

ABSTRAK

Diabetes melitus merupakan salah satu penyakit tidak menular dengan prevalensi yang terus meningkat dan menjadi ancaman serius bagi kesehatan masyarakat. Deteksi yang umumnya dilakukan setelah muncul gejala klinis menyebabkan banyak kasus terlambat ditangani sehingga meningkatkan risiko komplikasi. Kondisi ini mendorong perlunya sistem deteksi dini yang mampu mengklasifikasikan risiko diabetes berdasarkan kombinasi faktor klinis dan gaya hidup. Permasalahan dalam penelitian ini adalah bagaimana membangun sistem klasifikasi yang mampu menangani ketidakpastian data medis sekaligus memanfaatkan pola historis dari data populasi dalam satu mekanisme yang terintegrasi.

Penelitian ini mengembangkan *Hybrid Expert System* menggunakan integrasi Logika Fuzzy Mamdani dan algoritma K-Nearest Neighbor (KNN) untuk mengklasifikasikan risiko diabetes ke dalam kategori Risiko Rendah, Risiko Sedang, dan Risiko Tinggi. Dataset yang digunakan berasal dari *National Health and Nutrition Examination Survey* (NHANES) dengan total 4.024 data hasil pembersihan dari 9.813 data awal. Sistem dirancang menggunakan arsitektur paralel yang memisahkan parameter klinis ke dalam modul Fuzzy Mamdani Klinis dan parameter gaya hidup ke dalam modul Fuzzy Mamdani Lifestyle. Setiap modul menerapkan proses fuzzifikasi, inferensi aturan, agregasi, dan defuzzifikasi metode Centroid. Secara paralel, algoritma KNN dilatih menggunakan seluruh fitur dengan pemilihan nilai K optimal melalui 5-fold cross-validation. Untuk menangani ketidakseimbangan kelas pada dataset, diterapkan teknik SMOTE pada data training. Seluruh keluaran sistem kemudian digabungkan menggunakan mekanisme *Weighted Decision Fusion* dengan bobot 30% untuk komponen klinis, 50% untuk komponen gaya hidup, dan 20% untuk probabilitas KNN berdasarkan validasi pakar medis.

Hasil penelitian menunjukkan bahwa model KNN dengan nilai $K=3$ menghasilkan nilai F1-macro sebesar 0,8668 pada tahap validasi. Pada data uji, model memperoleh akurasi sebesar 75,28% dengan nilai ROC-AUC sebesar 0,6669 serta recall sebesar 0,45 pada kelas berisiko. Hasil validasi pakar juga menunjukkan bahwa variabel, fungsi keanggotaan, aturan fuzzy, dan mekanisme penggabungan keputusan yang digunakan telah sesuai dengan pendekatan klinis dalam penilaian risiko diabetes. Penelitian ini berkontribusi dalam menyediakan sistem skrining dini diabetes yang lebih transparan, interpretatif, dan relevan secara klinis sebagai pendukung proses pemeriksaan medis konvensional.

Kata Kunci: Sistem Pakar Hibrida, Fuzzy Mamdani, K-Nearest Neighbor, Klasifikasi Risiko Diabetes, Decision Fusion

ABSTRACT

Diabetes mellitus is one of the non-communicable diseases with a continuously increasing prevalence and has become a serious public health concern. Detection is generally performed after clinical symptoms appear, causing many cases to be identified too late and increasing the risk of complications. This condition highlights the need for an early detection system capable of classifying diabetes risk based on a combination of clinical and lifestyle factors. The main problem addressed in this study is how to develop a classification system that can handle uncertainty in medical data while simultaneously utilizing historical population patterns within an integrated framework.

This study develops a Hybrid Expert System by integrating Mamdani Fuzzy Logic and the K-Nearest Neighbor (KNN) algorithm to classify diabetes risk into Low Risk, Moderate Risk, and High Risk categories. The dataset used was obtained from the National Health and Nutrition Examination Survey (NHANES), consisting of 4,024 cleaned records from an initial 9,813 data entries. The system was designed using a parallel architecture that separates clinical parameters into a Clinical Mamdani Fuzzy module and lifestyle parameters into a Lifestyle Mamdani Fuzzy module. Each module implements fuzzification, rule inference, aggregation, and Centroid defuzzification. In parallel, the KNN algorithm was trained using all features with optimal K selection through 5-fold cross-validation. To address class imbalance, the SMOTE technique was applied to the training data. The outputs from all modules were then combined using a Weighted Decision Fusion mechanism with weights of 30% for the clinical component, 50% for the lifestyle component, and 20% for KNN probability based on medical expert validation.

The results showed that the KNN model with $K=3$ achieved an F1-macro score of 0.8668 during validation. On the testing data, the model obtained an accuracy of 75.28%, an ROC-AUC score of 0.6669, and a recall value of 0.45 for the at-risk class. Expert validation also confirmed that the variables, membership functions, fuzzy rules, and decision fusion mechanism used in the system were consistent with clinical approaches to diabetes risk assessment. This study contributes by providing an early diabetes screening system that is more transparent, interpretable, and clinically relevant as a supporting tool for conventional medical examinations.

Keywords: *Hybrid Expert System, Mamdani Fuzzy, K-Nearest Neighbor, Diabetes Risk Classification, Decision Fusion*