

Perbandingan Penerapan *Feature Selection* Melalui Pendekatan *Embedded* dalam Klasifikasi Penyakit Diabetes Melitus Menggunakan Metode *K-Nearest Neighbor*

Imam Nurul Ihsan¹, Widi Astuti, S.T., M. Kom.², Prof. Dr. Adiwijaya, S.Si., M.Si.³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹imamnihsan@student.telkomuniversity.ac.id, ²widiwdu@telkomuniversity.ac.id,

³adiwijaya@telkomuniversity.ac.id

Abstract

Diabetes is a degenerative disease that continues to increase every year. According to the International Diabetes Federation, it is predicted that by 2045 people with this disease will increase by 700 million people in the world. With the increasing number of cases of people with diabetes, the patient data will increase and there is no cheap system for diagnosing this disease. Through the application of data mining, we can classify and find information from very large data in a short time at a relatively low cost. The data classification method applied is k-Nearest Neighbor by comparing two different scenarios, namely data that is applied through feature selection through an embedded approach with data that is not applied to feature selection. The results of this study, based on the data applied to the feature selection process, got the best accuracy value of 89.36% at the value of $k=9$. This proves that the feature selection process can improve accuracy for diagnosing diabetes mellitus.

Keywords: diabetes melitus, *data mining*, *k-nearest neighbor*, *feature selection*
