

ABSTRAK

Brian Adytia Hutapea, NIM 4201230011 (2020). PENGELOMPOKAN PASIEN YANG MEMILIKI FAKTOR PENYEBAB PENYAKIT JANTUNG DENGAN MENGGUNAKAN METODE PRINCIPAL COMPONENT ANALYSIS (PCA) DAN K NEAREST NEIGHBORS (KNN)

Tujuan penelitian ini adalah mengelompokkan pasien dengan faktor penyebab penyakit jantung menggunakan metode Principal Component Analysis (PCA) dan K-Nearest Neighbors (KNN). Metode penelitian yang digunakan adalah penelitian kuantitatif dengan data pasien rawat inap yang bertempat di Rumah Sakit Umum (RSU) Haji pada rentang tahun 2021-2023. Metode PCA digunakan untuk mereduksi dimensi data, sehingga hanya faktor-faktor yang paling signifikan yang digunakan dalam analisis. Dalam penelitian ini dilakukan melalui beberapa tahapan penting, seperti uji *KMO*, *Bartlett's Test*, serta eigenvalue dan eigenvektor. Berdasarkan hasil nilai eigenvektor di dapatkan 3 faktor yang paling signifikan dari 7 faktor yang menyebabkan penyakit jantung. Setelah hasil analisis PCA di dapatkan selanjutnya digunakan metode KNN untuk mengelompokkan pasien berdasarkan faktor-faktor yang telah direduksi. Beberapa langkah penting yang harus dilakukan dalam metode KNN tersebut seperti partisi data, perhitungan jarak Euclidean, serta melihat akurasi dari data yang kita uji. Berdasarkan hasil akurasi di dapat nilai 90% pada pengujian data x_{test} nilai akurasi tersebut menjelaskan bahwa pendeteksian dini dengan mengelompokkan pasien yang memiliki faktor penyebab penyakit berhasil di deteksi.

Kata kunci : Statistika, Pengelompokan, Reduksi, PCA, KNN.