Chapter 13
Enhanced K-Means Clustering Algorithms in Pattern Detection of Human Freedom Index Dataset

F.Mohamed Ilyas
Bharath Institute of Higher Education and Research, India

S. Silvia Priscila
Bharath Institute of Higher Education and Research, India

ABSTRACT

The human freedom index (HFI) evaluates the universal state of social liberty using a wide metric that includes individual, public, and financial liberty. Human freedom is a public notion that affirms a person's self-respect and is described here as undesirable freedom or the nonappearance of coercion. Since liberty is fundamentally valued and contributes to social development, it is worth measuring cautiously. This study emphasizes using the k-means clustering technique to locate clusters in data, with the inconstant k representing the number of clusters. After the groups have been gathered, this method will be tested with several k-clusters defining metrics in order to find the best k value for the model and group the data into the correct cluster counts. This study aims to examine existing data mining approaches for k-means clustering and mini batch k-means clustering and develop ways to improve accuracy by looking at a large number of statistics and choosing those with a specific shape using the human freedom index (HFI).

INTRODUCTION

HFI assesses financial liberties such as the ability to craft and custom sound change, as and notch to which individuals in the nations surveyed are allowed to exercise major civic rights such as liberty of expression, religious conviction, connotation, and gathering (Vásquez & Mcmahon, 2018; Tak et al., 2023). In contrast to same-sex couples, there are other indices for the rule of commandment, corruption

DOI: 10.4018/979-8-3693-5951-8.ch013

and wildness, liberty of travel, and lawful discernment. It includes characteristics of women's liberty that start across the index groups (Aditya Komperla, 2023). The HFI is a device for perceiving associations between liberty and other public financial spectacles. The scopes of liberty neutrally interrelate (Marar et al., 2023).

The core aim of this declaration is to make available a broad but fairly exact image of global liberty (Angeline et al., 2023). A bigger goal is to know what liberty means and how it relates to various public and commercial issues (Kumar et al., 2023). This study could also aid us in observing more quantitatively how several liberties relate to one another, such as economic and civic liberties (Zong et al., 2020). We anticipate this directory will become a valuable source for academics, policymakers, and similar curious individuals. Its worth will grow as it is reorganized yearly, permitting us to track various interactions across periods. Unsubstantiated clustering is a well-known technique (Bala Kuta & Bin Sulaiman, 2023). This divides the data into many linked classes irrespective of whether or not you have any former facts of class characterizations and hierarchy, used to find groups or clusters of things in a massively large number of records (Kaur & Shah, 2022). It's one of the most basic data structure analyses used in various fields, including ecology, finances, and thinking (Bose et al., 2023).

The goal is to organize things (Yuxin et al., 2019). Items in the same group should be the same as far as possible and in a manner that is distinct from the target of the investigation (Sapkota et al., 2019) as many opposed clusters or groups as feasible clustering of files used to arrange text documents that can be utilized for a variation of resolutions such as DM, information retrieval. Data contains a variety of concealed outlines (Qukai & Chi, 2019). With one, among others, in a neutral fashion (Obaid et al., 2023). With the quick advancement of big data, the measure and length of high-dimensional data contain a variety of concealed outlines (Pandit, 2023). These concealed patterns may be gathered and support us in getting perceptions through them using clustering systems (Nirmala et al., 2023). K –means clustering, which has proven useful in finding new information. Human Freedom Index (HFI) demographics are illustrated in the figure 1. The most widely used partition strategies are K-mean and its variations. The temporal difficulty of segmentation algorithms is almost direct and linear (Kuragayala, 2023; Sengupta et al., 2023). The K-mean clustering procedure is a data partitioning process that divides data into groups. It begins by randomly initializing group centroids and then assigning data opinions to the centroids that are neighboring (and most parallel) (Guihua et al., 2018; Yinyang, 2016; Yalavarthi & Boussi Rahmouni, 2023).

The same technique is repeated until the finish requirement is met (either a particular number of iterations are finished or clusters exhibit no change after a specified number of iterations) (Priscila & Hemalatha, 2018). It is a centroid-based approach in which data arguments are assumed to be spherically distributed about the cluster's center, with the group represented by a solitary center idea and the points (Liu, 2017; Yu et al., 2018).

The following goals are linked with this research:

- To learn about the different K-Means Clustering algorithms and to conduct experimental data examination on the data to gain perceptions and knowledge from it.
- The data will be examined using the K-Means clustering and Mini Batch K-Means clustering algorithms to establish the investigation strategy for finding the pattern in the data (Regin et al., 2023a).