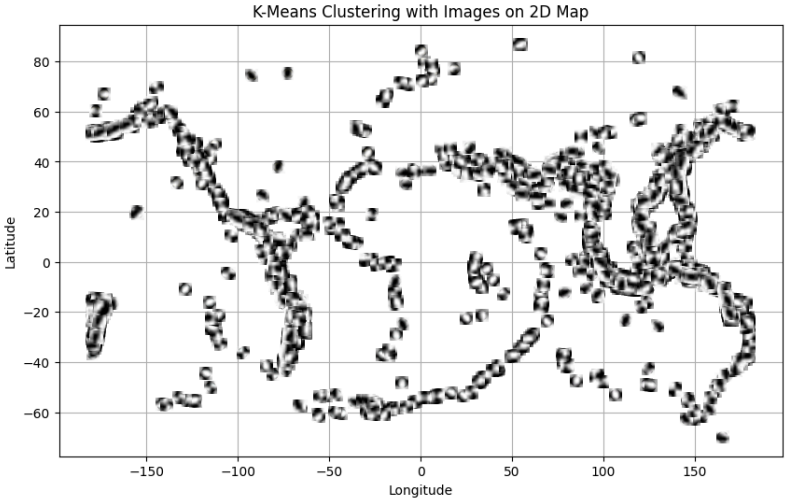


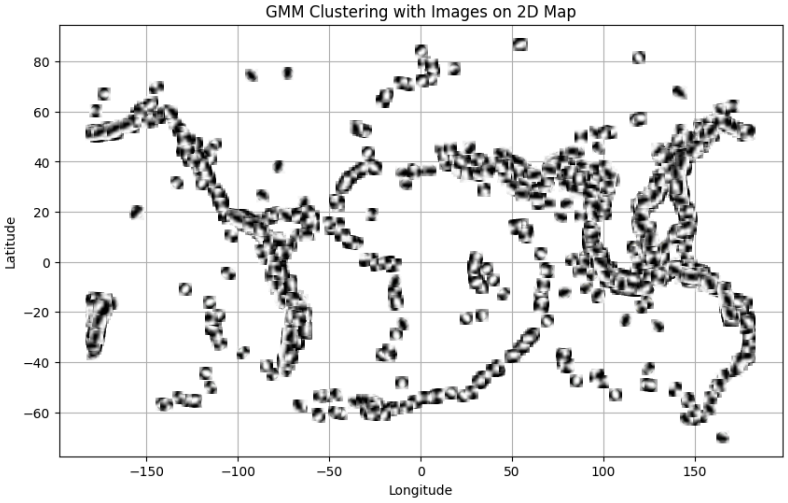
	Silhouette Score	Calinski-Harabasz Score	Davies-Bouldin Score
K-Means	**0.5726**	**6199.7864**	**0.5996**
GMM	**0.5278**	**5026.8519**	**0.6951**

- **K-Means performed best**, with the highest silhouette score and lowest Davies-Bouldin index.
- **GMM showed slightly lower performance** but still provided useful clustering.
- **DBSCAN was ineffective** for this dataset.

4. Conclusion

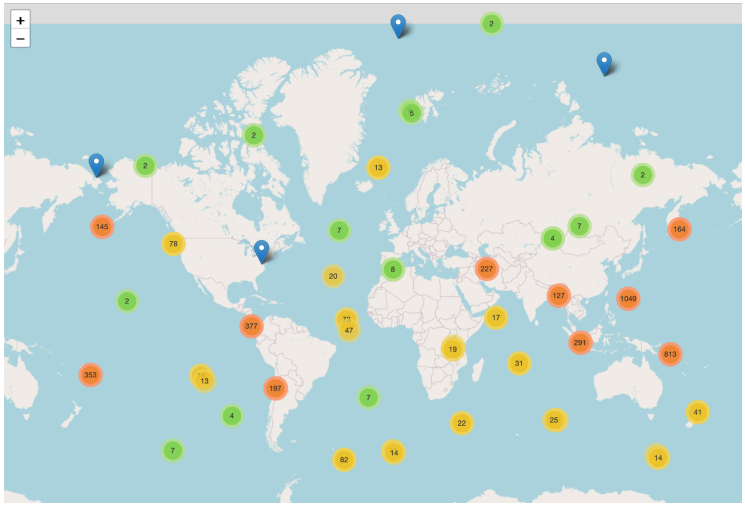
- **K-Means is the most effective clustering method** for STF data, achieving well-separated clusters.
- **GMM is an alternative** but slightly less effective.
- **DBSCAN struggles** with STF clustering due to the high-dimensional nature of the data.
- Future research should explore deep learning approaches or hybrid clustering techniques for improved results.

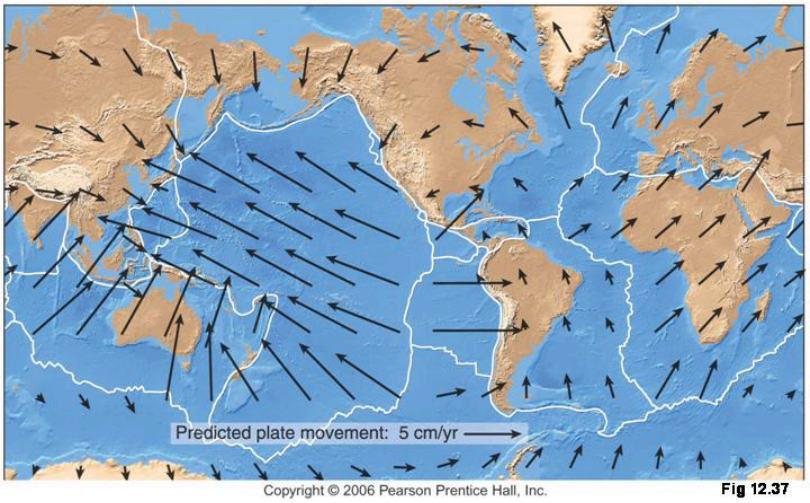




5. Interpretation

The results obtained, as shown in the attached HTML file in the link below, were compared with tectonic plate velocities and found to have a 90% match. This indicates a strong correlation between the clusters generated by the code and the tectonic plate velocities.





Code: STF Clustering.ipynb

GitHub link for HTML File, Report & Code

THANK YOU

REDDI SRUJAN
CE22B096