

ABSTRAK

Semakin ketatnya persaingan dalam dunia bisnis mendorong perusahaan perusahaan untuk intensif dalam meningkatkan performa karyawan mereka. Hal ini dipicu oleh peran krusial teknologi dalam mengoptimalkan efisiensi waktu kerja dan akurasi pengolahan data. Oleh karena itu, banyak perusahaan yang beralih dari sistem manual ke sistem digital guna mengatasi lonjakan data perusahaan yang terus meningkat setiap tahun. Tantangan utama terletak pada pengelolaan bank data, terutama dalam bagian penjualan, yang jika tidak ditangani dengan baik dapat berdampak fatal bagi kelangsungan perusahaan.

Penelitian ini memiliki tujuan untuk menerapkan algoritma *forgy initialization* dan *Kmeans++* untuk inisialisasi awal centroid dalam metode *K-Means Clustering*, serta melakukan evaluasi jumlah cluster yang paling optimal pada *clustering* dengan menggunakan metode *silhouette coefficient*. Implementasi metode yang dilakukan pada penelitian diharapkan mampu meningkatkan tingkat akurasi pada hasil *clustering*.

Berdasarkan hasil penelitian yang dilakukan, penerapan algoritma *Kmeans++* dalam metode *K-Means Clustering* dengan 2 (dua) cluster mendapatkan nilai *silhouette coefficient* paling tinggi sebesar 0.9232095. Hasil tersebut menunjukkan pengaruh dari penerapan *Kmeans++* sebagai metode inisialisasi yang membuat hasil *clustering* menjadi lebih optimal.

Penelitian ini dapat dijadikan sebagai acuan untuk pemilihan metode inisialisasi awal cluster dalam *K-Means Clustering* sehingga hasil yang didapatkan lebih optimal sehingga resiko terhadap kesalahan strategi pemasaran dapat diminimalisir.

Kata Kunci : *Forgy Initialization, Kmeans++, K-means Clustering, Silhouette coefficient*

ABSTRACT

The increasing competition in the business world has driven companies to intensively improve their employees' performance. This is triggered by the crucial role of technology in optimizing work efficiency and data processing accuracy. Therefore, many companies have shifted from manual systems to digital systems in order to manage the growing volume of company data that continues to increase every year. The main challenge lies in managing the data bank, particularly in the sales department, which, if not handled properly, could have a fatal impact on the company's continuity.

This research has a goals to implementing the forgy initialization and Kmeans++ algorithm for first centroid initialization in K-Means Clustering method, and evaluate the optimal number of cluster in the clustering process using silhouette coefficient method. The implementation of these methods in this study is expected to improve the accuracy of the clustering results.

Based on the results of the research, the application of the Kmeans++ algorithm in the K-Means Clustering method with 2 (two) clusters has the highest silhouette coefficient value of 0.9232095. this result indicates the implementing of Kmeans++ as an initialization method leads to more optimal clustering results.

Keywords : *Forgy Initialization, Kmeans++,K-means Clustering, Silhouette coefficient*