

SCHIZOPHRENIA CLASSIFICATION USING FUZZY K-NEAREST NEIGHBOR ON PATIENT DATA FROM RSJD Dr. AMINO GONDOHUTOMO

KLASIFIKASI SKIZOFRENIA MENGGUNAKAN FUZZY K-NEAREST NEIGHBOR PADA DATA PASIEN RSJD Dr. AMINO GONDOHUTOMO

Ozagastra Caluella Prambudi¹, Ajib Susanto², Christy Atika Sari³

^{1,2,3} Universitas Dian Nuswantoro, Jl. Imam Bonjol No.207, Semarang, Indonesia
111202113451@mhs.dinus.ac.id¹, ajib.susanto@dsn.dinus.ac.id², atika.sari@dsn.dinus.ac.id³

Abstract - Schizophrenia is a complex mental disorder with overlapping symptoms, making subtype diagnosis uncertain. This study aims to develop an automated classification method for schizophrenia subtypes using the Fuzzy K-Nearest Neighbour (FKNN) algorithm, which effectively handles uncertainty in medical data. The dataset includes 300 patients from RSJD Dr. Amino Gondohutomo, Central Java, aged 18–60 years, with balanced gender distribution. Four subtypes—paranoid, catatonic, hebephrenic, and undifferentiated—were classified. Symptom and demographic data were encoded and normalised using min-max scaling. The model was trained using $k = 5$ and evaluated via 10-fold cross-validation. The results achieved 94% accuracy with high precision and recall across all classes. However, limitations include a relatively small and single-source dataset and the lack of ROC/AUC analysis. These findings suggest that FKNN has strong potential as a data-driven decision support system for schizophrenia diagnosis, suitable for integration into psychiatric hospital information systems. Future research should explore oversampling techniques such as SMOTE and threshold tuning to improve model sensitivity.

Keywords: Schizophrenia, Classification, Diagnosis, Fuzzy K-Nearest Neighbour.

Abstrak - Skizofrenia merupakan gangguan mental kompleks dengan gejala yang tumpang tindih, sehingga diagnosis jenisnya sering mengalami ketidakpastian. Penelitian ini bertujuan untuk mengembangkan metode klasifikasi otomatis jenis skizofrenia menggunakan algoritma Fuzzy K-Nearest Neighbor (FKNN), yang mampu menangani data medis yang tidak pasti. Dataset terdiri dari 300 pasien di RSJD Dr. Amino Gondohutomo Jawa Tengah, dengan rentang usia 18–60 tahun dan distribusi jenis kelamin seimbang. Empat jenis skizofrenia diklasifikasikan: paranoid, katatonik, hebefrenik, dan undifferentiated. Data gejala dan demografis diproses menggunakan encoding dan normalisasi min-max. Model dilatih dengan parameter $k = 5$ dan divalidasi menggunakan teknik 10-fold cross-validation. Hasil menunjukkan akurasi sebesar 94%, dengan precision dan recall tinggi pada seluruh kelas. Namun, keterbatasan penelitian ini terletak pada ukuran data yang relatif kecil dan bersumber tunggal, serta belum dilakukan analisis ROC/AUC. Temuan ini menunjukkan potensi penerapan FKNN sebagai sistem pendukung keputusan diagnosis berbasis data, yang dapat diintegrasikan ke dalam sistem informasi rumah sakit jiwa. Penelitian selanjutnya disarankan untuk mengeksplorasi teknik oversampling seperti SMOTE dan tuning threshold untuk meningkatkan sensitivitas model.

Kata Kunci: Skizofrenia, Klasifikasi, Diagnosis, Fuzzy K-Nearest Neighbor.

I. PENDAHULUAN

Skizofrenia merupakan gangguan otak kronis yang memengaruhi kemampuan seseorang untuk berpikir secara jernih [1]. Penderita biasanya menunjukkan ketidaksesuaian antara pola pikir dan perilakunya. Istilah skizofrenia berasal dari bahasa Yunani *schizein* (memisahkan) dan *phren* (pikiran), yang menggambarkan kondisi perpecahan antara pikiran dan emosi [2]. Gangguan ini termasuk dalam kategori masalah kesehatan mental yang serius dan ditandai dengan gejala seperti berbicara sendiri, rasa curiga berlebihan, delusi, halusinasi, pola pikir yang tidak teratur, serta perubahan perilaku yang signifikan [3]. Skizofrenia umumnya muncul pada awal masa dewasa, tetapi dapat terjadi pada berbagai rentang usia. Risiko tertinggi untuk didiagnosis terjadi pada usia 16 hingga 25 tahun.

Penyakit ini dikelompokkan berdasarkan gejala-gejala utama yang ditunjukkan oleh pasien. Skizofrenia dibagi ke dalam beberapa jenis, antara lain paranoid, hebefrenik, katatonik, dan undifferentiated. Proses klasifikasi ini penting karena kesulitan dalam menentukan jenis gangguan secara akurat sangat memengaruhi penanganan medis yang tepat. Kompleksitas gejala yang tumpang tindih antar jenis skizofrenia menjadikan diagnosis sebagai proses yang menantang dan rawan kesalahan [5].

Salah satu pendekatan yang dapat digunakan untuk mengatasi masalah ini adalah algoritma Fuzzy K-Nearest Neighbor (FKNN), sebuah teknik pembelajaran mesin yang memanfaatkan informasi kedekatan antara data baru dan data yang telah dikelompokkan sebelumnya [6]. FKNN dirancang untuk menangani ketidakpastian dan ketidakjelasan dalam proses diagnosis skizofrenia, dengan memberikan bobot keanggotaan terhadap tiap kelas berdasarkan kedekatan fitur. Algoritma ini memiliki keunggulan dalam mengurangi pengaruh ketidakpastian terhadap hasil klasifikasi [6]. Penelitian ini bertujuan untuk mengembangkan metode klasifikasi otomatis jenis skizofrenia menggunakan algoritma Fuzzy K-Nearest Neighbor (FKNN), berdasarkan data gejala dan demografi pasien, guna meningkatkan akurasi dan objektivitas diagnosis dalam lingkungan klinis.

II. SIGNIFIKASI STUDI

A. Studi Literatur

Penelitian mengenai skizofrenia sebagai salah satu gangguan mental berat telah menjadi perhatian dalam berbagai bidang, baik medis maupun teknologi informasi. Skizofrenia merupakan gangguan otak kronis yang ditandai oleh gejala seperti halusinasi, delusi, emosi datar, dan disorganisasi perilaku. Diagnosis skizofrenia pada praktik klinis masih menghadapi tantangan besar karena gejala antar jenisnya sering tumpang tindih dan bersifat subjektif. Oleh karena itu, dibutuhkan pendekatan yang dapat membantu proses klasifikasi diagnosis secara lebih objektif dan berbasis data.

Beberapa penelitian terdahulu telah mengusulkan berbagai pendekatan berbasis sistem pakar dan logika klasifikasi untuk membantu proses diagnosis. Penelitian Priskilla Puji Larasati (2019) mengembangkan sistem pakar berbasis metode Certainty Factor untuk menentukan jenis skizofrenia secara otomatis. Sistem ini menunjukkan hasil diagnosis yang akurat dan sesuai dengan hasil perhitungan manual [7]. Penelitian lain oleh Khoirun Nisa dan Sony Kartika Wibisono (2023) menggunakan pendekatan algoritma Logistic Regression untuk klasifikasi skizofrenia, namun mengalami kesulitan dalam membedakan data pasien sehat dengan pasien skizofrenia karena bias prediksi [8].

Sementara itu, pendekatan menggunakan logika fuzzy mulai diterapkan dalam kasus diagnosis penyakit yang melibatkan ketidakpastian data. Hardyman Zalfi et al. mengaplikasikan metode FKNN untuk diagnosis penyakit pada hewan, dan menunjukkan bahwa FKNN mampu menangani variasi data yang tidak terstruktur dengan baik [9]. Di bidang pertanian, Wildan Gita Akbari et al. juga menggunakan FKNN dalam diagnosis penyakit tanaman cabai dan berhasil memperoleh hasil klasifikasi yang cukup akurat melalui proses fuzzifikasi dan defuzzifikasi [10]. Penelitian serupa di bidang neurologi oleh Siburian et al. menunjukkan bahwa FKNN dapat meningkatkan performa diagnosis pada penyakit berbasis gejala kompleks [11].

Selain itu, studi oleh Paramita dan Setyani menegaskan bahwa diagnosis berbasis algoritma sangat dibutuhkan dalam gangguan mental karena masih banyak dilakukan secara manual dan rentan terhadap kesalahan [12]. Algoritma FKNN dianggap cocok karena dapat menghitung nilai keanggotaan suatu data terhadap seluruh kelas yang ada secara proporsional, bukan hanya klasifikasi biner [13]. Dalam penerapannya, FKNN juga telah terbukti bekerja dengan baik pada data yang memiliki ketidakseimbangan kelas dengan sedikit pelatihan parameter [14]. Sebuah studi oleh Silaban (2020) juga menyoroti pentingnya sistem pendukung keputusan otomatis dalam diagnosis gangguan jiwa karena banyak penderita enggan memeriksakan diri secara langsung [15].

Salah satu metode yang mulai digunakan dalam kasus klasifikasi dengan data tidak pasti adalah FKNN. Metode ini merupakan pengembangan dari algoritma KNN, yang memperhitungkan tingkat keanggotaan (fuzzy membership) suatu data terhadap kelas-kelas yang ada, bukan hanya mengambil keputusan secara biner. Pendekatan ini dinilai sesuai untuk kasus medis, karena gejala yang muncul sering kali tidak mutlak masuk dalam satu kategori. Beberapa penelitian sebelumnya telah menunjukkan bahwa Fuzzy KNN cocok digunakan pada data diagnosis yang kompleks dan sulit dibedakan.

Kontribusi utama dari penelitian ini adalah pengembangan sistem klasifikasi jenis skizofrenia menggunakan metode FKNN yang belum banyak diterapkan dalam konteks diagnosis gangguan mental di rumah sakit jiwa. Sistem ini dirancang untuk memberikan hasil diagnosis yang lebih objektif dan akurat berdasarkan data klinis aktual dari pasien. Berbeda dengan penelitian sebelumnya yang lebih bersifat eksperimental dan terbatas pada validasi teknis, penelitian ini juga mengusulkan rencana implementasi sistem klasifikasi FKNN dalam bentuk prototipe sistem pendukung keputusan (decision support system) yang dapat diintegrasikan ke dalam sistem informasi rumah sakit. Implementasi ini akan difokuskan di RSJD Dr. Amino Gondohutomo, sehingga kontribusi nyata dari penelitian ini tidak hanya terbatas pada pengembangan model klasifikasi, tetapi juga pada kesiapan penerapan teknologi dalam praktik klinis untuk mendukung proses diagnosis tenaga medis secara langsung.

Sistem yang diusulkan diharapkan dapat membantu tenaga medis dalam mempercepat proses identifikasi jenis skizofrenia, meminimalkan bias subjektif, serta meningkatkan efisiensi layanan kesehatan mental. Dengan sistem ini, proses diagnosis di RSJD dapat menjadi lebih konsisten, data-driven, dan dapat digunakan sebagai dasar pertimbangan klinis untuk pengambilan keputusan lanjutan dalam terapi atau perawatan pasien.

B. Data Penelitian

Penelitian ini menggunakan data sekunder yang diperoleh dari RSJD Dr. Amino Gondohutomo Provinsi Jawa Tengah. Rumah sakit ini dipilih karena merupakan salah satu fasilitas kesehatan jiwa rujukan di wilayah Jawa Tengah yang memiliki basis data pasien skizofrenia yang luas, lengkap, dan terpercaya. Lokasi ini relevan digunakan dalam penelitian karena menyediakan data medis yang memadai untuk kebutuhan analisis klasifikasi jenis skizofrenia.

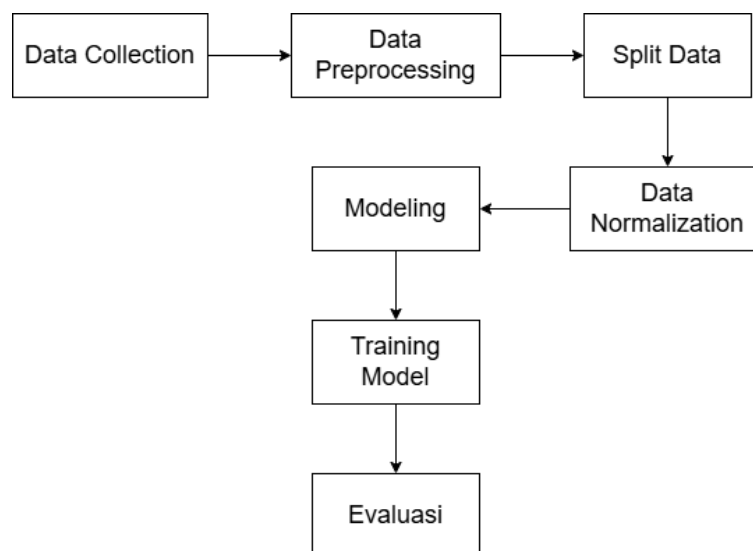
Data yang digunakan dalam penelitian ini meliputi informasi demografis pasien, gejala klinis yang dialami, serta diagnosis jenis skizofrenia. Data diperoleh dalam bentuk rekam medis yang telah disesuaikan untuk kepentingan penelitian dan dikumpulkan secara anonim untuk menjaga kerahasiaan identitas pasien. Penggunaan data dari RSJD Dr. Amino Gondohutomo Provinsi Jawa Tengah dilakukan dengan izin dan tetap mengedepankan etika penelitian serta privasi pasien.

Hasil dari penelitian ini nantinya akan diimplementasikan dalam bentuk prototipe sistem pendukung keputusan yang dapat diintegrasikan ke dalam sistem informasi rumah sakit. Dengan demikian, tenaga medis di RSJD dapat memanfaatkan sistem ini sebagai alat bantu dalam proses diagnosis jenis skizofrenia secara cepat dan akurat. Implementasi ini diharapkan dapat meningkatkan efisiensi kerja klinis, mengurangi beban tenaga medis, serta memperbaiki kualitas layanan kesehatan jiwa di rumah sakit. Pemilihan lokasi ini diharapkan dapat memberikan kualitas data yang valid untuk mendukung pembangunan sistem klasifikasi berbasis metode Fuzzy K-Nearest Neighbor.

C. Metode Penelitian

1. Metode

Penelitian ini dirancang untuk mengklasifikasikan jenis skizofrenia berdasarkan gejala pasien menggunakan metode FKNN. Proses klasifikasi dilakukan secara bertahap dan sistematis, mulai dari pengumpulan data hingga evaluasi model. Tahapan yang digunakan dalam penelitian ini terdiri dari: pengumpulan data, preprocessing data, pembagian data (split data), normalisasi data numerik, pembuatan model, pelatihan model (training), dan evaluasi hasil klasifikasi. Setiap tahap memiliki alur metode sebagai berikut:



Gambar 1. Diagram Alur Metode Klasifikasi Skizofrenia Menggunakan Fuzzy K-Nearest Neighbor

Metode penelitian ini terdiri dari beberapa tahapan dimulai dengan pengumpulan data, yaitu data pasien skizofrenia yang mencakup informasi demografis seperti usia dan jenis kelamin, serta data gejala yang diamati. Selanjutnya, data melalui tahap preprocessing, yang meliputi konversi data kategorikal menggunakan label encoding, penanganan nilai kosong, dan penyesuaian tipe data agar sesuai format input model. Setelah data bersih, dilakukan split data menjadi data latih 70% dan data uji 30 % secara stratifikasi agar distribusi kelas tetap seimbang. Tahap berikutnya adalah normalisasi fitur numerik menggunakan metode StandardScaler, yang bertujuan untuk menyamakan skala setiap fitur sehingga perhitungan jarak menjadi akurat. Setelah proses normalisasi, dilakukan pembuatan model klasifikasi dengan menerapkan algoritma Fuzzy KNN. Model ini kemudian dilatih menggunakan data latih pada tahap training model. Setelah model selesai dilatih, proses selanjutnya adalah evaluasi menggunakan berbagai metrik seperti akurasi,

classification report, confusion matrix, serta analisis keanggotaan fuzzy untuk mengukur seberapa yakin model terhadap setiap prediksi kelas yang dihasilkan. Evaluasi ini penting untuk mengetahui seberapa efektif model dalam mengenali masing-masing jenis skizofrenia secara akurat.

2. Ilustrasi Dataset

Untuk memberikan gambaran lebih jelas mengenai struktur data yang digunakan dalam penelitian ini, berikut ditampilkan representasi visual dari data pasien skizofrenia:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
jenis_kelamin	usia	berdiam_diri_seperti_patum	berperilaku_aneh	bicara_kacau	bicara_kasar	bicara_sendiri	curiga	delusi	ingin_bunuh_diri	halusinasi	marah_marah	tertawa_sendiri	tidak_mau_bersosialisasi	diagnosis
0	51	0	1	1	1	0	1	1	1	1	1	0	0	paranoid
1	47	0	0	0	1	0	1	1	0	1	1	1	0	paranoid
0	53	0	0	0	1	0	1	1	0	1	1	1	1	paranoid
0	59	0	0	1	1	0	1	1	1	1	1	1	0	paranoid
0	58	0	0	0	1	0	1	1	0	1	1	1	1	paranoid
0	60	0	0	1	1	0	1	1	0	1	1	0	0	paranoid
0	34	0	0	1	1	0	1	1	0	1	1	0	0	paranoid
1	45	0	0	0	1	1	1	1	0	1	1	0	0	paranoid
0	44	0	0	0	1	1	1	1	0	1	1	0	0	paranoid
0	27	0	0	0	1	1	1	1	0	1	1	0	0	paranoid
0	29	0	0	0	1	0	1	1	1	1	1	0	0	paranoid
1	62	0	0	0	1	1	1	1	0	1	1	1	0	paranoid
0	21	0	0	0	1	1	1	1	0	1	1	0	0	paranoid
0	41	0	0	1	1	0	1	1	1	1	1	0	0	paranoid
0	58	0	0	1	1	1	1	1	0	1	1	0	0	paranoid
0	30	0	0	1	1	0	1	1	0	1	1	0	0	paranoid

Gambar 2 Tampilan Dataset Pasien Skizofrenia

Dataset yang digunakan dalam penelitian ini berisi informasi pasien dengan jenis skizofrenia, yaitu paranoid, hebefrenik, katatonik, dan undifferentiated. Setiap baris pada dataset mewakili satu pasien dengan sejumlah fitur yang menunjukkan kondisi klinis dan demografis. Fitur gejala meliputi halusinasi, delusi, emosi datar, bicara kacau, tertawa sendiri, serta tidak mau bersosialisasi. Sedangkan data demografis mencakup jenis kelamin yang dikodekan sebagai angka (0 atau 1) dan usia dalam satuan tahun. Kolom terakhir pada dataset adalah diagnosis, yang menunjukkan label jenis skizofrenia sesuai hasil pemeriksaan medis. Dataset ini digunakan sebagai data latih dan data uji dalam pengembangan model klasifikasi FKNN.

3. Data Preprocessing

Data preprocessing merupakan tahap awal yang sangat penting untuk memastikan bahwa data siap digunakan dalam proses klasifikasi menggunakan algoritma FKNN. Proses ini diawali dengan melakukan label encoding pada fitur kategorikal seperti jenis kelamin dan diagnosis, agar dapat diubah menjadi bentuk numerik. Hal ini diperlukan karena FKNN hanya dapat bekerja dengan data dalam format angka. Selain itu, dilakukan juga pemeriksaan terhadap nilai kosong (missing value) serta pembersihan data untuk menjaga kualitas input dan mencegah kesalahan dalam pemrosesan lebih lanjut. Setelah data dibersihkan, dilakukan pembagian dataset menjadi dua bagian, yaitu data latih sebesar 70% dan data uji sebesar 30%, menggunakan metode stratified train-test split. Teknik ini digunakan untuk menjaga keseimbangan distribusi jenis skizofrenia pada kedua subset data. Dengan pembagian ini, model dapat belajar dari data yang mencerminkan kondisi nyata setiap jenis secara proporsional dan mampu dievaluasi secara objektif. Tahapan preprocessing ini berperan besar dalam meningkatkan kualitas data yang digunakan, sehingga model dapat bekerja secara lebih optimal dan menghasilkan klasifikasi yang akurat.

4. Data Normalization

Data normalization merupakan langkah penting dalam tahap pemrosesan data untuk memastikan bahwa setiap fitur memiliki kontribusi yang setara dalam proses perhitungan jarak antar data. Pada penelitian ini, proses normalisasi dilakukan menggunakan metode StandardScaler, yaitu sebuah teknik transformasi yang mengubah nilai fitur numerik ke dalam distribusi standar dengan rata-rata (mean) 0 dan standar deviasi 1. Proses ini diterapkan pada semua fitur numerik dalam dataset,

seperti usia dan berbagai gejala yang telah dikodekan secara biner (0 dan 1). Tujuan dari normalisasi ini adalah agar setiap fitur memiliki skala yang setara, sehingga tidak ada fitur yang mendominasi dalam proses pembelajaran model. Penerapan normalisasi sangat krusial dalam metode FKNN karena algoritma ini melakukan klasifikasi berdasarkan jarak Euclidean antara data uji dan data latih. Jika skala fitur tidak diseragamkan, fitur dengan nilai besar seperti usia dapat memberi pengaruh lebih besar terhadap perhitungan jarak, sehingga mengganggu akurasi hasil klasifikasi. Dengan normalisasi, setiap fitur diberikan bobot yang adil dalam proses perhitungan jarak, yang pada akhirnya meningkatkan stabilitas, akurasi, dan keandalan model dalam mengklasifikasikan jenis skizofrenia secara konsisten

5. Fuzzy K-Nearest Neighbor (FKNN)

Fuzzy K-Nearest Neighbor (FKNN) adalah algoritma klasifikasi yang merupakan pengembangan dari metode K-Nearest Neighbor (K-NN) dengan mengadopsi prinsip logika fuzzy. Berbeda dengan K-NN tradisional yang hanya menentukan kelas mayoritas dari tetangga terdekat, FKNN memberikan nilai derajat keanggotaan suatu data terhadap setiap kelas. Nilai keanggotaan tersebut kemudian digunakan untuk menentukan hasil klasifikasi berdasarkan kelas dengan derajat keanggotaan tertinggi. Pendekatan ini dinilai lebih fleksibel dalam menangani ketidakpastian dan tumpang tindih antar kelas, terutama dalam domain medis seperti diagnosis skizofrenia. Proses klasifikasi dengan metode FKNN dimulai dengan penginputan data latih dan data uji. Langkah awal adalah menentukan nilai K, yaitu jumlah tetangga terdekat yang digunakan dalam penghitungan. Selanjutnya, dilakukan pengukuran jarak antara setiap data uji dengan data latih menggunakan metode Euclidean Distance, yang digunakan untuk menentukan kedekatan antar titik data berdasarkan nilai atribut. Adapun rumus Euclidean Distance ditunjukkan sebagai berikut:

$$d(x_1, y_2) = \sqrt{\sum_{i=1}^n (x_{i1} - y_{i2})^2} \quad (1)$$

Di mana x_1 adalah vektor atribut dari data latih, y_2 adalah vektor atribut dari data uji, dan n adalah jumlah fitur. Hasil dari perhitungan jarak ini digunakan untuk memilih K data latih terdekat dari data uji yang sedang diuji.

Setelah jarak dihitung dan tetangga terdekat ditentukan, langkah berikutnya adalah menghitung derajat keanggotaan awal (initial membership) dari masing-masing tetangga terhadap setiap kelas. Perhitungan nilai awal keanggotaan u_{ij} yaitu derajat keanggotaan data latih ke-j terhadap kelas ke-i, dilakukan dengan rumus berikut:

$$u_{ij} = \begin{cases} 0.51 + \left(\frac{n_j}{n} * 0.49\right), & \text{Jika } j=i \\ \frac{n_i}{n} * 0.49, & \text{Jika } j \neq i \end{cases} \quad (2)$$

Di mana n_j adalah jumlah data latih yang termasuk dalam kelas ke-i, dan n adalah total data latih. Nilai u_{ij} ini nantinya digunakan dalam perhitungan derajat keanggotaan data uji terhadap masing-masing kelas.

Proses penentuan nilai keanggotaan akhir dari data uji terhadap kelas ke-i dihitung menggunakan rumus:

$$u_i(x) = \frac{\sum_{j=1}^k u_{ij} \left(\|x - x_j\|^{\frac{-2}{(m-1)}} \right)}{\sum_{j=1}^k 1 \left(\|x - x_j\|^{\frac{-2}{(m-1)}} \right)} \quad (3)$$

Hasil akhir dari klasifikasi ditentukan berdasarkan nilai keanggotaan tertinggi dari data uji terhadap kelas yang tersedia. Dengan demikian, prediksi jenis skizofrenia dari masing-masing pasien akan diperoleh secara lebih fleksibel dan informatif dibanding metode klasifikasi biner biasa.

6. Evaluasi

Model klasifikasi berbasis Fuzzy K-Nearest Neighbor (FKNN) ini dievaluasi menggunakan berbagai metrik performa untuk mengukur tingkat keberhasilan dalam mengklasifikasikan jenis skizofrenia berdasarkan gejala pasien. Proses evaluasi ini bertujuan untuk menilai sejauh mana model dapat memberikan prediksi yang akurat, tepat, dan relevan sesuai dengan label diagnosis aktual. Metrik evaluasi yang digunakan dalam penelitian ini mencakup akurasi, presisi, recall, F1-score, dan confusion matrix. Akurasi menunjukkan seberapa sering model memberikan prediksi yang benar secara keseluruhan, presisi mengukur proporsi prediksi benar di antara semua prediksi positif, sedangkan recall menilai sejauh mana model mampu menangkap seluruh data yang seharusnya diklasifikasikan ke dalam kelas tertentu. F1-score berperan dalam menyeimbangkan antara presisi dan recall, terutama saat menangani data yang tidak seimbang antar kelas. Sementara itu, confusion matrix digunakan untuk memahami distribusi klasifikasi yang dilakukan oleh model, termasuk jumlah prediksi yang benar maupun salah untuk masing-masing kelas. Hasil evaluasi dalam penelitian ini menunjukkan bahwa FKNN mampu memberikan performa klasifikasi yang sangat baik, dengan akurasi tinggi dan kesalahan klasifikasi yang rendah. Oleh karena itu, model ini dinilai efektif untuk mendukung proses diagnosis skizofrenia secara objektif dan efisien dalam lingkungan klinis maupun sistem pendukung keputusan berbasis data.

III. HASIL DAN PEMBAHASAN

Hasil evaluasi menunjukkan bahwa metode FKNN merupakan pendekatan yang efektif dalam klasifikasi jenis skizofrenia berdasarkan gejala pasien. Model ini menghasilkan akurasi tinggi serta nilai precision, recall, dan F1-score yang seimbang di seluruh kelas diagnosis. Meskipun demikian, perlu diperhatikan adanya potensi kesalahan klasifikasi, terutama pada kasus false negatives, yaitu saat model mengklasifikasikan pasien dengan gangguan tertentu sebagai tipe lain yang lebih ringan.

Dalam konteks medis, kesalahan ini dapat berdampak signifikan terhadap penanganan pasien yang tidak sesuai dengan kondisi sebenarnya. Untuk mengurangi risiko tersebut, beberapa strategi seperti penyesuaian parameter K, penambahan variasi data latih, serta pengayaan fitur gejala yang lebih relevan dapat diterapkan. Pendekatan ini diharapkan dapat meningkatkan sensitivitas model terhadap pola gejala yang tumpang tindih antar jenis skizofrenia, sehingga prediksi menjadi lebih akurat dan lebih bermanfaat dalam mendukung diagnosis klinis.

A. Hasil Evaluasi Model

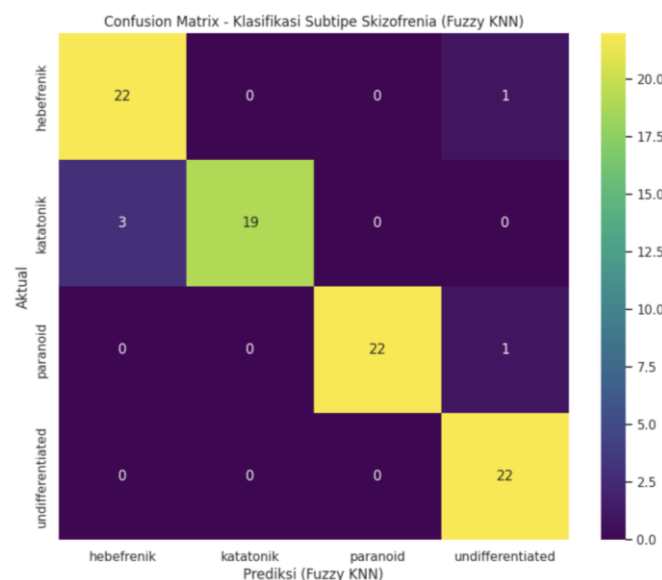
TABEL 1. Evaluasi Kinerja Model Fuzzy K-Nearest Neighbor

	Precision	Recall	F1-Score	Support
Hebefrenik	0.88	0.92	0.92	23
Katatonik	1.00	0.86	0.93	22
Paranoid	1.00	0.96	0.98	23
Undifferentiated	0.92	1.00	0.96	22
Accuracy			0.94	90
Macro AVG	0.95	0.94	0.94	90
Weighted AVG	0.95	0.94	0.94	90

Hasil evaluasi model klasifikasi menggunakan algoritma FKNN menunjukkan performa yang sangat baik dengan akurasi keseluruhan mencapai 94%. Ini berarti dari 90 data uji, 94% berhasil diklasifikasikan dengan benar ke dalam jenis skizofrenia yang tepat. Kelas paranoid memperoleh precision dan recall tertinggi, masing-masing 1.00 dan 0.96, serta F1-score 0.98, menandakan model sangat mampu mengenali jenis skizofrenia ini secara konsisten. Kelas undifferentiated juga menunjukkan hasil kuat dengan recall sempurna (1.00) dan F1-score 0.96, menandakan tidak ada data dari kelas ini yang salah klasifikasi. Sebaliknya, kelas katatonik memiliki recall paling rendah (0.86) walaupun precision-nya sempurna (1.00), menunjukkan adanya beberapa kesalahan klasifikasi yang mengarahkan data katatonik ke kelas lain, yang merefleksikan tantangan dalam membedakan gejala yang mirip.

Nilai macro dan weighted average yang mendekati 0.94–0.95 menunjukkan model bekerja secara konsisten dan tidak bias terhadap kelas tertentu. Namun, evaluasi ini belum membahas aspek trade-off antara sensitivitas dan spesifisitas lebih jauh, seperti analisis ROC/AUC, yang dapat menjadi fokus pengembangan model berikutnya. Jika dibandingkan dengan penelitian terdahulu yang menggunakan metode lain, seperti Logistic Regression yang mengalami kesulitan dalam membedakan pasien sehat dan pasien skizofrenia, FKNN menunjukkan potensi yang lebih baik dalam menangani ketidakpastian dan tumpang tindih gejala. Meskipun demikian, studi lebih lanjut dengan baseline metode lain sangat dianjurkan untuk menguatkan klaim performa ini.

B. Confusion Matrix



Gambar 4 Confusion Matrix

Confusion matrix memperlihatkan performa model FKNN dalam mengklasifikasikan jenis skizofrenia secara detail. Pada kelas hebefrenik, 22 dari 23 data diklasifikasikan dengan benar, sementara 1 data salah diklasifikasikan sebagai undifferentiated. Untuk kelas katatonik, terdapat 19 prediksi benar, namun 3 data salah diklasifikasikan sebagai hebefrenik, menandakan adanya kemiripan pola gejala yang mempengaruhi akurasi klasifikasi. Kelas paranoid berhasil diklasifikasikan dengan sangat baik, 22 data benar dan 1 salah masuk kelas undifferentiated. Sedangkan pada kelas undifferentiated, seluruh 22 data berhasil diklasifikasikan tanpa kesalahan. Dominasi angka pada diagonal utama menunjukkan bahwa model ini memiliki tingkat akurasi tinggi dan mampu mengenali ciri khas masing-masing jenis skizofrenia. Kesalahan klasifikasi yang muncul relatif kecil dan tersebar, sehingga model menunjukkan kestabilan dan keseimbangan hasil pada seluruh kelas diagnosis.

IV. KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma Fuzzy K-Nearest Neighbor (FKNN) mampu mengklasifikasikan jenis skizofrenia dengan akurasi tinggi sebesar 94%, membedakan empat tipe yaitu paranoid, katatonik, hebephrenik, dan undifferentiated berdasarkan gejala klinis dan data demografis pasien. Model ini menunjukkan performa yang konsisten pada metrik precision, recall, dan F1-score, sehingga berpotensi digunakan sebagai alat bantu diagnosis yang lebih cepat dan objektif dalam praktik medis. Namun, keterbatasan masih ada terutama dalam menangani gejala yang tumpang tindih antar jenis skizofrenia dan ketidakseimbangan data, yang menyebabkan kesalahan klasifikasi terutama pada kelas katatonik. Untuk itu, penelitian selanjutnya disarankan mengimplementasikan teknik oversampling seperti SMOTE dan penyesuaian threshold untuk meningkatkan sensitivitas model. Integrasi model ini ke dalam sistem informasi rumah sakit jiwa seperti RSJD Dr. Amino Gondohutomo dapat meningkatkan efisiensi dan akurasi proses diagnosis, sekaligus menjawab kebutuhan akan metode diagnosis berbasis data yang lebih objektif pada gangguan mental berat.

REFERENSI

- [1] Firmansyah, Dian Palupi Rini, Sukemi. "Klasifikasi Data Penderita Skizofrenia Menggunakan CNN-LSTM dan CNN-GRU pada Data Sinyal EEG 2D," *Jurnal Teknologi Informasi dan Komunikasi*, 2023; 642-650
- [2] Muhammad Khairi Lubis. "Sistem Pakar Diagnosa Penyakit Skizofrenia dengan CSP dan Certainty Factor," *Jurnal Sains dan Teknologi Informasi*, 2023; Vol 2 No 2. 66–75
- [3] Triandini Paramita, Setyani Alfinuha. "Dinamika Pasien dengan Gangguan Skizofrenia" *Jurnal Psikologi*, Maret 2021; Vol. 17 No 1, 12–19.
- [4] Ai Siti Mardiah, Aulia Rahmawati, Tahrizi Fathul Aliim, Sahadi Humaedi. "Praktik Support Group Bagi Orang dengan Skizofrenia," *Focus: Jurnal Pekerjaan Sosial*, Juli 2022; Vol5 No.1 37-47
- [5] Rebriyani Riyanda, Imam Cholissodiin, Sutrisno "Klasifikasi Gangguan Jiwa Skizofrenia Menggunakan Algoritma Decision Tree C5.0." *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* Oktober 2019; 10176-10182
- [6] Nadia Siburian, Imam Cholissodin, Putra Pandu Adikara. "Penerapan Fuzzy K-Nearest Neighbor pada Klasifikasi Penyakit Menular Seksual Pria," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, November 2020; 4096-4102
- [7] Priskilla Puji Larasati. "Sistem Pakar Diagnosis Penyakit Skizofrenia Menggunakan Metode

- Certainty Factor Berbasis Web,” *Jurnal Mahasiswa Teknik Informatika*, Maret 2019; Vol.3 No.1
- [8] Khoirun Nisa, Sony Kartika Wibisono. “Klasifikasi Penyakit Skizofrenia Menggunakan Algoritma Logistic Regression.” *Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK)* Agustus 2023; Vol. 8 No,2 696-704
- [9] Hardyan Zalfi, Agus Wahyu Widodo, Bayu Rahayudi. “Penerapan Metode Fuzzy K-Nearest Neighbour (FK-NN) untuk Diagnosis Penyakit pada Kucing” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Oktober 2019; Vol.3 No.10 10078-10085
- [10] Wildan Gita Akbari, Nurul Hidayat, Nurudin Santoso. “Diagnosis Penyakit Cabai Menggunakan Metode Fuzzy K-Nearest Neighbor (FKNN)” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Januari 2019; Vol.3 No. 1 1070-1074
- [11] Made Amalia Kristanto Dewi, Luh Made Karisma Sukmayanti S. “Dukungan Sosial dan Skizofrenia” *Psikobuletin: Buletin Ilmiah Psikologi*, September 2020; Vol 1. No. 3 179-186
- [12] Ningnurani, Muslimah Zahro Romas, Fx. Wahyu Widianoro. “Studi Kasus Penderita Skizofrenia Paranoid” *Jurnal Psikologi*, 2022; Vol.18 No.1 25–29.
- [13] Ike Asana, Amnan, B Fitria Maharani. “Skizofrenia: Suatu Studi Literatur,” *Journal of Public Health and Medical Studies*, Juni 2022; Vol.1 No.1
- [14] Cynthia Silaban “Penerapan Certainty Factor dalam Mendiagnosa Skizofrenia Berbasis Komputer,” *Journal of Pharmaceutical and Health Research*, Juni 2020; Vol.1 No.2 32-36
- [15] Maria Adelvin Londa, Kristina Sara, Melky Radja. “Penerapan Logika Fuzzy dalam Evaluasi Kinerja Dosen.” *Matrix: Jurnal Manajemen Teknologi dan Informatika*, Juli 2020; Vol.10 No.2