ABSTRACT

This research develops a novel approach for learning group formation within the Jigsaw Cooperative Learning model in higher education settings, aiming to address challenges related to group homogeneity and lack of personalization. Utilizing the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology, the study begins with business and data understanding phases, where student data—including identity, academic performance (quiz scores and GPA), personality (MBTI), and learning styles—were collected via Microsoft Forms questionnaires. In the data preparation stage, data cleaning was performed to handle missing values, and features were encoded and standardized using Z-score normalization. Dimensionality reduction with Principal Component Analysis (PCA) was employed to simplify the data structure, with the first two principal components chosen for their ability to capture most of the variance. For modeling, the K-means clustering algorithm was implemented. The optimal number of clusters was determined to be 4 using the Elbow method. The clustering results were visualized in two dimensions using PCA, demonstrating a reasonably clear separation of clusters. The evaluation of clustering quality using the Silhouette Score yielded a value of 0.36, indicating that the cluster structure formed is adequately representative for separating data into distinct groups, despite some data points being close to cluster boundaries. Based on these clustering results, heterogeneous groups were formed by sequentially drawing one member from each cluster, resulting in 7 groups with 4-5 members each. This process aims to enhance collaboration and material comprehension, while providing a more equitable and personalized learning experience.

*Keywords—**Jigsaw**, **K-means**, **clustering**, **Cooperative Learning**, **CRISP-DM**.*