

Abstrak

Dengan semakin berkembangnya teknologi, maka kemampuan manusia pun semakin berkembang dalam mengumpulkan dan mengolah data secara terkomputerisasi yang mengakibatkan semakin banyak data yang berbentuk digital. Data yang tersimpan pasti memiliki informasi yang beraneka ragam. Untuk memudahkan pengambilan informasi dari data tersebut diperlukan adanya pengelompokan data secara otomatis. *Clustering* adalah proses untuk mengelompokkan data ke dalam suatu klaster, sehingga objek pada suatu klaster memiliki kemiripan yang sangat besar dengan objek lain pada klaster yang sama, tetapi sangat tidak mirip dengan objek pada klaster yang lain. Untuk melakukan pengelompokan tersebut, digunakan salah satu algoritma *clustering* yaitu *Jarvis-Patrick clustering*. *Jarvis-Patrick clustering* merupakan metode *clustering* berdasarkan kesamaan antara tetangga terdekat. Satu atau lebih tetangga pada umumnya adalah digunakan untuk menilai anggota *cluster* dari objek yang diteliti. *Cluster* yang terbentuk tergantung dari nilai parameter J dan K. Untuk mengukur similarity antar dokumen sebelum proses *clustering* digunakan *Euclidean distance*.

Pada tugas akhir ini kualitas *cluster* yang dihasilkan diukur menggunakan parameter *silhouette coefficient*. Berdasarkan percobaan yang dilakukan bahwa rata-rata nilai *silhouette coefficient* adalah 0.220. Dengan begitu kualitas *cluster* yang dihasilkan bersifat *no structure*, maksudnya hampir mustahil menemukan pusat *cluster* yang signifikan dan jelas menetapkan mayoritas titik data.

Kata kunci: *Jarvis Patrick, Euclidean distance, Silhouette Coefficient, Clustering*