

## ABSTRACT

*In this digital era, data is at the core of a business. The need for data practitioners is increasing in today's industry. The data itself can be useful information if data is done. Humans are faced with an abundance of data that can be obtained without having the ability to extract information in it. By using data science and machine learning algorithms, companies can process data and provide more value for business strategy purposes. One example of the results of data processing in business is by making customer segmentation, customer segmentation is useful for recognizing and filtering customers with the specified data. Analysis of segmentation is able to provide more target market allocation, more effective marketing or promotion strategies, and much more. Because segmentation aims to separate customers into several groups or clusters, clustering can be used..*

*In this final project, credit card customer data will be grouped using the K-Means algorithm, to determine the optimal number of clusters, 2 metrics are used, namely the elbow method and the silhouette score to provide a curve that shows the optimal number of clusters to be used in the data. Principal Component Analysis (PCA) is a metric that will be used in this Final Project to reduce the dimensions of the data into 2 principal components that have the highest variance value, PCA 1 with 28% and PCA 2 with 23%, if added up it becomes 51% of 100% data.*

*After preprocessing data and features selection, from 18 features and 8950 records contained in credit card customer data, only 14 features will be used for the modeling process, the elbow and silhouette score method gives the result that the 14 features used produce the optimal number of clusters. is 2 clusters, the results of the grouping will be made in the visualization of histograms and scatter plots to facilitate the analysis of each cluster.*

**Keywords:** *Data Science, Machine Learning, Clustering, K-Means.*