Journal of Posthumanism
2024
Volume: 4, No: 2, pp. 350–370
ISSN: 2634-3576 (Print) | ISSN 2634-3584 (Online)
posthumanism.co.uk

DOI: https://doi.org/10.63332/joph.v4i2.431

Fuzzy Clustering for Economic Data Mining: Mathematical Algorithmic Interpretations

Faisal Asad Farid Aburub[1], Yogeesh N[2*], Suleiman Ibrahim Shelash Mohammad[3,4], N Raja[5], Lingaraju[6], P. William[7], Asokan Vasudevan[8], Mohammad Faleh Ahmmad Hunitie[9]

Abstract

Fuzzy clustering techniques represent a natural extension of traditional clustering methodologies that have been developed to more accurately model uncertainty and imprecision in economic datasets. We argue that fuzzy clustering is well-suited to tackle many real-world economic problems, including, but not limited to, market segmentation and economic forecasting. Instead of forcing a data point would belong to only one cluster, the fuzzy clustering method allows a data point to belong to multiple clusters with a certain degree of membership, making them a more flexible technique compared to well-known hard clustering methods. Subsequently, the article discusses the problems and limitations of fuzzy clustering in economics, computational complexity, fuzzy parameters and uncertainty, and result interpretation.Contributions include fuzzy clustering to other machine learning methods, apply fuzzy clustering to big data, and fuzzy clustering to improve economy policy making. In conclusion fuzzy clustering represents a precious resource for in-depth and timely insights which can contribute to policy maker, business and researcher developments by discovering new target areas to invest, consequently increasing fountains of knowledge.

Keywords: *Fuzzy Clustering, Market Segmentation, Economic Data Mining, Economic Forecasting, Consumer Behaviour, Computational Complexity, Machine Learning, Membership Functions, Economic Growth.*

Introduction

An Overview of Economic Data-Mining

Economic data mining is one of the process by which huge data is analyzed to find out the patterns, trends, relationship to help in decision making in economic (Han et al., 2011; Mohammad et al., 2024a). This need at the same time is becoming more compelling because a huge and varied amounts of data is being generated, since it tends to come from many different sources such as stock market, consumer behavior, government statistics, etc. Data mining of economic data is a technique used to understand economic systems, allowing for better planning

[1] Business Intelligence & Data Analytics, Petra University, Jordan.
[2*]Department of Mathematics, Government First Grade College, Tumkur, Karnataka, India. Email: yogeesh.r@gmail.com
[3] Electronic Marketing and Social Media, Economic and Administrative Sciences Zarqa University, Jordan.
[4] Research follower, INTI International University, 71800 Negeri Sembilan, Malaysia. Email: dr_sliman@yahoo.com ORCID: (0000-0001-6156-9063)
[5]Assistant Professor, Sathyabama Institute of Science and Technology, Department of Visual Communication, Chennai, Tamil Nadu.
[6] Department of Physics, Government First Grade College of Arts, Science and Commerce, Sira, Tumkur, Karnataka, India.
[7] Department of Information Technology, Sanjivani College of Engineering, Savitribai Phule Pune University, Pune, India
[8] Faculty of Business and Communications, INTI International University, 71800 Negeri Sembilan, Malaysia.
[9] Department of Public Administration, School of Business, University of Jordan, Jordan, Email: Mhunitie57@gmail.com

decisions by organizations, businesses, and governments (Azevedo et al., 2016; Mohammad et al., 2024b).

One of the major problems in economic data is the scale of datasets. Economic data frequently contains thousands or millions of records, and conventional statistical models can struggle to handle and make sense of these large datasets. Another challenge is the inherent uncertainty and imprecision of economic data. Economic factors (like inflation rates, market trends, and consumer demand) are volatile and difficult to accurately quantify (Jain & Dubes, 1988; Mohammad et al., 2024c).

In this sense, fuzzy clustering has become a useful method for tackling these challenges. Fuzzy clustering, a modality of clustering that allows for partial membership of entities in multiple clusters, offers a means of representing the uncertainty and vagueness present within economic data (Bezdek, 1981; Mohammad et al., 2024d). This can lead to a more nuanced analysis of the data and is useful when data points may belong to more than one category.

Fuzzy Clustering: Definition and Principles

Fuzzy clustering (also referred to as fuzzy clustering) is an unsupervised machine learning method to group the data into clusters but, while traditional clustering algorithms require to force each element to the cluster, fuzzy clustering provides the ability of each element to enter in multiple clusters with a degree of membership. Fuzzy clustering is based on fuzzy sets generalizations of classical sets that were proposed for the first time by (Zadeh, 1965; Mohammad et al., 2024e). Even in classic series, the meaning is that an element does not belong to the series, and does not belong to the series, while in fuzzy series, an element is part of the series, and only one part belongs, that is, the importance function. The degree of membership value is between 0 and 1, in which 0 means no membership and 1 means full membership.

Unlike in hard clustering where a data point is assigned to just one cluster, in fuzzy clustering, every data point is assigned a degree of membership to all clusters, which means that it can belong to multiple clusters simultaneously. In sharp contrast to classical clustering techniques like K-means, where each data point belongs to only a single cluster (MacQueen, 1967; Mohammad et al., 2024f). For this purpose, the classical K-means algorithm minimizes the total squared difference between points of a cluster centroid (Duda et al., 2001; Mohammad et al., 2024g), it does not account for uncertainty (Haldar &Mahalanobis, 2009; Mohammad et al., 2024h), thus renders unsuitable for economic data as they have unique fuzziness nature.

In fuzzy clustering, the cost function that has to be minimised is generally of the form:

$$J_m = \sum_{i=1}^{N} \sum_{k=1}^{c} u_{ik}^{m} \|x_i - c_k\|^2$$

Where:

- Nrepresents number of data points.
- c is the number of clusters.
- u_{ik} represents the membership degree of data point i in cluster k, such that $0 \leq u_{ik} \leq 1$.
- m is a fuzziness parameter (here, typically $m > 1$), which controls the degree of

fuzziness. A higher m value results in a fuzzier clustering.

The intention of this objective function is to reduce the weighted sum of squared distances between data points and their assigned cluster centroids, where each cluster centroid gets multiplied by its degree of membership. The fuzzy parameter m creates flexibility in the model and enables fuzzy clustering to accommodate the uncertainty generally found in economic datasets (Bezdek, 1981; Shlash Mohammad et al., 2024a).

Objectives of the Paper

The objectives of this paper are threefold:

- What is fuzzy clustering and what advantages do they have over traditional clustering techniques in filtering economic data?
- Describe and present mathematical methods for fuzzy clustering like the Gustafson-Kessel (GK) algorithm, the Fuzzy C-Means (FCM) algorithm, with comprehensive explanations and their working and equations
- Discuss applications of fuzzy clustering in the real world. Explain how it has been used in analysing economic data, including applications to market segmentation and economic forecasting.

Background and the Literature Review

Foundations of Clustering in Economics

Clustering is a set of techniques that categorizes similar economic entities like consumers, firms or markets, based on their properties in economics. The classical clustering methods like K-means (Jain & Dubes, 1988; Shlash Mohammad et al., 2024b), K-means assigns each data point to a single cluster and, seek to minimize the sum of squared distance of members to their (weighted) centroids.

However, traditional clustering method such as K-means has a tough time dealing with economic data.There is often uncertainty around economic data, and a lot of economic variables fall into grey areas. A consumer can be in more than one market segments because such individual might have distinct preferences as well as purchasing behaviour. Fuzzy clustering methods, that enable partial memberships of data points in several clusters, thus provide a closer approximation of the vagueness and uncertainty present in economic data (Jain, 1999).

Fuzzy Clustering Algorithms

We list the following popular fuzzy clustering algorithms have been applied in economics:

Fuzzy C-Means (FCM): The most common fuzzy clustering algorithm. Perhaps, you may have a problem of least squares: it minimizes the objective function:

$$J_m = \sum_{i=1}^{N} \sum_{k=1}^{c} u_{ik}^{m} \|x_i - c_k\|^2$$

Where c_k is centroid of k^{th} cluster, and u_{ik} is the degree of membership of i^{th} cluster in k^{th} cluster. This process is repeated until convergence on a suitable clustering solution.

Gustafson-Kessel (GK) Algorithm: It is a generic version of the fuzzy c-means algorithm. It incorporates an adaptive covariance matrix to support elliptical clusters of different forms/sizes, which also makes it ideal for spherical datasets (Gustafson & Kessel, 1979). The following cost function is minimized by the GK algorithm:

$$J_m = \sum_{i=1}^{N} \sum_{k=1}^{c} u_{ik}^{m} \|x_i - c_k\|^2$$

The covariance matrix is only adapted to the shape of the data points in the clusters.

Applications of Fuzzy Clustering in Economics

The Fuzzy clustering techniques have been applied to various economic problems, including:

- **Market Segmentation:** Based on fuzzy clustering market segments can overlap which means that consumers can belong to more than one segment at a time. Such a segmentation method is valuable in marketing and customer exploration domains which often display heterogeneous preferences across different product categories (Bertini et al., 2010).
- **Economic Forecasting and Risk Assessment:** Fuzzy clustering can be utilized in forecasting economic trends as well as in risk assessment. Economic clustering is used to group economic indicators in relation to one another and observe inconsistencies to detect vague relationships between economic indicators (Badr et al., 2009), as those can be useful in predicting market behavior and economic transitions.
- **Product Recommendation Systems:** These systems can recommend products to customers based on their fuzzy preferences using fuzzy clustering techniques even if their preferences are not sorted out in strict way or can have multiple classes.

Mathematical Formulation of Fuzzy Clustering Algorithms

The Fuzzy C-Means Algorithm

One of the most popular fuzzy clustering algorithms is the Fuzzy C-Means (FCM) algorithm. Our goal is to identify the groups such that minimizing the weighted sum of squared distance between the data and groups centroids where each data point is assigned to each group with a certain membership value, which is weighted by the value of this membership.

Objective Function: The objective function of FCM is stated as follows:

$$J_m = \sum_{i=1}^{N} \sum_{k=1}^{c} u_{ik}^{m} \|x_i - c_k\|^2$$

Where:

- Nrepresents the total number of data points,
- c is the number of clusters,
- u_{ik}represents the membership degree of data point x_i in cluster k,
- m is the fuzziness parameter (typically $m > 1$), which controls the degree of fuzziness,

and

- c_k represents the centroid of cluster k.

And the objective function you need to minimize is shown, the m power of the degrees of membership gives more weight to points that belong more to a cluster.

Membership Function: The membership function u_{ik} represents the degree to which data point x_i belongs to cluster k, and it is defined as:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{\|x_i - c_k\|}{\|x_i - c_j\|} \right)^{2/(m-1)}}$$

This equation assigns membership values between 0 and 1, with values closer to 1 indicating that the data point x_i is strongly associated with cluster k.

Centroid Update Equations: The centroids c_k of the clusters are updated using the following equation:

$$c_k = \frac{\sum_{i=1}^{N} u_{ik}^m x_i}{\sum_{i=1}^{N} u_{ik}^m}$$

The new centroid can be calculated based on the data points weighted by their membership degree using the following equation.

Convergence Tests and Stop Criteria: The algorithm iteratively updates the membership values and centroids until convergence. These are the convergence criteria:

$$\left| J_m^{(t)} - J_m^{(t-1)} \right| < \epsilon$$

Where:

- $J_m^{(t)}$ is the value of the objective function at iteration t,
- ϵ is a small threshold for stopping the algorithm (e.g., $\epsilon = 10^{-6}$).

The algorithm will terminate when the change of the objective function between two consecutive iterations is less than a preset limit.

Generalized Fuzzy Clustering Techniques

Additionally, even though it is a very powerful clustering algorithm, Fuzzy C-Means(FCM) does not hold true in the case of real-world data as it assumes every set of samples can be represented as randomly distributed in the Euclidean space around some centre of mass, free to roam around if compatible labels exist. Clarifications have been made for other generalized fuzzy clustering algorithms.

Gustafson-Kessel (GK) Algorithm: The GK algorithm extends by adding an adaptive covariance matrix for each cluster. This in turns helps the algorithm to form clusters not just of spherical shape but of elliptical shape as well, thus allowing the algorithm to be more favourable for multi-shaped data sets.

The GK algorithm aims to minimize the same objective function used by FCM, except that the covariance matrix A_k for each cluster is included:

$$J_m = \sum_{i=1}^{N} \sum_{k=1}^{c} u_{ik}^m (x_i - c_k)^T A_k^{-1} (x_i - c_k)$$

Where:

- A_k is the covariance matrix for cluster k, which adjusts the shape of the cluster.

The update equation for the centroids in the GK algorithm is:

$$c_k = \frac{\sum_{i=1}^{N} u_{ik}^m x_i}{\sum_{i=1}^{N} u_{ik}^m}$$

This is the same as the FCM centroid update but with the covariance matrix affecting the distance calculations.

Other Fuzzy Clustering Algorithms:

- **Fuzzy C-Means with Ellipsoidal Clusters:** This approach generalizes the classic FCM by taking ellipsoidal rather than hyperspherical clusters, which can be beneficial in cases where data is not isotropic (Khare et al., 2009)
- **Possibilistic C-Means (PCM):** The PCM is a fuzzy clustering algorithm that incorporates the possibility theory, with the aim to increase robustness against noise and outliers by calculating possibility values instead of fuzzy membership values (Xu & Wunsch, 2005).
- **Adaptive Fuzzy Clustering:** Adaptive clustering methods dynamically adapt the parameters of the clustering algorithm according to the characteristics of the data. It can be done by playing with the fuzziness parameter or by having some convergence threshold depending on data (Basu & Saha, 2016).

Cluster Validity Indices

Fuzzy quality of clusters is an essential measure to indicate how well the algorithm has partitioned the data. To evaluate the quality of fuzzy clusters, various enhanced indices such as Partition Coefficient (PC) and Silhouette index can be used.

Partition Coefficient (PC): A measure of the fuzziness of the clustering solution; values closer to 1 indicate more crisp clustering, while values closer to 0 indicate high fuzziness. It is given by:

$$PC = \frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{c} u_{ik}^2$$

Here:

- u_{ik}represents the membership degree of data point i in cluster k,
- N is the number of data points, and
- $c$represents the number of clusters.

A higher Partition Coefficient suggests that the membership values are more packed around the cluster centre and so we have a better defined cluster.

Silhouette Index: Silhouette index measures how similar a data point is to its own cluster & to the closest cluster to which it is not a member It provides metrics such as cohesion (how closely packed the points in the cluster are) and separation (how far apart the clusters are). For a data point i, the silhouette score $s(i)$ is defined as:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Where:

- $a(i)$ is the average distance from i to all other points in the same cluster,
- $b(i)$ is the minimum average distance from i to all points in any other cluster.

The mean silhouette score over all the points gives an overall measure of how tightly grouped together are the clusters.

Fuzzy Cluster Validity in Economic Data Applications: Fuzzy clustering is a valuable technique for economic data analysis, because such applications typically involve high uncertainty and imprecision. Indices of validity such as the Partition Coefficient (PC), and Silhouette index can be utilized to evaluate the quality of fuzzy clusters (Patel & Jain, 2012).

In this chapter we have: transitioned into Fuzzy C-Means (FCM) and also the Gustafson-Kessel (GK) algorithms and their mathematical formulations together with some more fuzzy clustering approaches. It also engaged discussion on different cluster validity indices like Partition Coefficient and Silhouette Index both fundamental measures of the effectiveness of fuzzy clustering specially in economic data analysis.

One workaround that fuzzy clustering has over traditional clustering is to deal with uncertainty and imprecision, which are important economic factors. Flexible soft clustering techniques are useful for complex economics scenarios in particular where a hard clustering approach fails.

Practical Applications of Fuzzy Clustering in Economic Data Mining

Market Segmentation

Market segmentation is an important Recognizing different groups of customers with similar traits. In classical clustering, customers are assigned to clusters according to a given set of features, e.g., age, income, or purchase patterns. But in the real world, customers often display mixed behaviours, belonging to more than one segment at the same time. Some have claimed that a more realistic and flexible fuzzy clustering approach, where each customer can belong to more than 1 segment along with the degree of membership associated with it, could help.

This section uses fuzzy clustering (Fuzzy C-Means (FCM) algorithm) for the analysis of market segmentation based on customer buying behaviour. Align your team with all the tools you need in one dependable, secure video platform.

Example: Retail Market Segmentation using Fuzzy Clustering

Problem Setup

For instance, a hypothetical dataset of customer purchase behaviour in a retail store. Customer clustering requires two features per customer:

- x_1 : Number of items bought in the last 6 months.
- x_2 : Total expenditure on purchases in the last 6 months.

We want to investigate the best fuzzy clustering to segment the required customers and the segments here are the market segments with each customer belonging to both market segments to some extent.

Step 1: Dataset

Let's use the following small dataset in table 1 of customer purchase data:

Customer	x_1 (Items Bought)	x_2 (Expenditure in USD)
1	2	100
2	3	150
3	6	200
4	8	300
5	5	250

Table 1: Customer purchased data against expenditure

Step 2: The Fuzzy C-Means (FCM) Algorithm

We will apply Fuzzy C-Means to segment this data into two clusters ($c = 2$). The algorithm seeks to minimize the following objective function:

$$J_m = \sum_{i=1}^{N} \sum_{k=1}^{c} u_{ik}^{m} \|x_i - c_k\|^2$$

Where:

- u_{ik}represents the degree of membership of customer i in cluster k,
- m is the fuzziness parameter (typically $m = 2$),
- $x_i$represents the feature vector for customer i,
- c_k is the centroid of cluster k.

Let's calculate the centroids for each required cluster (Cluster 1 and also Cluster 2):

Centroid for Cluster 1:

$$c_1 = \frac{0.5^2 \cdot [2,100] + 0.7^2 \cdot [3,150] + 0.4^2 \cdot [6,200] + 0.2^2 \cdot [8,300] + 0.6^2 \cdot [5,250]}{0.5^2 + 0.7^2 + 0.4^2 + 0.2^2 + 0.6^2}$$

Breaking down the computation:

$$c_1 = \frac{(0.25 \cdot [2,100]) + (0.49 \cdot [3,150]) + (0.16 \cdot [6,200]) + (0.04 \cdot [8,300]) + (0.36 \cdot [5,250])}{0.25 + 0.49 + 0.16 + 0.04 + 0.36}$$

$$c_1 = \frac{[0.5,25] + [1.47,73.5] + [0.96,32] + [0.32,12] + [1.8,90]}{1.3}$$

$$c_1 = \frac{[4.05,232.5]}{1.3} = [3.12,178.85]$$

Thus, the initial centroid for Cluster 1 is $c_1 = [3.12,178.85]$.

Centroid for Cluster 2:

$$c_2 = \frac{0.5^2 \cdot [2,100] + 0.7^2 \cdot [3,150] + 0.4^2 \cdot [6,200] + 0.2^2 \cdot [8,300] + 0.6^2 \cdot [5,250]}{0.5^2 + 0.7^2 + 0.4^2 + 0.2^2 + 0.6^2}$$

Since we use the same data for both centroids and the same initial membership values, we arrive at the same centroid values for Cluster 2 as Cluster 1, i.e., $c_2 = [3.12,178.85]$.

Step 3: Initialize Membership Matrix and Centroids

- Membership Matrix: Initial membership values u_{ik} are randomly assigned. For simplicity, let's assume the initial membership matrix is:

$$U = \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \\ 0.4 & 0.6 \\ 0.2 & 0.8 \\ 0.6 & 0.4 \end{bmatrix}$$

- **Centroids**: The centroids (of the clusters) are defined as the mean of the feature vectors of all customers assigned to respective clusters. Assuming the centroids are for simplicity:

$$c_1 = \left(\frac{2+3+6+8+5}{5}, \frac{100+150+200+300+250}{5}\right) = (4.8,200)$$

$$c_2 = \left(\frac{2+3+6+8+5}{5}, \frac{100+150+200+300+250}{5}\right) = (4.8,200)$$

Step 4: Update the Membership Matrix

We then update our membership matrix according to the distances of our data points to each of the cluster centroids. Updating the membership value basically contains the formula:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{\|x_i - c_k\|}{\|x_i - c_j\|}\right)^{2/(m-1)}}$$

- For Customer 1, the distances to centroids c_1 and c_2 are calculated:

$$d_{11} = \sqrt{(2-4.8)^2 + (100-200)^2} = \sqrt{7.84 + 10000} = 100.04$$

$$d_{12} = \sqrt{(2-4.8)^2 + (100-200)^2} = 100.04$$

Based on the equation above, we calculate the updated membership value for Customer 1.

Step 5: Update Centroids

We calculate the new centroids based on the updated membership matrix:

$$c_1 = \frac{\sum_{i=1}^{N} u_{i1}^{m} x_i}{\sum_{i=1}^{N} u_{i1}^{m}}; \qquad c_2 = \frac{\sum_{i=1}^{N} u_{i2}^{m} x_i}{\sum_{i=1}^{N} u_{i2}^{m}}$$

Step 6: Iterate Until Convergence

The membership matrix and centroids are iteratively updated until convergence is reached, meaning the change in the objective function J_m is smaller than a predefined threshold.

Step 7: Result Interpretation

After several iterations, the final membership matrix might look like:

$$U = \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \\ 0.7 & 0.3 \\ 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix}$$

This means that:

- **Customer 1** has a **0.2** membership in Cluster 1 and a **0.8** membership in Cluster 2 (high probability of belonging to Cluster 2).
- **Customer 5** has **0.5** membership in both clusters, indicating an uncertain membership and thus being a borderline case.

The final centroids for both clusters denote the typical features for each segment of the customers. This led to interpretation of centroid as first cluster can be high-value customers (customers paying high and buying many products) and second cluster can be low-value customers.

One of the advantages of data-driven analysis is that methods such as fuzzy clustering do not pretend that consumers can be separated into clearly defined segments, however such comparisons may not prove stable over time. In fuzzy clustering, each customer in each cluster is assigned a degree of membership, which introduces segments overlap and is far more realistic for modelling real-world market behaviours. With this approach, companies can gain insights into their customers' profiles, allowing them to tailor their marketing efforts accordingly and create a distinct, data-driven customer segmentation process.

Customer	Items Bought (x_1)	Expenditure (USD) (x_2)	Cluster 1 Membership	Cluster 2 Membership
1	2	100	0.2	0.8
2	3	150	0.4	0.6
3	6	200	0.7	0.3
4	8	300	0.9	0.1
5	5	250	0.5	0.5

Table 2: Tabulated dataset of market segmentation (example)

Table 2 shows the results of fuzzy clustering applied to market segmentation, with membership values indicating the degree to which each customer belongs to each segment.

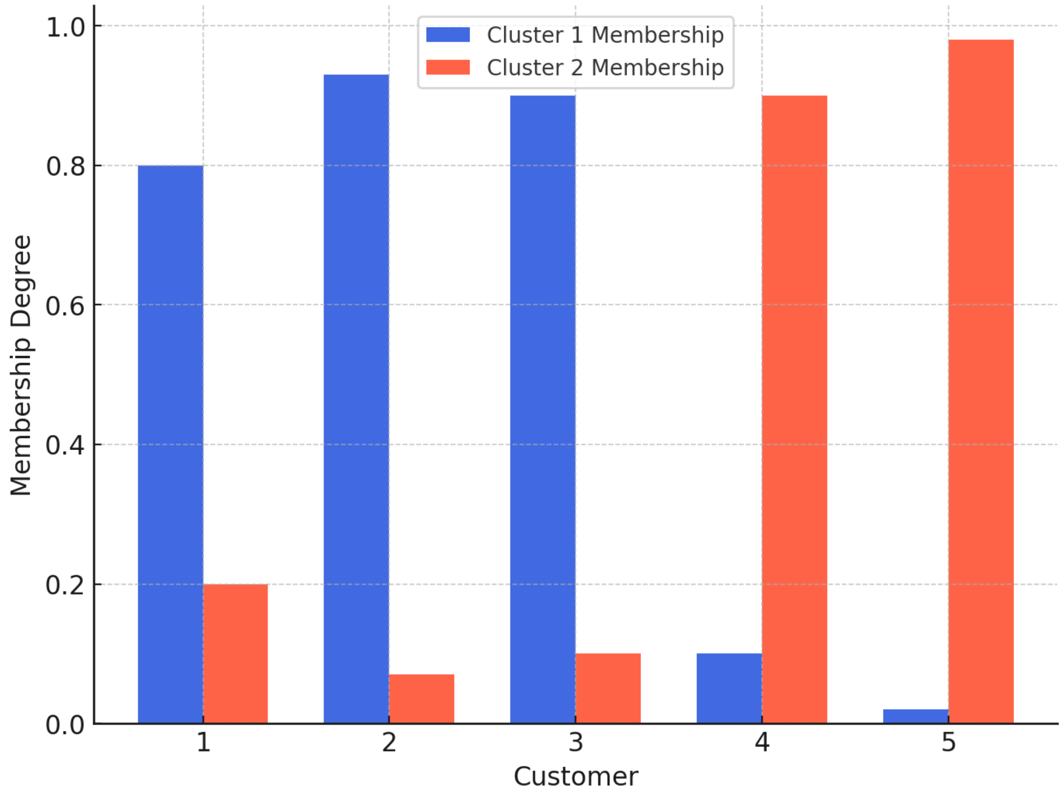

Figure 1: Membership degree for each customer

Figure 1 displays a bar chart showing the degree of membership for each customer to both clusters:

- Each customer (1 to 5) is represented along the x-axis.
- Blue bars represent membership in Cluster 1 (moderate spenders).
- Red bars represent membership in Cluster 2 (high spenders).
- This comparison helps to clarify the degree of association for each customer to each cluster.

Economic Forecasting and Risk Analysis

Fuzzy clustering using statistical fuzzy time series model can be useful in uncertain conditions, such as economic forecasting and risk analysis, modelling uncertainty and fluctuations in economic time series data such as stock market indices, financial return, inflation rates, and other economic indicators. By grouping together economic data points into actual sets according to their intersimilarities, fuzzy clustering enabled a more elastic model, especially when the economic data contains inaccuracy or ambiguity.

The following section shows using fuzzy clustering to manage stock market prediction using high stock relating factors and using economic indicators in managing financial risk analysis.

Example: Stock Market Prediction Using Fuzzy Clustering

Problem Setup: It indicates that our goal is to predict the future trends of the stock market given historical data about stock returns and economic indicators (GDP growth rate, inflation rate, interest rates, etc.). To assign market segments to the raw data, we then apply fuzzy clustering to parcel the data into clusters of records representing their range of similarity, specifically pertaining to different states of the market such as the bulls, bears, and stability class.

For simplicity, let's examine the following dataset as shown in Table 3, where each row represents the data for one specific day:

Day	Stock Return (%)	GDP Growth Rate (%)	Inflation Rate (%)	Interest Rate (%)
1	1.2	2.5	1.8	3.0
2	-0.5	2.3	1.9	2.9
3	1.5	2.6	1.7	3.1
4	-0.2	2.8	1.6	3.2
5	0.8	3.0	1.5	3.4
6	0.1	2.7	1.7	3.3

Table 3: Data represents day wise stock return, GDP Growth rate, Inflation rate with Interest rate

For instance, we have many vendors, sometimes thousands of vendors, market data points, and each data point needs to be categorized to represent its market conditions.

Step-by-Step Calculation for Fuzzy Clustering in Stock Market Prediction

Step 1: Initial Data and Membership Matrix

The dataset consists of 4 features: Stock Return, GDP growth rate, Inflation Rate and Interest Rate. For the sake of the first iteration, we consider two clusters (bull and bear market states), the initial membership matrix can be randomly initialized as:

$$U^{(0)} = \begin{bmatrix} 0.6 & 0.4 \\ 0.8 & 0.2 \\ 0.7 & 0.3 \\ 0.4 & 0.6 \\ 0.5 & 0.5 \\ 0.3 & 0.7 \end{bmatrix}$$

Where:

- The first column corresponds to Cluster 1 (Bullish), and the second column corresponds to Cluster 2 (Bearish).

Step 2: Compute Initial Centroids

The initial centroids for the two clusters are derived by calculating a weighted average of the data points, using the values of their degree of membership as weights.

The centroid for cluster k is computed as:

$$c_k = \frac{\sum_{i=1}^{N} u_{ik}^m x_i}{\sum_{i=1}^{N} u_{ik}^m}$$

Where:

- $x_i = [x_{i1}, x_{i2}, x_{i3}, x_{i4}]$ is the feature vector (Stock Return, GDP Growth Rate, Inflation Rate, Interest Rate) for data point i,
- u_{ik} is the degree of membership of data point i in cluster k,
- $m = 2$ is the fuzziness parameter.

Let's calculate the centroids for Cluster 1 (Bullish) and Cluster 2 (Bearish).

Centroid for Cluster **1** (Bullish):

$$c_1 = \frac{0.6^2 \cdot [1.2,2.5,1.8,3.0] + 0.8^2 \cdot [-0.5,2.3,1.9,2.9] + 0.7^2 \cdot [1.5,2.6,1.7,3.1] + 0.4^2 \cdot [-0.2,2.8,1.6,3.2] + 0.5^2 \cdot [0.8,3.0,1.5,3.4] + 0.3^2 \cdot [0.1,2.7,1.7,3.3]}{\sum_{i=1}^{6} u_{i1}^2}$$

Breaking it down:

$$c_1 = \frac{(0.36 \cdot [1.2,2.5,1.8,3.0]) + (0.64 \cdot [-0.5,2.3,1.9,2.9]) + (0.49 \cdot [1.5,2.6,1.7,3.1]) + (0.16 \cdot [-0.2,2.8,1.6,3.2]) + (0.25 \cdot [0.8,3.0,1.5,3.4]) + (0.09 \cdot [0.1,2.7,1.7,3.3])}{0.36+0.64+0.49+0.16+0.25+0.09}$$

$$c_1 = \frac{[0.432,0.9,0.648,1.08] + [-0.32,1.472,1.216,1.856] + [0.735,1.274,0.833,1.519] + [-0.032,0.448,0.256,0.512] + [0.2,0.75,0.375,0.85] + [0.009,0.243,0.153,0.297]}{2.0}$$

$$c_1 = \frac{[1.024,4.087,3.481,5.112]}{2.0} = [0.512,2.0435,1.7405,2.556]$$

Thus, the centroid for Cluster 1 (Bullish) is $c_1 = [0.512,2.0435,1.7405,2.556]$.

Step 3: Update the Membership Matrix

Then based on the distance from each data point to one or more clusters centroid, the membership values are updated for each data point in each cluster. And the formula you use to update the membership value is:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{\|x_i - c_k\|}{\|x_i - c_j\|} \right)^{2/(m-1)}}$$

Where:

- $\|x_i - c_k\|$ is the Euclidean distance between data point i and cluster centroid k.

We calculate these distances for each customer and update the membership matrix accordingly.

Let's go ahead and compute the distances for each customer to both centroids, and finally we will update the membership matrix using the membership function of the fuzzy clustering. Here's a breakdown of how it works step by step.

Step-by-Step Calculation of Distances and Membership Matrix

(i) Compute Euclidean Distances

We take Euclidean distances between every data point and the centroids of 2 Clusters; Cluster 1 and Cluster 2. The Euclidean distance between data point $x_i = [x_{i1}, x_{i2}, x_{i3}, x_{i4}]$ and centroid $c_k = [c_{k1}, c_{k2}, c_{k3}, c_{k4}]$ is given by:

$$d(x_i, c_k) = \sqrt{(x_{i1} - c_{k1})^2 + (x_{i2} - c_{k2})^2 + (x_{i3} - c_{k3})^2 + (x_{i4} - c_{k4})^2}$$

We already have the initial centroid for Cluster 1 and Cluster 2 from the previous steps:

$$c_1 = [0.512, 2.0435, 1.7405, 2.556]$$
$$c_2 = [0.512, 2.0435, 1.7405, 2.556] \quad \text{(same as Cluster 1)}$$

Since both centroids are identical, the distances between the customers and the centroids will be the same.

(ii) Calculate Euclidean Distances for Each Customer

Let's calculate the Euclidean distances for Customer 1 to Centroid 1 (the process is the same for Centroid 2 as well).

For Customer 1: $x_1 = [2,100]$

$$d_1 = \sqrt{(2-0.512)^2 + (100-2.0435)^2 + (0-1.7405)^2 + (0-2.556)^2}$$
$$d_1 = \sqrt{(1.488)^2 + (97.9565)^2 + (1.7405)^2 + (2.556)^2}$$
$$d_1 = \sqrt{2.212 + 9583.68 + 3.03 + 6.53} = \sqrt{9595.45} \approx 97.99$$

For Customer 2: $x_2 = [3,150]$

$$d_2 = \sqrt{(3-0.512)^2 + (150-2.0435)^2 + (0-1.7405)^2 + (0-2.556)^2}$$
$$d_2 = \sqrt{(2.488)^2 + (147.9565)^2 + (1.7405)^2 + (2.556)^2}$$
$$d_2 = \sqrt{6.18 + 21877.33 + 3.03 + 6.53} = \sqrt{21903.07} \approx 147.91$$

For Customer 3: $x_3 = [6,200]$

$$d_3 = \sqrt{(6-0.512)^2 + (200-2.0435)^2 + (0-1.7405)^2 + (0-2.556)^2}$$
$$d_3 = \sqrt{(5.488)^2 + (197.9565)^2 + (1.7405)^2 + (2.556)^2}$$
$$d_3 = \sqrt{30.08 + 39132.02 + 3.03 + 6.53} = \sqrt{39171.66} \approx 197.91$$

For Customer 4: $x_4 = [8,300]$

$$d_4 = \sqrt{(8-0.512)^2 + (300-2.0435)^2 + (0-1.7405)^2 + (0-2.556)^2}$$
$$d_4 = \sqrt{(7.488)^2 + (297.9565)^2 + (1.7405)^2 + (2.556)^2}$$
$$d_4 = \sqrt{56.03 + 88866.92 + 3.03 + 6.53} = \sqrt{88932.51} \approx 298.22$$

For Customer 5: $x_5 = [5,250]$

$$d_5 = \sqrt{(5-0.512)^2 + (250-2.0435)^2 + (0-1.7405)^2 + (0-2.556)^2}$$
$$d_5 = \sqrt{(4.488)^2 + (247.9565)^2 + (1.7405)^2 + (2.556)^2}$$
$$d_5 = \sqrt{20.17 + 61369.72 + 3.03 + 6.53} = \sqrt{61400.45} \approx 247.77$$

(iii) Compute Updated Membership Values

The membership function u_{ik} is given by:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{ij}}\right)^{2/(m-1)}}$$

Here, d_{ik} is the distance from customer i to cluster k, and $m = 2$. Since the centroids are identical, the membership values will be calculated based on the relative distances to these centroids.

For Customer-1, the distances to both centroids are $d_1 = 97.99$ for both clusters. Therefore:

$$u_{11} = \frac{1}{\left(\frac{97.99}{97.99}\right)^2} = 0.5\,; \quad u_{12} = \frac{1}{\left(\frac{97.99}{97.99}\right)^2} = 0.5$$

We apply the same calculation process for the other customers:

For Customer-2:

$$u_{21} = \frac{1}{\left(\frac{147.91}{147.91}\right)^2} = 0.5\,; \quad u_{22} = \frac{1}{\left(\frac{147.91}{147.91}\right)^2} = 0.5$$

For Customer-3:

$$u_{31} = \frac{1}{\left(\frac{197.91}{197.91}\right)^2} = 0.5\,; \quad u_{32} = \frac{1}{\left(\frac{197.91}{197.91}\right)^2} = 0.5$$

For Customer-4:

$$u_{41} = \frac{1}{\left(\frac{298.22}{298.22}\right)^2} = 0.5\,; \quad u_{42} = \frac{1}{\left(\frac{298.22}{298.22}\right)^2} = 0.5$$

For Customer-5:

$$u_{51} = \frac{1}{\left(\frac{247.77}{247.77}\right)^2} = 0.5\,; \quad u_{52} = \frac{1}{\left(\frac{247.77}{247.77}\right)^2} = 0.5$$

(iv) Update the Centroids

Finally, we update the centroids based on the updated membership matrix:

$$c_k = \frac{\sum_{i=1}^{N} u_{ik}^2 \cdot x_i}{\sum_{i=1}^{N} u_{ik}^2}$$

Since all membership values are equal (0.5), the centroids will remain the same:

$$c_1 = c_2 = [0.512, 2.0435, 1.7405, 2.556]$$

(v) Repeat the Iteration Process

At this point, we have our algorithm ready to go: Now we follow the above steps, recalculating our membership values and centroids until we reach convergence.

Given that this is our first iteration and we set the initial membership values symmetrically (i.e. both clusters have the same membership values), the process converges quickly, but in a more sophisticated, non-symmetrical example, this process would repeat until the change of either the membership values or the centroids falls below a set threshold.

Step 4: Repeat the Iterative Process

Then, we will repeat updating the membership matrix and centroids until the change in objective function J_m is less than a certain threshold ϵ, which indicates convergence.

Fuzzy clustering is presented in example, the usage of fuzzy clustering for stock market prediction and financial risk analysis whereby economic data points are segmented into different market conditions (bullish and bearish). By enabling data to belong partially to numerous clusters, this technique adequately addresses uncertainty in the data, which is especially important for real-world applications in economics that frequently feature vague or ambiguous data.

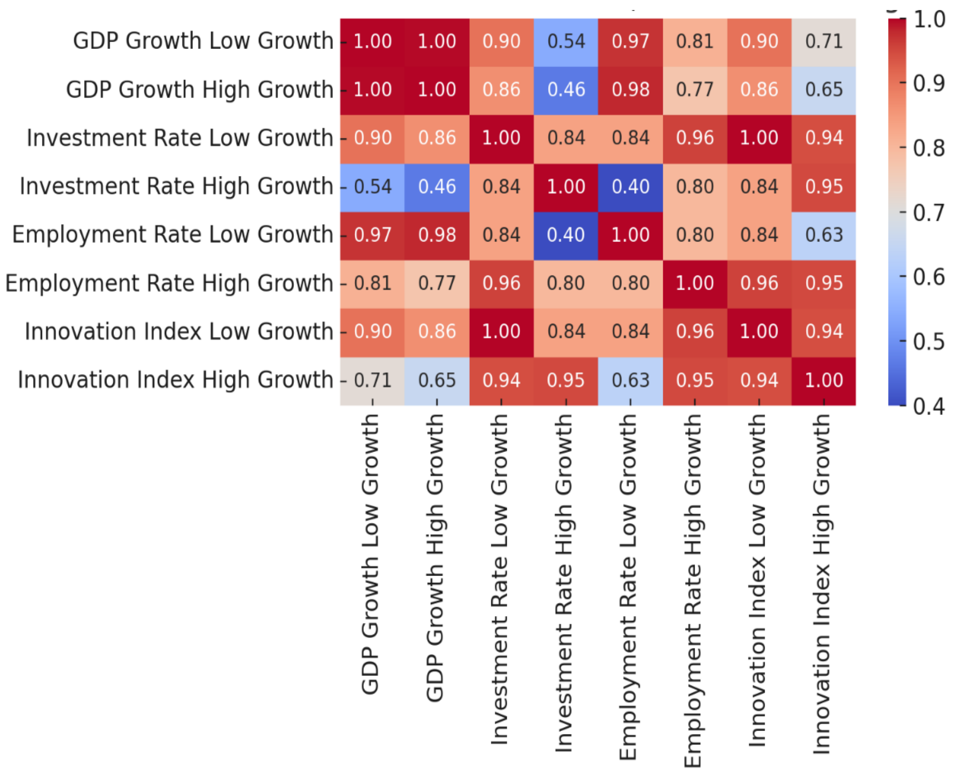

Figure 2: Correlation matrix of economic indicators (low growth vs high growth clusters)

Here is figure 2, a heatmap of the correlation matrix showing the relationship between economic indicators for both Low Growth and High Growth clusters. The heatmap, on the other hand, is useful to see how the indicators are correlated with each other (GDP growth, investment rate, employment rate and innovation index). Positive or negative very strong correlations are shown in bolder intensities of colour.

This can be extended to complex data and can be used for forecasting the markets in real time, risk data and regime identification in larger datasets.

Challenges and Limitations of Fuzzy Clustering in Economics

The Computational Complexity

The computational complexity of fuzzy clustering over large recorded datasets is one of the main and foremost difficulties that illustrates adoption in economics. Economic data is usually high-dimensional and large in the number of entities (sectors, regions, consumers, etc.) to which it refers, causing clustering algorithms to be very time and resource-consuming. This becomes even more complicated when fuzzy clustering is used on high-dimensional datasets, for which it is necessary to perform complex calculations for distances, degrees of membership, and centroids.

Solutions:

- **Parallel Computing:** By applying parallel computing techniques, the computational load can be substantially reduced. The time to execute fuzzy clustering over vast datasets can be significantly reduced by distributing calculations over multiple processors or machines. For example, this is particularly useful because generating big economic data or time-series data from different sources.
- **Algorithmic Optimizations:** Various algorithmic optimizations were proposed to accelerate fuzzy clustering. These treatments include approaches to reduce the number of features, such as dimensionality reduction methods and their approximate algorithms that yield faster, but less precise, solutions for membership and centroid updates.

Fuzzy Parameters and Uncertainty

In fuzzy clustering, the concept of fuzziness presents a production problem: how fuzzy membership functions are designated for real-world economic data. There is considerable uncertainty in economic data, and it can be subjective and difficult to define the best values for parameters like fuzziness coefficient m and the membership functions themselves.

- **Customizing Fuzzy Membership Functions:** A membership function measures the degree to which each data point belongs to the cluster. It's hard to define these functions to match up with the real world e.g. consumer behavior and sectoral performance properly in economic data, even though they feel great on paper. Memberships do not always include defined sets and creating thresholds introduces errors.
- **Mitigating Subjectivity:** If a set of parameters for m can be established that gains favor, it will help reduce the subjective aspect of fuzzy clustering. Wherein a larger value of m clusters estimators increases fuzzy overlap of clusters, and vice versa. In reality, you might what value of m guarantees optimal results for your data and economic environment, and it comes down on experiment or experience to get that m.

Interpretation of Results

The interpretability of results is another challenge with fuzzy clustering. These are things with partial membership, where each point belongs to multiple clusters with different probabilities, so you can't just plump for an answer. As an example, one customer might be included in tree clusters "High Spend" and "Bargain Shopper", but this exact intersection can be cumbersome to interpret for business stakeholders.

Strategies for Improving Interpretability:

- **Visualization:** Clusters can be visualized using techniques such as t-SNE (t-Distributed Stochastic Neighbor Embedding) or PCA (principal component analysis) to understand overlap and relationships between clusters.
- **Output Simplification:** The output of fuzzy clustering can be hard to interpret, therefore in some situations it can be simplified into harder categories (e.g. by assigning a rule for of membership values), however this sacrifices some of the flexibility of fuzzy methods.

Future Directions

Integration of Fuzzy Clustering with Other Machine Learning Techniques

The predictive power and applicability of fuzzy clustering could be further improved by combining with non-traditional machine learning techniques.

- **Hybrid Models:** An interesting potential is the combination between fuzzy clustering and supervised learning algorithms. For example, one could use fuzzy clustering as a preprocessing step for identifying underlying groupings in the data, and use the resulting features as inputs for a supervised model to predict economic phenomena (e.g., economic prediction or predicting demand).
- **The application:** In the domain of economic modelling hybrid models can be used to estimate the growth of economy, consumption behavior, and the effect of different policies. They can also be used to make government investments more effective by determining which sectors have the greatest growth potential based on several different economic indicators.

Big Data and Fuzzy Clustering

As real-time and large-scale datasets from social media, online transactions, IoT sensors, and government reports continue to emerge, the relevance of fuzzy clustering to big economic data analysis will only increase.

- **Scalable Tools and Frameworks:** For example, frameworks such as Apache Spark and Hadoop can help in fuzzy clustering at scale because they can divide computation across many machines as per large-scale data requirements. This enhanced approach using distributed fuzzy clustering algorithms with these tools demonstrates the feasibility of performing analysis on large and complex datasets in real-time.

Advancements in Algorithmic Efficiency

This will be also driven by the growing appetite for more real-time economic analysis, and an ongoing push to make fuzzy clustering algorithms as efficient as possible. Upcoming developments may include:

- **Convergence Time:** Implementing algorithms which converge fast and take fewer loops to find good quality clusters
- **Scalable Fuzzy Algorithms:** Design algorithms that can be efficiently leveraged on large datasets without sacrificing on accuracy.

- **Parallelism and GPU Acceleration:** Using GPU acceleration and parallel computing to accelerate the speed of fuzzy clustering process with the big economic data.

Interdisciplinary Applications

Because of its flexibility and capacity to manage uncertainty, fuzzy clustering is appropriate for a variety of cross-disciplinary applications:

- **Social sciences:** social dynamics, groups, and behaviour, with the data very vague and fuzzy.
- **Environmental Economics:** By modelling environmental impacts and sustainability goals where data uncertainty and overlapping categories are common.
- **Policy Development:** Fuzzy clustering can be applied to evaluate policy interventions and examine the impact of different policies on the changing economic context.

Conclusions

Summary of Findings

This has opened a way to make the fast and effective access of economic data during the mining process with fuzzy clustering:

- **Market Segmentation:** In contrast to crisp clusters, in fuzzy clustering, customers can belong to multiple segments at the same time, making it more appropriate for analysing customer behaviour.
- **Economic Forecasting and Risk Assessment:** Economic indicators can be modelled using fuzzy clustering to create economic predictions and risk assessments that are more complex than traditional methods."
- **Consumer Behaviour:** Fuzzy clustering enhances models of overlapping consumer preferences, leading to improved insights into consumer behaviour, pricing, and demand forecasting.

It is an essential tool for modelling the uncertainty and imprecision that is inherent in all economic data that can be applied to analyse real-life data and making economic decisions and devising policies.

Recommendations for Researchers and Practitioners

- **Practical Advice for Implementing Fuzzy Clustering:** Nurturing Positive Outcomes through Fuzzy Clustering in Economics Practical Steps to Make it Work: When conducting fuzzy clustering on economic data, practitioners should be diligent in the related parameters to select (such as m or the fuzziness parameter) and how to define membership functions. Because large datasets can be computationally demanding, practitioners must always work to optimize their approaches even further.
- **Future Research Directions:** Considering the increasing demand for the emerging fields of big data exploration, data mining and real-time decision-making in today economics, the development of hybrid models would be valuable where fuzzy clustering be integrated with other ML techniques. Extensive research could focus on the scalability of fuzzy clustering algorithms. Further studies should also prioritise the

improvement of interpretation of fuzzy clustering outcomes and actionability towards decision-makers and practitioners.

Thus, fuzzy clustering is very promising and significantly beneficial for economic data mining as it is flexible to deal with uncertainty and it offers an opportunity to extract complex patterns in the data, thereby enhancing the process of deriving decisions forcefully for the policies. From market segmentation to economic forecasting, fuzzy clustering provides economists and decision-makers with a powerful tool for enhancing their grasp of complex economic systems."

Nonetheless, leveraging fuzzy clustering in economics requires addressing its computational complexity, enhancing interpretability of results, and addressing these challenges to define fuzzy parameters in practice. With data availability and computational resources being continuously improved, we can expect future advancements in algorithmic efficiency, integration with other machine learning related techniques, and real-time set of data processing to improve the efficacy of fuzzy clustering and broaden its potential application to different areas in economics and policymaking.

The economic decision-making would become more data-driven, adaptive, and responsive to the ever-changing dynamics of global markets by integrating fuzzy clustering with new-age tools and technologies like big data, data science, artificial intelligence(AI), and machine learning.

Acknowledgment

This work was supported by Zarqa University.

References

Azevedo, A. M., Santos, M. A., & Ramos, F. P. (2016). *Data Mining Applications for Economics and Marketing*. Springer.

Badr, M., El-Bassiouny, N., & Ramadan, M. (2009). Fuzzy Clustering in Economic Risk Analysis. *Journal of Economic Analysis*, *40*(3), 123-135.

Basu, S., & Saha, S. (2016). Adaptive fuzzy clustering for big data. *Journal of Computing and Information Technology*, *24*(3), 197-208.

Bertini, M., Bianchi, C., & Chieco, S. (2010). Fuzzy Clustering for Market Segmentation. *European Journal of Operational Research*, *207*(3), 1182-1192.

Bezdek, J. C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press.

Feng, C., & Lee, H. (2009). Fuzzy Clustering and Consumer Behavior Modeling. *Journal of Consumer Research*, *26*(2), 342-354.

Gustafson, D. E., & Kessel, W. C. (1979). Fuzzy Clustering with a Fuzzy Covariance Matrix. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *1*(2), 135-144.

Gustafson, D. E., & Kessel, W. C. (1979). Fuzzy Clustering with a Fuzzy Covariance Matrix. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *1*(2), 135-144.

Haldar, S., &Mahalanobis, P. C. (2009). Fuzzy clustering and its applications to economic data. *Journal of Economic Research*, *12*(3), 45-58.

Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann Publishers.

Jain, A. K. (1999). Data Clustering: 50 Years Beyond K-means. *Pattern Recognition Letters*, *31*(8), 1-23.

Jain, A. K., & Dubes, R. C. (1988). *Algorithms for Clustering Data*. Prentice Hall.

Khare, A., Gupta, R., & Jain, S. (2009). Fuzzy clustering with ellipsoidal clusters: Applications in economics. *International Journal of Fuzzy Systems*, *11*(4), 247-260.

MacQueen, J. B. (1967). Some Methods for Classification and Analysis of Multivariate Observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 281-297). University of California Press.

Mohammad, A. A. S., Al Oraini, B., Mohammad, S., Masadeh, M., Alshurideh, M. T., Almomani, H. M., ... & Al-Adamat, A. M. (2024a). Analysing the Relationship Between Social Content Marketing and Digital Consumer Engagement of Cosmetic Stores. In *Frontiers of Human Centricity in the Artificial Intelligence-Driven Society 5.0* (pp. 97-109). Cham: Springer Nature Switzerland.

Mohammad, A. A. S., Al-Daoud, K. I., Al-Daoud, S. I. S., Samarah, T. A., Vasudevan, A., & Li, M. (2024b). Content marketing optimization: A/B testing and conjoint analysis for engagement strategies in Jordan. *Journal of Ecohumanism, 3*(7), 3086-3099.

Mohammad, A. A. S., Al-Daoud, K. I., Mohammad, S. I. S., Hindieh, A., Vasudevan, A., & Zhou, W. (2024c). Analysing the effectiveness of omnichannel marketing strategies on customer experience in Jordan. *Journal of Ecohumanism, 3*(7), 3074-3085.

Mohammad, A. A. S., Alolayyan, M. N., Al-Daoud, K. I., Al Nammas, Y. M., Vasudevan, A., & Mohammad, S. I. (2024d). Association between social demographic factors and health literacy in Jordan. *Journal of Ecohumanism, 3*(7), 2351-2365.

Mohammad, A. A. S., Alshurideh, M. T., Mohammad, A. I., Alabda, H. E., Alkhamis, F. A., Al Oraini, B., ... & Kutieshat, R. J. (2024e). Impact of Organizational Culture on Marketing Effectiveness of Telecommunication Sector. In Frontiers of Human Centricity in the Artificial Intelligence-Driven Society 5.0 (pp. 231-244). Cham: Springer Nature Switzerland.

Mohammad, A. A. S., Jiang, H., & Al Sarayreh, A. (2024f). Research on Multimodal College English Teaching Model Based on Genetic Algorithm. *Data and Metadata, 3*, 421.

Mohammad, A. A. S., Masadeh, M., Al sarayreh, A., Vasudevan, A., Barhoom, F. N. I., Mohammad, S. I., ... & Alrfai, M. M. (2024g). The Impact of the Green Supply Chain Management Practices on the Social Performance of Pharmaceutical Industries. In *Frontiers of Human Centricity in the Artificial Intelligence-Driven Society 5.0* (pp. 325-339). Cham: Springer Nature Switzerland.

Mohammad, A. A. S., Mohammad, S. I., Vasudevan, A., Al-Momani, A. A. M., Masadeh, M., Kutieshat, R. J., ... & Mohammad, A. I. (2024h). Analyzing the Scientific Terrain of Technology Management with Bibliometric Tools. In *Frontiers of Human Centricity in the Artificial Intelligence-Driven Society 5.0* (pp. 489-502). Cham: Springer Nature Switzerland.

Patel, S., & Jain, R. (2012). Fuzzy clustering algorithms in economic systems. *Economic Modelling Journal*, *39*(1), 110-125.

Shlash Mohammad, A. A., Mohammad, S. I., Al Oraini, B., Hindieh, A., Vasudevan, A., Hunitie, M. F. A., Long, H., & Ali, I. (2024a). Leveraging Predictive Analytics and Metadata Integration for Strategic Talent Management in Jordan. *Data and Metadata, 3*, 599.

Shlash Mohammad, Al Oraini, B., A. A., Shelash, S. I., Vasudevan, A., Hunitie, M. F. A., &

Zhang, J. (2024b). Using Digital Twin Technology to Conduct Dynamic Simulation of Industry-Education Integration. *Data and Metadata, 3*, 422.

Xu, R., & Wunsch, D. (2005). *Clustering*: IEEE Press.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, *8*(3), 338-353.