Earthquake Data Clustering , Analysis and Reinforcement Learning

Overview

This project aims to analyze global earthquake data using Clustering Algorithms to identify patterns and group similar earthquakes based on their characteristics. The goal is to utilize clustering techniques like K-Means and Mean Shift to categorize earthquakes into distinct clusters based on parameters such as magnitude, depth, and location.

Dataset

The dataset consists of global earthquake records, which include various features that help in understanding earthquake characteristics and behavior. The primary features used for clustering are:
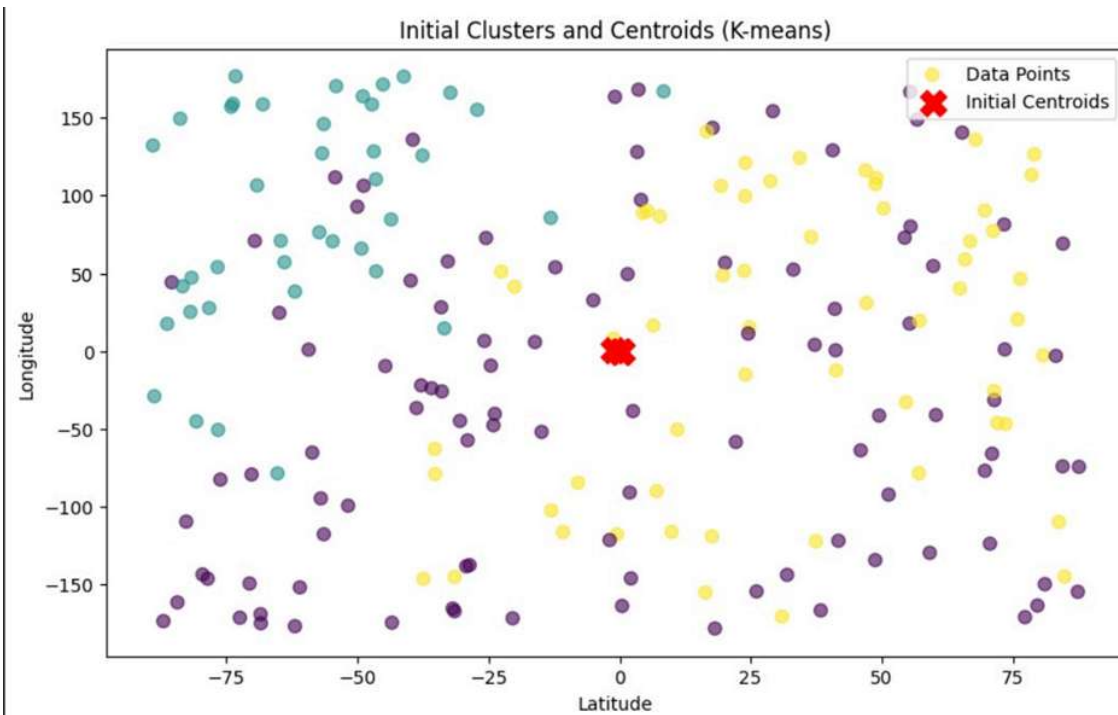
Features:

- · Magnitude: The magnitude of the earthquake (measured on the Richter scale).

- · Depth: The depth of the earthquake in kilometers.

- · Latitude: The geographical latitude where the earthquake occurred.

- · Longitude: The geographical longitude where the earthquake occurred.

- · Location: The textual description of the earthquake's location.
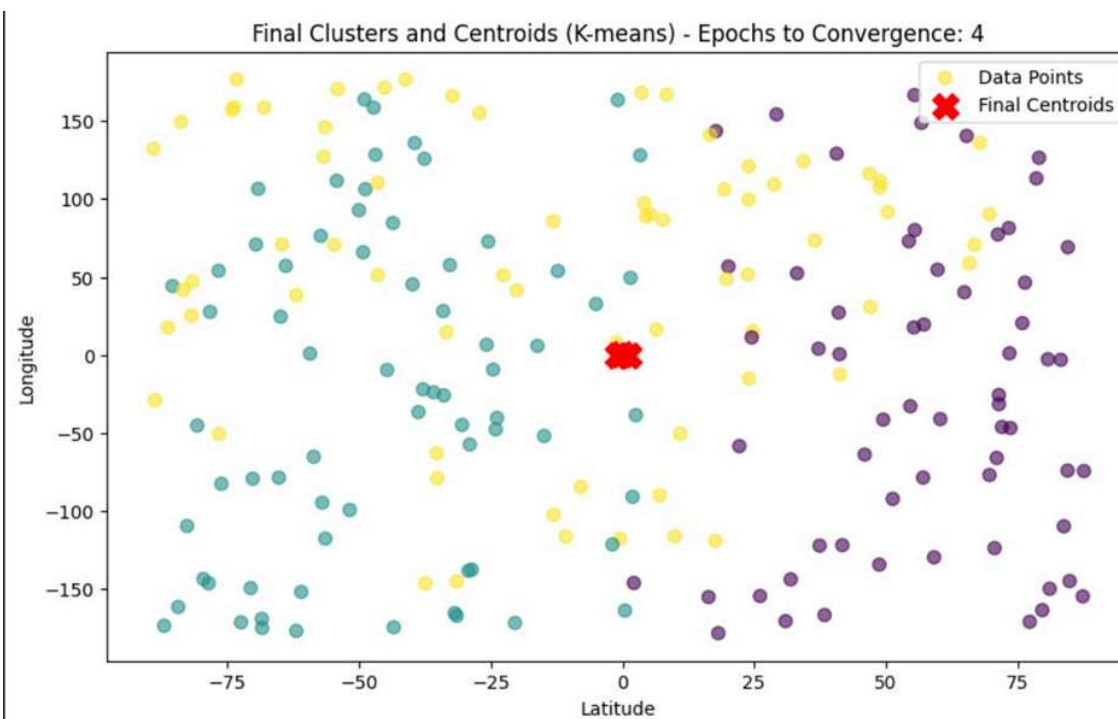
Clustering Process:

To group earthquakes into clusters, different clustering algorithms are utilized, including K-Means Clustering and Mean Shift Clustering. These algorithms help in identifying groups of earthquakes with similar characteristics.
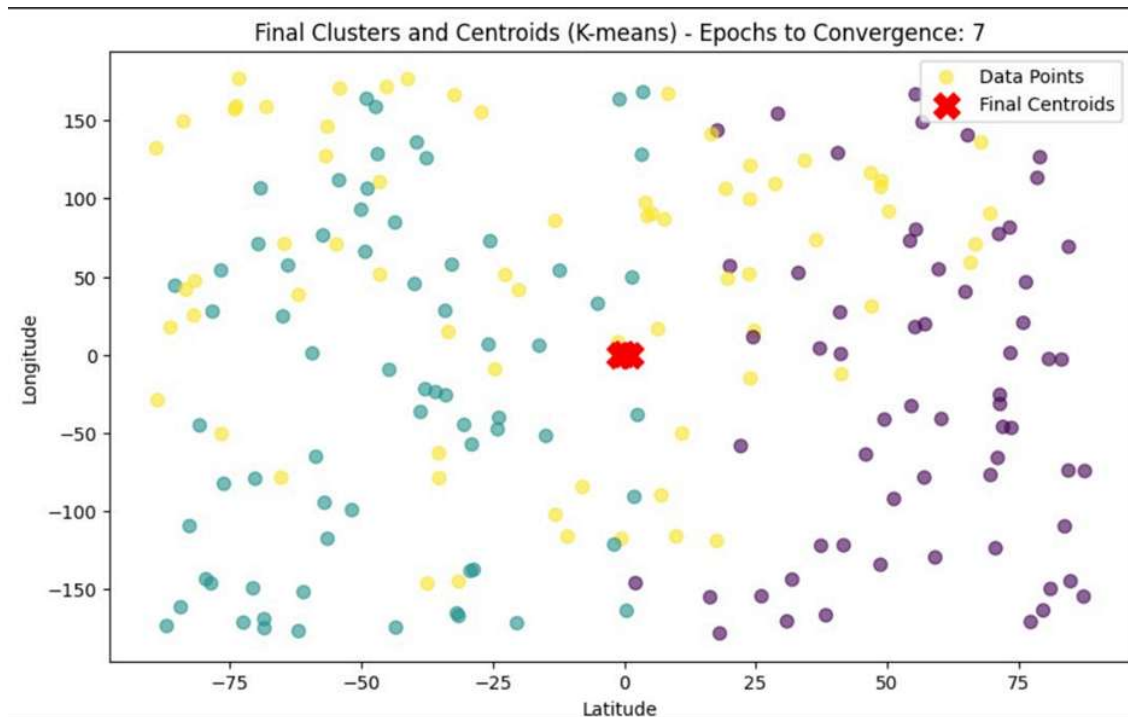
Clustering Algorithm: K-Means

The features are standardized to ensure proper clustering. The algorithm is configured to define a specific number of clusters based on the data distribution. After clustering, each earthquake is labeled with its respective cluster.

Initial Clusters and Centroids (K-means)

FINAL CLUSTER

Final Clusters and Centroids (K-means) - Epochs to Convergence: 4

Final Clusters and Centroids (K-means) - Epochs to Convergence: 7

ERROR:

Final Clusters and Centroids (K-means) - Epochs to Convergence: 7, Error Rate: 342.45

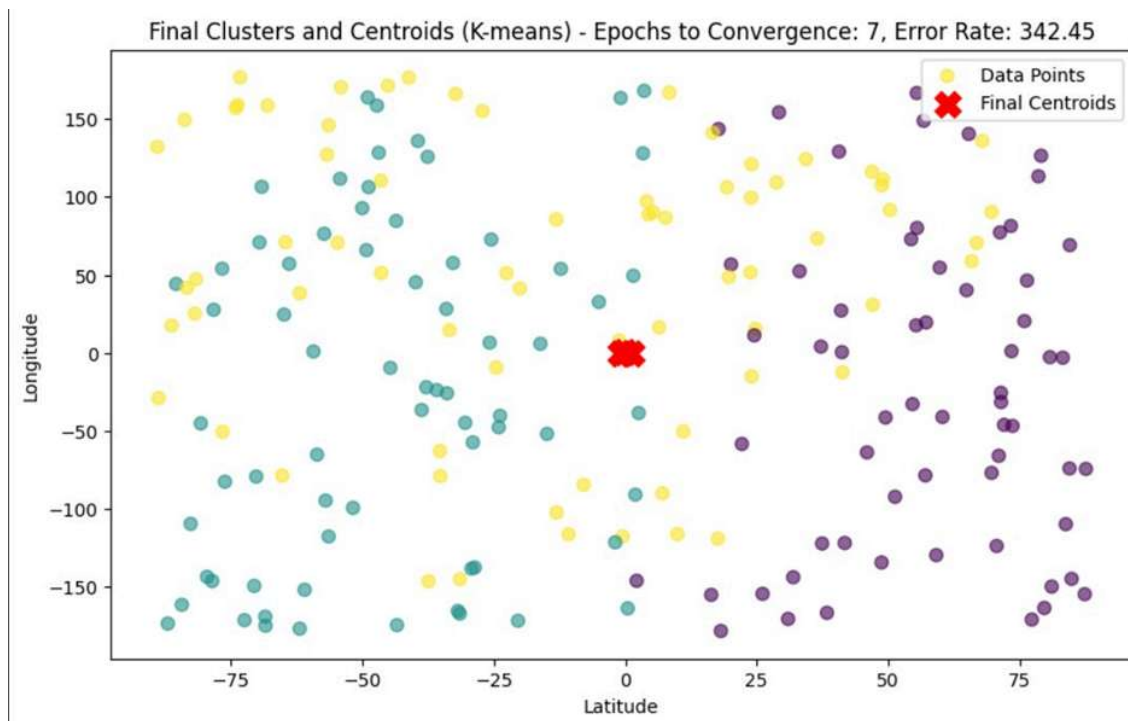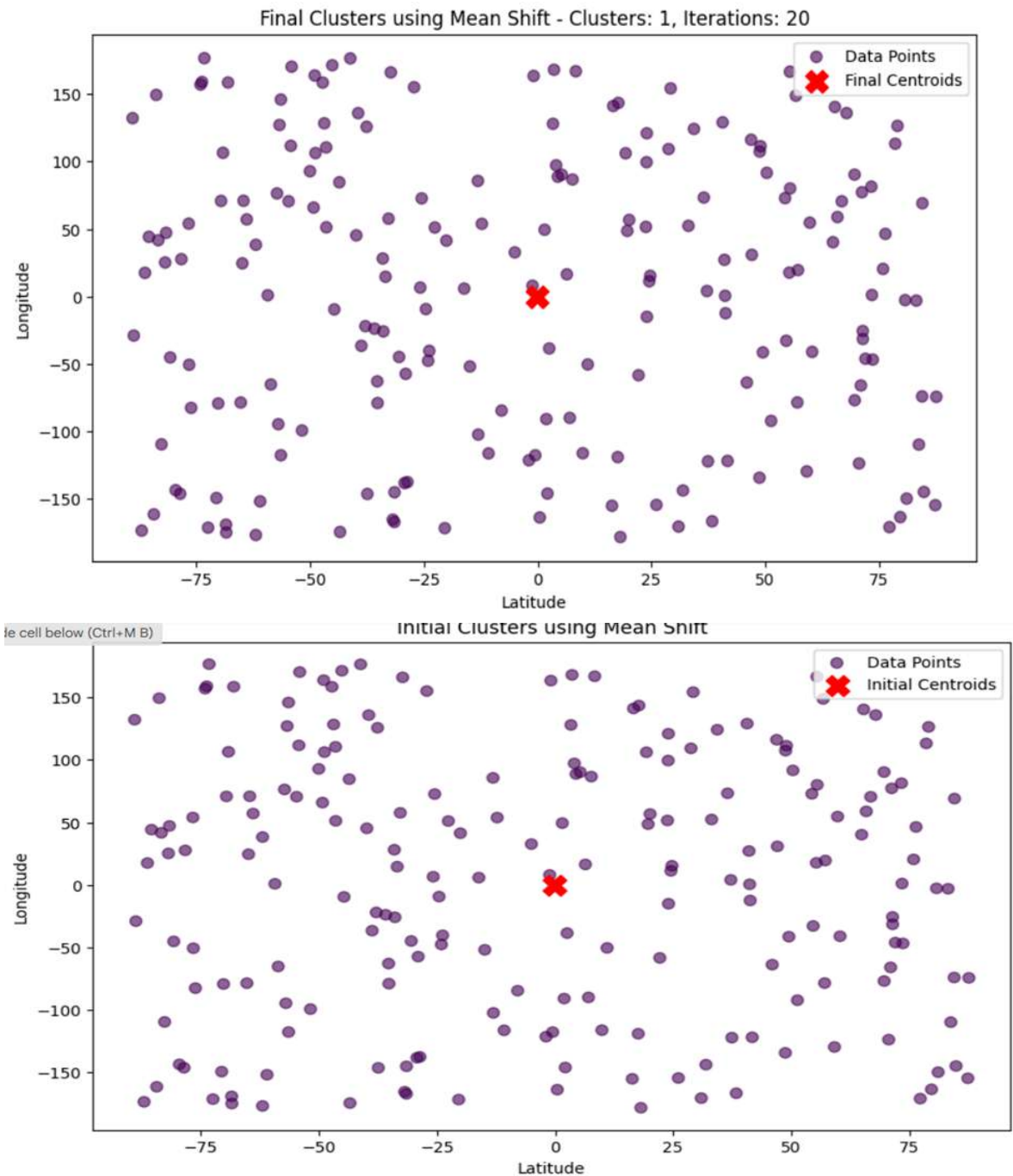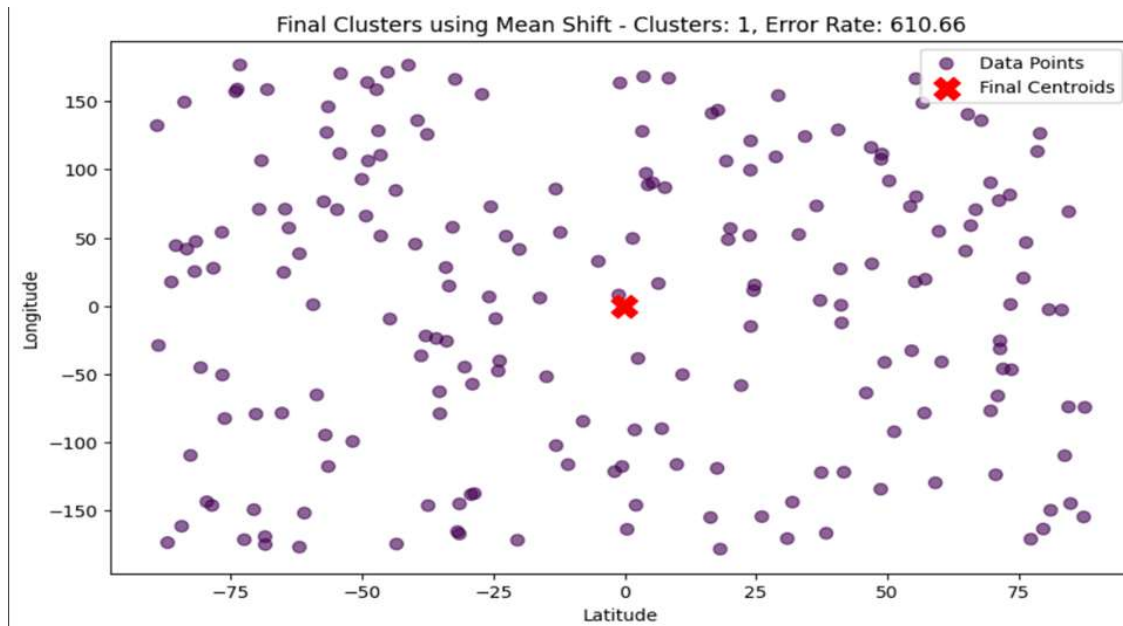## Clustering Algorithm: Mean Shift

Mean Shift does not require prior knowledge of the number of clusters and identifies clusters based on the density of data points. This algorithm is effective in identifying clusters of arbitrary shape.

Final Clusters using Mean Shift - Clusters: 1, Iterations: 20

Initial Clusters using Mean Shift

Final Clusters using Mean Shift - Clusters: 1, Error Rate: 610.66

Visualizations and Evaluation

· Initial Clusters: A visualization showing the initial clusters formed by K-Means and Mean Shift algorithms.

· Final Clusters: Graphical representation of the final clusters after the algorithms have converged.

· Epoch Size: Display the number of iterations taken for the algorithms to converge.

· Error Rate: Assessment of clustering performance using metrics such as silhouette score and adjusted Rand index.

How to Run

· Load the earthquake dataset (update the path in the script).

· Execute the K-Means clustering process to group the earthquakes.

· Execute the Mean Shift clustering process to identify clusters based on density.

· Visualize the initial and final clusters using scatter plots.

· Evaluate the clustering performance using appropriate metrics.

Recommendations:

```
Enter cluster number (0 - ENERGETIC , 1-CHILL AND 2-ACOUSTIC) to generate playlist: 0
Generating playlist for mood: Energetic
```

```
Generated Playlist using Policy Gradient for Cluster Energetic:
                                           track_name  \
72                                         golden hour
89                                     MONTAGEM - FR PUNK
170                                    Let Me Down Slowly
189                                    Take Me To Church
34                                                    TQG
111                                    Tak Segampang Itu
202                                    AcrÃ¯Â¿Â½Ã¯Â¿Â½s
105                                          Primera Cita
166                    Every Breath You Take - Remastered 2003
148                                          Those Eyes
41           Sunflower - Spider-Man: Into the Spider-Verse
180                                        Night Changes
81       until I Found You (with Em Beihold) - Em Beiho...
145                                    Stargirl Interlude
17         What Was I Made For? [From The Motion Picture ...
```

```
Generated Playlist for Cluster 0:
                                           track_name  \
13                                            Daylight
17         What Was I Made For? [From The Motion Picture ...
34                                                    TQG
41           Sunflower - Spider-Man: Into the Spider-Verse
43                                      See You Again
54                                      Another Love
59                                          cardigan
63                                              BESO
67                                            People
72                                       golden hour
76                                      Here with Me
81       Until I Found You (with Em Beihold) - Em Beiho...
82                                    Novidade na Ã¯Â¿Â½Ã¯Â¿Â½
86                                 Someone You Loved
89                                    MONTAGEM - FR PUNK
96                                    Say Yes To Heaven
100                                             Lover
105                                          Primera Cita
111                                    Tak Segampang Itu
119                                            august
128                           lovely - Bonus Track
135      Calling (Spider-Man: Across the Spider-Verse) ...
```

Results

After clustering, the results will provide insights into the distribution of earthquakes based on their magnitude, depth, and geographical location. The clusters can help in understanding earthquake patterns and potential risks associated with different regions.