

## Sistem Otomatis untuk Deteksi Penyakit Parkinson Menggunakan Fuzzy K-NN

Ery Permana Yudha<sup>1\*</sup>, Nibras Faiq Muhammad<sup>2</sup>

<sup>1,2</sup>Universitas Duta Bangsa Surakarta

<sup>1\*</sup>[ery\\_permanayudha@udb.ac.id](mailto:ery_permanayudha@udb.ac.id)

<sup>2</sup>[nibras\\_faiqmuhammad@udb.ac.id](mailto:nibras_faiqmuhammad@udb.ac.id)

Diterima	Direvisi	Disetujui
01-01-2020	01-02-2020	01-03-2020

**Abstrak** - Penyakit Parkinson merupakan salah satu penyakit neurodegeneratif. Penyakit ini muncul karena beberapa faktor resiko yang berkontribusi seperti usia, jenis kelamin, dan beberapa faktor lain. Oleh sebab itu, identifikasi penyakit Parkinson sejak dini diperlukan untuk mencegah terjadinya resiko yang lebih buruk. Dalam membangun sistem identifikasi penyakit Parkinson secara otomatis membutuhkan dataset rekam medis. Dataset tersebut merupakan sekumpulan data frekuensi dan amplitudo gelombang suara dari beberapa subjek. Analisis data yang kuat merupakan hal yang paling menantang dalam mendeteksi penyakit Parkinson. Selain itu, juga diperlukan sebuah sistem yang cepat atau otomatis dalam menganalisis data klinis tersebut. Maka dari itu, kami mengusulkan sebuah sistem otomatis yang bertujuan untuk mendeteksi penyakit Parkinson menggunakan metode Fuzzy K-NN (K-Nearest Neighbor). Fuzzy K-NN merupakan modifikasi dari algoritma K-NN dengan menambahkan teknik fuzzy logic di dalamnya. Teknik ini mampu meningkatkan kinerja sistem dalam menentukan seseorang terdiagnosis penyakit Parkinson atau tidak. Metode yang kami usulkan menghasilkan nilai akurasi sebesar 90% di mana lebih baik 8%, 10%, 14.5%, dan 20% daripada metode Naïve Bayes, SVM, K-NN, Decision Tree.

Kata Kunci: sistem otomatis, penyakit parkinson, fuzzy k-nn

**Abstract** – Parkinson’s disease is a neurodegenerative disease. This disease appears due to several factors such as age, gender, and others. Therefore, early identification of Parkinson’s disease is needed to prevent a worse risk in the future. In building a Parkinson’s disease identification system automatically requires a medical record dataset. The dataset is a collection of frequency and amplitude data of sound waves from several subjects. Medical data analysis is the most challenging thing especially in detecting Parkinson’s disease. In addition, a system that is fast and automatically in analyzing medical data is also needed. Therefore, we propose an automatic system that aims to detect Parkinson’s disease using the Fuzzy K-NN (K-Nearest Neighbor) method. Fuzzy K-NN is a modification of the normal K-NN algorithm by adding the fuzzy logic technique in it. This technique is able to improve the performance of the system in determining whether a person is diagnosed with Parkinson’s disease or not. The proposed method yields 90% of accuracy which is better than Naïve Bayes, SVM, K-NN, Decision Tree with 8%, 10%, 14.5%, and 20% respectively.

Keywords: automatic system, parkinson’s disease, fuzzy k-nn

### PENDAHULUAN

Penyakit Parkinson merupakan salah satu penyakit neurodegeneratif. Neurodegeneratif merupakan suatu kondisi di mana sel saraf mengalami kehilangan struktur atau fungsi. Penyakit ini disebabkan karena kurangnya dopamin di otak. Dopamin merupakan neurotransmitter yang mempengaruhi dalam gerakan, motivasi, memori, dan fungsi lainnya. Tingkat kinerja

otak pada penderita penyakit Parkinson akan menurun yang disebabkan oleh kerusakan sel dopamin. Sebagian besar penderita penyakit ini mengalami demensia, depresi, dan gejala gangguan otak lainnya (Pandya et al., 2019). Penyakit ini berbahaya karena otak merupakan organ vital yang berfungsi untuk mengatur gerakan seluruh tubuh. Banyak faktor yang mempengaruhi seseorang dapat



terkena penyakit Parkinson. Beberapa faktor tersebut yaitu stress, radikal bebas, dan zat beracun di udara. Namun, semua faktor itu memiliki tingkat signifikansi yang berbeda-beda dalam menyumbang risiko seseorang terkena penyakit Parkinson atau tidak. Untuk mengetahui seseorang merupakan penderita Parkinson dapat dilihat dari data rekam medis. Data rekam medis berupa data frekuensi yang dihasilkan dari rekaman suara setiap subyek. Kemudian dataset tersebut digunakan untuk mendiagnosis. Oleh sebab itu, diperlukan analisis yang kuat dalam menentukan fitur-fitur apa saja yang memiliki dampak paling besar. Biasanya analisis data rekam medis dilakukan oleh seorang dokter untuk memastikan pasien terkena penyakit atau tidak. Namun, pada saat ini sudah ada beberapa penelitian mengembangkan sistem deteksi penyakit secara otomatis dan dengan hasil yang akurat.

Berbagai macam penelitian telah mengembangkan sebuah sistem deteksi penyakit secara otomatis. Beberapa penelitian tersebut yaitu deteksi penyakit diabetes menggunakan citra retina (Ingle & Ambad, 2022), deteksi tumor otak menggunakan citra MRI (Magnetic Resonance Imaging) (Indraswari et al., 2022), prediksi penyakit jantung menggunakan teknik machine learning (Gavhane et al., 2018), dan deteksi penyakit Parkinson menggunakan SVM (Support Vector Machine) (Soumaya et al., 2021). Penelitian-penelitian tersebut menggunakan berbagai macam teknik dalam machine learning.

Teknik machine learning digunakan agar dapat menemukan hubungan antara faktor-faktor yang terdapat dalam data klinis terhadap pendeteksian penyakit. Berbagai macam teknik machine learning diterapkan dalam beberapa penelitian salah satunya algoritma K-NN (K-Nearest Neighbor). Algoritma ini pernah digunakan untuk mendiagnosis kanker paru-paru (Maleki et al., 2021). Penelitian tersebut mampu menghasilkan nilai akurasi yang sangat tinggi. Berdasarkan hasil penelitian tersebut algoritma K-NN merupakan salah satu algoritma machine learning yang handal.

Model machine learning dengan algoritma K-NN mempunyai banyak kelebihan seperti sederhana, efektif, intuitif, dan kinerja klasifikasi yang kompetitif (Imandoust & Bolandraftar, 2013). Selain itu, handal terhadap data training yang kotor atau noisy dan efektif jika ukuran data training besar. Namun, algoritma ini juga masih memiliki kekurangan. Beberapa kekurangan itu yaitu komputasi yang cukup lama jika ukuran data training yang sangat besar, sangat sensitif terhadap terhadap fitur-fitur yang tidak relevan atau berulang (Imandoust & Bolandraftar, 2013). Fitur-fitur yang tidak relevan tersebut mempengaruhi kinerja klasifikasi dari algoritma K-NN.

Pemilihan fitur perlu diterapkan pada dataset untuk menangani penurunan kinerja klasifikasi. Selain itu, dimensi fitur subset yang sangat besar akan mengakibatkan waktu komputasi yang lama. Oleh

sebab itu, perlu teknik untuk menyederhanakan dimensi fitur subset tersebut namun tanpa mengurangi kualitas dataset. Teknik ini dapat mempercepat komputasi tanpa mengurangi nilai akurasi.

Beberapa peneliti menerapkan beberapa metode penyederhanaan dimensi fitur subset salah satunya PCA (Principal Component Analysis) dalam penelitiannya. Penyederhanaan dimensi fitur subset juga diterapkan pada penelitian di bidang kesehatan. Beberapa penelitian tersebut yaitu memanfaatkan PCA sebagai metode ekstraksi fitur untuk deteksi penyakit Alzheimer (Biagetti et al., 2021) dan seleksi fitur menggunakan PCA untuk deteksi penyakit jantung (Gárate-Escamila et al., 2020). Kedua penelitian itu mampu mengurangi waktu komputasi dibandingkan tanpa menerapkan teknik penyederhanaan fitur subset. Efisiensi waktu komputasi yang dihasilkan bahkan mencapai 15,53%. Terlebih lagi nilai akurasi yang didapatkan mencapai 93,18% atau lebih tinggi 40% dibandingkan tanpa menggunakan PCA.

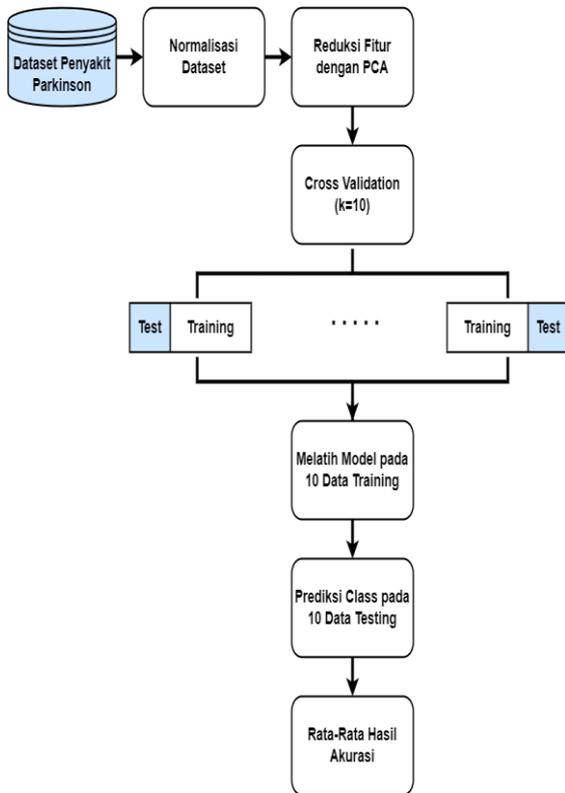
Peningkatan nilai akurasi tidak hanya dipengaruhi oleh seleksi fitur atau reduksi fitur subset saja. Masih ada faktor lain yaitu model klasifikasi yang digunakan. Model klasifikasi yang berbeda akan menghasilkan nilai akurasi yang berbeda pula. Oleh sebab itu, kita perlu mencari sebuah model yang mampu bekerja lebih baik dibandingkan model klasifikasi yang lain.

Pada penelitian ini, kami mengusulkan sebuah metode machine learning menggunakan Fuzzy K-NN untuk meningkatkan kinerja dalam mendeteksi penyakit Parkinson. Kinerja dari sistem yang kami bangun diukur berdasarkan nilai akurasi. Kemudian hasil kinerja dari metode yang kami usulkan dilakukan perbandingan dengan hasil kinerja metode klasifikasi lain seperti SVM, Decision Tree, Naïve Bayes, dan K-NN.

## METODE PENELITIAN

Pada bagian ini, kami menjelaskan dataset dan alur kerja dari metode yang kami usulkan. Dataset yang kami gunakan adalah dataset penyakit Parkinson yang terdiri dari 195 subyek, 24 fitur, dan 2 kelas. Kelas tersebut digolongkan sebagai penderita penyakit Parkinson atau normal. Dataset yang digunakan merupakan dataset publik. Alur kerja metode kami dapat dilihat pada Gambar 1. Metode yang kami usulkan terbagi menjadi tiga tahapan.

Pertama, tahap *preprocessing* yang terdiri dari normalisasi dan mengatasi *missing value* untuk mempersiapkan data untuk siap diolah menggunakan algoritma yang diusulkan. Kedua, tahap reduksi fitur menggunakan PCA (*Principal Component Analysis*) untuk mengeliminasi fitur-fitur yang kurang signifikan untuk meningkatkan kinerja klasifikasi. Ketiga, tahap membangun model klasifikasi Fuzzy K-NN dengan melatih data training menggunakan cross validation agar mendapatkan parameter yang optimal.



Gambar 1. Alur kerja metode yang diusulkan

Model Fuzzy K-NN yang sudah optimal kemudian digunakan untuk melakukan tugas klasifikasi.

### 1. Normalisasi

Normalisasi merupakan salah satu tahap penting dalam data mining untuk memastikan setiap data pada dataset tetap konsisten. Pada proses ini diperlukan transformasi data menjadi format yang lebih rapi agar menghasilkan pemrosesan yang lebih efisien. Salah satu bentuk transformasi yaitu membuat beberapa variabel atau fitur memiliki rentang nilai yang sama (0 hingga 1) tanpa mengurangi kualitas informasi dalam dataset.

Tahap ini bertujuan untuk mengubah data ke dalam bentuk yang lebih terstruktur, konsisten, dan efisien. Selain itu, bertujuan untuk mengubah nilai kolom numerik dalam himpunan data tanpa mendistorsi perbedaan dalam rentang nilai.

### 2. Reduksi fitur

Pada tahap ini dataset yang masuk kemudian diekstrak dan dipisahkan antara fitur dengan label. Sekumpulan fitur-fitur biasanya disebut dengan *feature subset* yang merupakan sekumpulan nilai-nilai dari variabel independen dari dataset. Kemudian *feature subset* tersebut dilakukan reduksi fitur. Tujuannya untuk meningkatkan efisiensi komputasi sehingga kinerja dari model yang dibangun bisa lebih cepat.

Banyak penelitian yang berhasil menunjukkan bahwa reduksi fitur berperan penting dalam meningkatkan

efektivitas kinerja model klasifikasi khususnya di bidang kesehatan (Gárate-Escamila et al., 2020), (Cao et al., 2022). Pada tahap ini memiliki dua tugas utama yaitu mengekstrak fitur dari dataset, mentransformasikan fitur ke dalam *feature subset*, dan mereduksi dimensi *feature subset*. Kami menggunakan PCA karena merupakan salah satu metode reduksi fitur sering digunakan peneliti.

PCA mampu mencari variansi data dari *feature subset* dengan baik. Secara keseluruhan jika algoritma reduksi fitur mampu mendefinisikan *feature subset* dengan baik maka proses klasifikasi akan lebih efektif dan efisien. Pada penelitian ini, kami ingin mencari tahu efektifitas PCA dalam reduksi fitur untuk deteksi penyakit Parkinson.

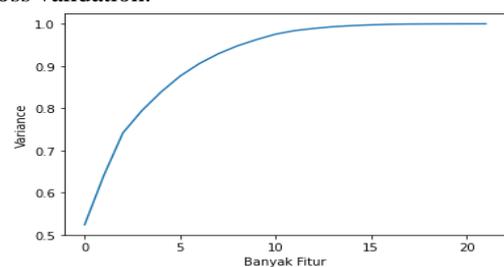
Dataset yang kami gunakan terdiri dari 24 fitur dan 195 baris. Fitur-fitur tersebut tidak semua dibutuhkan dalam membangun model klasifikasi. Oleh karena itu, kami menerapkan PCA pada dataset kami untuk menghasilkan suatu *feature subset* yang lebih efisien dan representatif dalam mewakili data.

*Feature subset* yang lebih efisien mampu mengurangi waktu komputasi sehingga mampu mendiagnosis lebih cepat. Berdasarkan Gambar 2 dapat didefinisikan bahwa semakin banyak fitur maka semakin besar pula variansinya. Variansi merupakan parameter untuk mengukur persebaran data. Namun, dengan jumlah fitur 10 sudah memiliki variansi yang sangat tinggi. Oleh sebab itu, pada penelitian ini kami memangkas jumlah fitur awal menjadi 10 fitur.

### 3. Klasifikasi

Pada tahap ini, model Fuzzy K-NN melakukan tugas klasifikasi menggunakan *features subset* yang sudah direduksi oleh PCA. Tahap ini terdiri dari beberapa proses, pertama kali kita harus mengatur nilai parameter yang optimal untuk model Fuzzy K-NN. Jika nilai parameter sudah optimal tentu akan berperan penting terhadap kinerja Fuzzy K-NN.

Oleh karena itu, kita mendesain sebuah penelitian untuk menemukan parameter fuzzy yang paling optimal pada model yang diusulkan. Pengaturan nilai parameter yang kami gunakan yaitu parameter kekuatan fuzzy  $m=1,05$  dan untuk validasi kinerja model klasifikasi menggunakan cross validation  $cv=10$  pada sejumlah  $k$  ketetanggan. Untuk setiap pilihan  $m$ , kita menguji rata-rata akurasi yang dihasilkan oleh model Fuzzy K-NN dengan analisis cross validation.



Gambar 2. Reduksi fitur dengan PCA

Hasil akurasi rata-rata yang tertinggi dipilih sebagai parameter *fuzzy* yang terbaik. Setelah memilih parameter *fuzzy* yang terbaik, model klasifikasi *Fuzzy K-NN* diuji pada data *testing* yang sudah direduksi *feature subset*. Kemudian hasil dari setiap *fold* pada *cross validation* dihitung rata-ratanya. Langkah-langkah metode *fuzzy k-nn* dapat dilihat pada pseudocode di bawah ini.

**Langkah 1:** Tentukan nilai *k* (banyaknya tetangga terdekat).

**Langkah 2:** Hitung kemiripan antara *feature vector* dan *class vector* menggunakan *Euclidean distance*.

**Langkah 3:** Urutkan berdasarkan hasil penghitungan kemiripan dari yang terkecil hingga terbesar.

**Langkah 4:** Hitung derajat keanggotaan setiap data terhadap seluruh kelas.

**Langkah 5:** Pilih kelas berdasarkan nilai derajat keanggotaan terbesar.

#### 4. Evaluasi kinerja

Bagian ini menjelaskan metode untuk mengukur kinerja model. Metode evaluasi yang digunakan yaitu *accuracy*, *sensitivity*, dan *specificity*. Ketiga metode tersebut didefinisikan sesuai dengan persamaan berikut ini:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{FP + TN} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (6)$$

Akurasi merupakan rasio prediksi benar (positif dan negatif) dengan keseluruhan data. Penghitungan nilai akurasi dapat dilihat pada Persamaan 1. *Sensitivity* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Penghitungan nilai *sensitivity* dapat dilihat pada Persamaan 2. *Specificity* merupakan kebenaran memprediksi negatif dibandingkan dengan keseluruhan data negatif. Penghitungan nilai *specificity* dapat dilihat pada Persamaan 3.

Selain itu, kita juga mengukur kinerja model menggunakan *recall*, *precision*, dan *f1-score*. *Recall* merupakan perbandingan antara *True Positive* (TP) dengan banyaknya data yang sebenarnya positif, rumus dapat dilihat pada Persamaan 4. *Precision* merupakan perbandingan antara TP dengan



Gambar 3. Perbandingan tingkat akurasi

banyaknya data yang diprediksi positif, rumus *precision* dapat dilihat pada Persamaan 5. Kemudian, *f1-score* merupakan perbandingan rata-rata presisi dan *recall* yang dibobotkan, penghitungan *f1-score* dapat dilihat pada Persamaan 6.

Di mana *True Positive* (TP) adalah hasil prediksi positif dan memang benar. *True Negative* (TN) adalah hasil prediksi negatif dan memang benar. *False Positive* (FP) adalah hasil prediksi positif tetapi tidak benar. *False Negative* (FN) adalah hasil prediksi negatif tetapi tidak benar.

## HASIL DAN PEMBAHASAN

Dalam rangka untuk menguji kinerja dari metode yang diusulkan, pada tahap ini metode *Fuzzy K-NN* akan dibandingkan dengan metode lain yang sudah ada seperti *K-NN*, *SVM*, *Decision Tree C4.5*, dan *Naive Bayes*. Kinerja dari metode *Fuzzy K-NN* dan metode lain kemudian divisualisasikan dalam bentuk grafik. Pengukuran kinerja dilihat dari tingkat akurasi, *specificity*, *sensitivity*, *recall*, *precision*, dan *f1-score*.

### 1. Akurasi

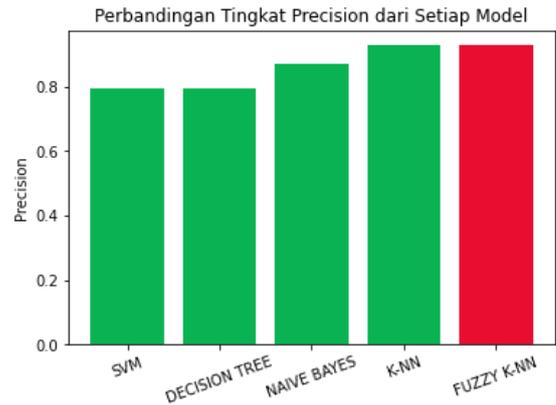
Pada pengujian tingkat akurasi, ternyata nilai yang dihasilkan *Fuzzy K-NN* jauh lebih baik dibandingkan metode lainnya. Parameter yang digunakan yaitu *k=3*. Hasil visualisasi perbandingan tingkat akurasi dapat dilihat pada Gambar 3. *Fuzzy K-NN* menghasilkan akurasi tertinggi yaitu 90% kemudian disusul *Naive Bayes* dengan 82%, *SVM* dengan 80%, *K-NN* dengan 75.5%, dan yang terakhir adalah *Decision Tree* dengan 70%.

### 2. Specificity

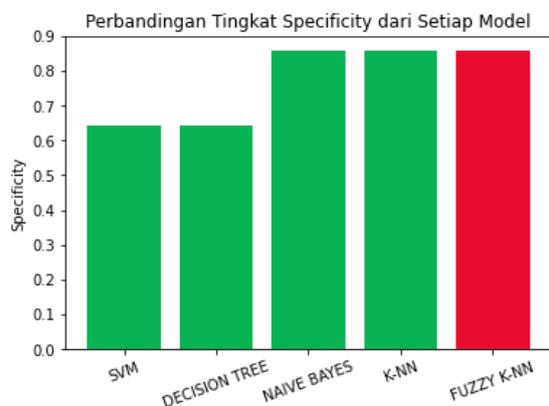
Pada pengukuran kinerja *specificity*, terdapat 3 metode dengan nilai *specificity* tertinggi. Ketiga metode tersebut adalah *Fuzzy K-NN*, *K-NN*, dan *Naive Bayes* yang sama-sama mendapatkan nilai 100%. Kemudian, untuk metode *Decision Tree* dan *SVM* sama-sama menghasilkan nilai *specificity* 64.28%. Hasil visualisasi perbandingan tingkat *specificity* dapat dilihat pada Gambar 4.



Gambar 5. Perbandingan tingkat *sensitivity*



Gambar 7. Perbandingan tingkat *precision*



Gambar 4. Perbandingan tingkat *specificity*.

### 3. Sensitivity

Pada pengukuran kinerja *sensitivity*, nilai tertinggi dihasilkan oleh metode *Fuzzy K-NN* dan *K-NN* yang sama-sama menghasilkan *specificity* sebesar 100%. Kemudian disusul dengan metode *Decision Tree* dengan 94.28%, *SVM* dengan 94.28, dan *Naïve Bayes* dengan 88.57%. Hasil visualisasi perbandingan tingkat *sensitivity* dapat dilihat pada Gambar 5.

### 4. Recall

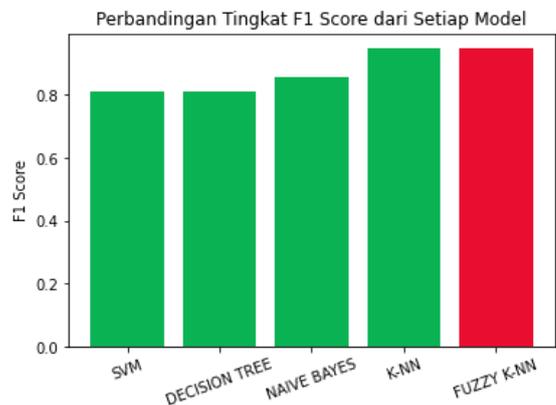
Pada pengujian tingkat *recall*, metode *Fuzzy K-NN* dan metode *K-NN* sama-sama mendapatkan hasil yang tertinggi dengan nilai *recall* sebesar 100%. Selanjutnya, disusul dengan metode *Decision Tree* dan *SVM* dengan nilai sebesar 81.8%. Hasil kinerja *recall* terendah yaitu metode *Naïve Bayes* dengan nilai sebesar 75%. Hasil visualisasi perbandingan tingkat *recall* dapat dilihat pada Gambar 6.

Pengujian selanjutnya adalah *precision*. Metode *Fuzzy K-NN* dan metode *K-NN* sama-sama mampu menghasilkan nilai *precision* terbaik yaitu sebesar 92.85%. Selanjutnya disusul dengan metode *Naïve Bayes* dengan nilai *precision* sebesar 87.14%. Di urutan terakhir diperoleh nilai *precision* sebesar 79.28% dengan menggunakan metode *Decision Tree* dan *SVM*. Hasil visualisasi perbandingan tingkat

*precision* dapat dilihat pada Gambar 7.

### 6. F1-Score

Pengujian yang terakhir yaitu *f1-score*, metode *Fuzzy*



Gambar 8. Perbandingan tingkat *f1-score*

*K-NN* dan metode *K-NN* masing-masing menghasilkan nilai yang sama sebesar 94.76%. Selanjutnya disusul dengan metode *Naïve Bayes* dengan nilai *f1-score* sebesar 85.88%. Terakhir nilai yang dihasilkan metode *Decision Tree* dan *SVM* yaitu 81.20%. Hasil visualisasi perbandingan tingkat *f1-score* dapat dilihat pada Gambar 8.

Berdasarkan hasil dari kinerja setiap metode dari beberapa skenario uji coba yang dapat dilihat pada Tabel 1 dan Tabel 2. Metode *Fuzzy K-NN* mendominasi seluruhnya dibandingkan metode lain. Untuk nilai akurasi *Fuzzy K-NN* memiliki nilai tertinggi dengan skor 90% atau 8% lebih baik dibandingkan *Naïve Bayes*. Selain itu, *Fuzzy K-NN* mampu meningkatkan akurasi 14.5% daripada metode normal *K-NN*.

Tabel 1. Perbandingan akurasi, specificity, dan sensitivity

Metode	Akurasi	Specificity	Sensitivity
Fuzzy K-NN	0.9	1	1
K-NN	0.755	1	1
SVM	0.8	0.642	0.942

Dec. Tree	0.7	0.642	0.942
Naïve Bayes	0.82	1	0.885

Tabel 2. Perbandingan *recall*, *precision*, dan *f1-score*

Metode	Recall	Precisioin	F1-Score
Fuzzy K-NN	1	0.928	0.947
K-NN	1	0.928	0.947
SVM	0.818	0.792	0.812
Dec. Tree	0.818	0.792	0.812
Naïve Bayes	0.75	0.871	0.858

## KESIMPULAN

Pada penelitian ini diusulkan sebuah metode *Fuzzy K-NN* yang merupakan modifikasi dari metode normal *K-NN* untuk mendeteksi penyakit Parkinson. Metode yang diusulkan terdapat penambahan logika *Fuzzy* di dalamnya yang bertujuan agar dapat meningkatkan kinerja klasifikasi. Selain itu, penelitian ini mengkombinasikan antara metode *Fuzzy K-NN* dengan teknik reduksi fitur yang diharapkan dapat meningkatkan efektifitas komputasi.

Hasil eksperimen menunjukkan metode yang diusulkan bekerja sangat baik dalam membedakan antara orang dengan penyakit Parkinson dengan orang yang sehat. Sementara, perbandingan metode antara *Fuzzy K-NN* dan metode yang sudah ada menunjukkan kinerja yang sangat baik dalam hal akurasi, *specificity*, *sensitivity*, *recall*, *precision*, dan *f1-score*. Oleh sebab itu, dapat disimpulkan bahwa sistem otomatis yang dikembangkan dapat berfungsi sebagai alat yang bermanfaat dalam pengambilan keputusan medis untuk diagnosis penyakit Parkinson.

## REFERENSI

- Biagetti, G., Crippa, P., Falaschetti, L., Luzzi, S., & Turchetti, C. (2021). Classification of Alzheimer's Disease from EEG Signal Using Robust-PCA Feature Extraction. *Procedia Computer Science*, 192, 3114–3122. <https://doi.org/10.1016/j.procs.2021.09.084>
- Cao, D., Lin, H., Liu, Z., Gu, Y., Hua, W., Cao, X., Qian, Y., Xu, H., & Zhu, X. (2022). Serum-based surface-enhanced Raman spectroscopy combined with PCA-RCKNCN for rapid and accurate identification of lung cancer. *Analytica Chimica Acta*, 1236, 340574. <https://doi.org/10.1016/j.aca.2022.340574>
- Gárate-Escamila, A. K., Hajjam El Hassani, A., & Andrés, E. (2020). Classification models for heart disease prediction using feature selection and PCA. *Informatics in Medicine Unlocked*, 19, 100330. <https://doi.org/10.1016/j.imu.2020.100330>
- Gavhane, A., Kokkula, G., Pandya, I., & Devadkar, K. (2018). Prediction of Heart Disease Using Machine Learning. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 1275–1278. <https://doi.org/10.1109/ICECA.2018.8474922>
- Imandoust, S. B., & Bolandraftar, M. (2013). Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background. 3(5).
- Indraswari, R., Ardan, I. S., Arifin, A. Z., Tjahyanto, A., Rakhmawati, N. A., & Kusumawardani, R. (2022). Brain Tumor Detection on Magnetic Resonance Imaging (MRI) Images Using Convolutional Neural Network (CNN). 2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 367–373. <https://doi.org/10.23919/EECSI56542.2022.9946622>
- Ingle, V., & Ambad, P. (2022). Diabetic retinopathy classifier with convolution neural network. *Materials Today: Proceedings*, S2214785322063052. <https://doi.org/10.1016/j.matpr.2022.09.480>
- Maleki, N., Zeinali, Y., & Niaki, S. T. A. (2021). A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection. *Expert Systems with Applications*, 164, 113981. <https://doi.org/10.1016/j.eswa.2020.113981>
- Pandya, S., Zeighami, Y., Freeze, B., Dadar, M., Collins, D. L., Dagher, A., & Raj, A. (2019). Predictive model of spread of Parkinson's pathology using network diffusion. *NeuroImage*, 192, 178–194. <https://doi.org/10.1016/j.neuroimage.2019.03.001>
- Soumaya, Z., Drissi Taoufiq, B., Benayad, N., Yunus, K., & Abdelkrim, A. (2021). The detection of Parkinson disease using the genetic algorithm and SVM classifier. *Applied Acoustics*, 171, 107528. <https://doi.org/10.1016/j.apacoust.2020.107528>