HÃY NHẬP TIÊU ĐỀ TIẾNG ANH VÀO ĐÂY ...

Hãy nhập tên của (các) tác giả vào đây ...

Hãy nhập đơn vị công tác của (các) tác giả vào đây ...

\*Corresponding author: authors@gmail.com

(Received: xxx xx, 2023; Accepted: xxx xx, 2023)

**Abstract -** ……………………………..

**Key words -** ………………………….

# Introduction

Custom segmentation is a process of classifying potential customers based on their mutual characteristics such as shopping habits, consumption trends, and demand. It can assist businesses in better understanding the behavior of customers and implementing various marketing campaigns for specific markets to increase sales and customer satisfaction [1]. The explosion of information technology, especially the development of Industry Revolution 4.0 is currently changing how businesses approach market segmentation. Digital versions of numerous products are now available. A new set of customer behaviors, identities, and expectations have been produced because of the two-way information flow between customers and providers that technology has enabled. Businesses must modify and adjust their segments to take into account new information on actual customer behavior as it changes over time.

# Related works

## Fuzzy C-Means Algorithm

FCM is a popular clustering method that uses fuzzy logic in data partitioning [24]. The fuzzy concept allows a data instance to belong to a cluster based on its membership degree from 0 to 1. A dataset } contains *n* data instances. FCM algorithm classifies *X* into *c* clusters to minimize the following objective function:

(1)

## Density Peak Fuzzy C-Means Algorithm

DPFCM algorithm is a combination of density peak clustering (DPC) approach and FCM algorithms. The concept behind the DPC is that the cluster centers are surrounded by neighbors who have lower local densities [25]. The DPC approach determines two values: the local density of each data point *i* and the distance between that point and the point with the highest density [26]. Combined with FCM algorithm, the DPFCM has several advantages such as effectively solving the problem of sensitivity to the initial centers, and improving the clustering accuracy [18]. The procedure of DPFCM is described as follows:

**Stage 1**: Identify the initial cluster centers.

## Forensic-Based Investigation Algorithm

The FBI algorithm proposed by Chou and Nguyen [23] Diagram

Description automatically generated

**Figure 1.** General procedure of the FBI algorithm [23]

# Proposed DP-FBI-FCM algorithm

The proposed DP-FBI-FCM algorithm, which combines DPFCM with FBI algorithms, aims to not only overcome the drawback of sensitivity of initial cluster centers in most of the clustering methods but also can explore the global optimal solution and improve the clustering accuracy. The procedure of DP-FBI-FCM algorithm is described in Figure 2.

# Result *analysis*

## Dataset and parameter setting

To illustrate the performance of the proposed DP-FBI-FCM algorithm in custom segmentation, wholesale customer data collected in UCI machine learning repository is used for analysis [27]. This dataset presents the yearly spending in monetary units (m.u.) on several product categories. There are eight features with a total of 440 data instances in this dataset. These features describe the annual spending on fresh products, milk products, groceries, frozen foods, detergent and paper products, delicatessen products, retail channels, and regions.

## Identify the number of clusters (k)

The number of clusters is a predetermined parameter to implement a clustering algorithm. This study employs one of the most common methods to identify the number of clusters, i.e., the elbow method. The FCM algorithm is used in the elbow method by implementing the dataset with different values of *k* and relying on the SSE (sum of squared errors) value to evaluate the clustering results and select the optimal *k*. Figure 3 shows the SSE values of the dataset implemented by the FCM algorithm with *k* selected from 2 to 20. According to the Elbow method, the optimal *k* is selected at 4.

## Clustering result

The section evaluates the clustering result of the proposed DP-FBI-FCM algorithm. Its result is compared with several benchmark algorithms such as *k*-means, FCM, and DPFCM algorithms. All algorithms are coded in Python and implemented in a Windows 10 computer, Intel Core i5 processor with 16 GB. Each algorithm was run 20 times with its set-up parameter. Then, the average value is presented for comparison.

**Table 1.** The comparison of clustering result

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Index** | **k-means** | **FCM** | **DPFCM** | **DP-FBI-FCM** |
| *Si* | 0.605 | 0.538 | 0.643 | 0.821 |
| DBI | 0.954 | 1.063 | 0.906 | 0.817 |

## Custom segmentation analysis

According to the clustering result of the proposed algorithm, the wholesale customer dataset is grouped into 4 clusters. The features of each cluster are presented in Figure 4.

# Conclusion

The FBI algorithm is first proposed in this research for market segmentation in a business organization. The FBI algorithm, DPC, and FCM algorithm are then combined to get more accurate segmentation results. To evaluate the performance of the proposed DP-FBI-FCM algorithm, a comparison with the benchmark algorithms, such as *k*-means, FCM, and DPFCM algorithms, was made to precisely assess how well the proposed method performed. The clustering result validated in terms of Silhouette and DBI indices shows that the proposed DP-FBI-FCM algorithm outperforms the benchmark algorithm to cluster the wholesale customer data. Each customer group is then analyzed based on its distinct characteristics to support a specific marketing strategy.

# Acknowledgments: This research is funded by …..

References

1. A. Ahani, M. Nilashi, O. Ibrahim, L. Sanzogni, and S. Weaven, "Market segmentation and travel choice prediction in Spa hotels through TripAdvisor’s online reviews”, *International Journal of Hospitality Management,* vol. 80, 2019, pp. 52-77.
2. H. Müller and U. Hamm, "Stability of market segmentation with cluster analysis–A methodological approach”, *Food Quality and Preference,* vol. 34, 2014, pp. 70-78.
3. H. Zhou *et al.*, "Market segmentation approach to investigate existing and potential aviation markets”, *Transport Policy,* vol. 99, 2020, pp. 120-135.
4. L. T. Nduna and C. van Zyl, "A benefit segmentation framework for a nature-based tourism destination: the case of Kruger, Panorama and Lowveld areas in Mpumalanga Province”, *International Journal of Tourism Cities,* vol. 6 (4), 2020, pp. 953-973.