

Penerapan Data Mining Untuk Klasifikasi Penyakit Arthritis Gout Menggunakan Algoritma KNN

Yuni Handayani¹, Abdul Rahman Ismail², Bagus Prindo Sugihartono Putro³

^{1,3} Teknik Informatika, Fakultas Komputer dan Desain dan Universitas Selamat Sri

² Sistem Informasi, Fakultas Ilmu Komputer dan Universitas Ichsan Gorontalo Utara

yuni.04.handayani@gmail.com¹, abdulrahmanismail123@gmail.com², bagusprindo96@gmail.com³

ABSTRACT

This study aims to apply the K-Nearest Neighbor (KNN) algorithm to classify Gouty Arthritis based on patient clinical data. Gouty arthritis is a type of arthritis caused by a buildup of uric acid crystals in the joints, which can cause severe pain and swelling. The research method uses a dataset consisting of 500 entries, including information such as clinical symptoms, risk factors, medical history, and response to treatment. This data is processed through pre-processing stages to clean invalid data and convert categorical data into a format suitable for analysis. Next, the dataset is divided into training and test data for model building and evaluation. The KNN algorithm is applied to training data to build a classification model that can recognize important patterns related to the diagnosis of Gouty Arthritis. The evaluation was carried out using a test dataset to verify the accuracy and reliability of the model predictions. The research results show that the KNN algorithm at a value of $K=5$ achieved an accuracy of 87.76%, recall of 92.31%, and precision of 90.48% in classifying Gouty Arthritis. The implication of this research is the potential use of data mining techniques to support more effective diagnosis and disease management in clinical practice

Keywords : Arthritis Gout, K-Nearest Neighbor (KNN), Data Mining

ABSTRAK

Penelitian ini bertujuan untuk menerapkan algoritma K-Nearest Neighbor (KNN) dalam mengklasifikasikan penyakit Arthritis Gout berdasarkan data klinis pasien. Arthritis Gout adalah jenis artritis yang disebabkan oleh penumpukan kristal asam urat di dalam sendi, yang dapat menyebabkan nyeri hebat dan pembengkakan. Metode penelitian menggunakan dataset yang terdiri dari 500 entri, termasuk informasi seperti gejala klinis, faktor risiko, riwayat medis, dan respons terhadap pengobatan. Data ini diproses melalui tahapan pra-pemrosesan untuk membersihkan data yang tidak valid dan mengubah data kategorikal menjadi format yang sesuai untuk analisis. Selanjutnya, dataset dibagi menjadi data latih dan uji untuk pembangunan serta evaluasi model. Algoritma KNN diterapkan pada data latih untuk membangun model klasifikasi yang dapat mengenali pola-pola penting terkait dengan diagnosis Arthritis Gout. Evaluasi dilakukan menggunakan dataset uji untuk memverifikasi keakuratan dan keandalan prediksi model. Hasil penelitian menunjukkan bahwa algoritma KNN pada nilai $K=5$ mencapai nilai akurasi 87,76%, recall 92,31%, dan precision 90,48% dalam mengklasifikasikan penyakit Arthritis Gout. Implikasi dari penelitian ini adalah potensi penggunaan teknik data mining untuk mendukung diagnosis dan manajemen penyakit secara lebih efektif dalam praktik klinis

Keywords : Arthritis Gout, K-Nearest Neighbor (KNN), Data Mining

1. PENDAHULUAN

Arthritis Gout merupakan salah satu jenis artritis yang disebabkan oleh penumpukan kristal asam urat di dalam sendi, yang sering kali menimbulkan rasa nyeri yang hebat, pembengkakan, dan peradangan. Penyakit ini umumnya menyerang sendi pada jempol kaki, namun bisa juga mempengaruhi sendi lain seperti pergelangan kaki, lutut, dan tangan (Tabalujan et al., 2023). Prevalensi Arthritis Gout terus meningkat di seluruh dunia, seiring dengan gaya hidup yang tidak sehat dan meningkatnya usia harapan hidup. Diagnosis Arthritis Gout secara konvensional biasanya dilakukan melalui pemeriksaan klinis dan laboratorium, seperti analisis cairan sendi untuk mendeteksi kristal urat. Namun, metode ini sering memakan waktu dan memerlukan keahlian khusus. Selain itu, gejala Arthritis Gout yang mirip dengan jenis artritis lainnya seringkali menyebabkan kesulitan dalam penentuan diagnosis yang tepat. Seiring dengan perkembangan teknologi informasi, data mining muncul sebagai solusi potensial untuk membantu proses diagnosis penyakit. Data mining adalah proses untuk menemukan pola dan pengetahuan yang bermanfaat dari data yang besar dan kompleks. Dalam konteks medis, data mining dapat digunakan untuk menganalisis data klinis pasien, mengidentifikasi pola penyakit, serta memberikan prediksi dan rekomendasi untuk penanganan pasien.

Algoritma K-Nearest Neighbor (KNN) merupakan salah satu teknik data mining yang efektif untuk klasifikasi. KNN (Dwi Fasnuari et al., 2022) bekerja dengan cara mencari sejumlah titik data terdekat (tetangga terdekat) dari titik data baru dan menentukan kelasnya berdasarkan mayoritas kelas tetangga tersebut. KNN dikenal karena kesederhanaannya dalam implementasi dan kemampuan adaptasinya pada berbagai jenis data. Pada penelitian yang dilakukan oleh (Dwi Fasnuari et al., 2022) Happy Andrian Dwi Fasnuari, Haris Yuana, M. Taofik Chulkamdi pada tahun 2022 yang berjudul Penerapan Algoritma K-Nearest Neighbor (K-NN) Untuk Klasifikasi Penyakit Diabetes Melitus Studi Kasus : Warga Desa Jatitengah yang membahas tentang klasifikasi penyakit diabetes melitus dengan menggunakan algoritma KNN pada penelitian ini menggunakan 8 variabel yaitu mudah haus, berat badan meskipun konsumsi makan teratur, tekanan darah tinggi, terdapat Riwayat diabetes dalam keluarga, luka yang sulit sembuh, sering buang air di malam hari, serta hasil cek gula darah dan usia. Pada hasil penelitian menunjukkan data yang digunakan sebanyak 108 data training dan 27 data testing menghasilkan akurasi 93% pada $K=9$, presisi 100%, recall 60% dan F1-Score 75%. Dengan diperoleh nilai akurasi sebesar 93% menunjukkan bahwa penelitian ini telah berhasil menerapkan metode KNN untuk melakukan klasifikasi terhadap penyakit diabetes mellitus. Penelitian ini bertujuan untuk menerapkan algoritma KNN dalam mengklasifikasikan

penyakit Arthritis Gout berdasarkan data klinis pasien. Dengan menggunakan data mining, diharapkan dapat meningkatkan akurasi diagnosis serta mempercepat proses identifikasi penyakit. Selain itu, penelitian ini juga bertujuan untuk mengevaluasi kinerja algoritma KNN dan membandingkannya dengan metode klasifikasi lainnya, guna menemukan pendekatan terbaik dalam mendiagnosis Arthritis Gout. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam bidang medis, terutama dalam meningkatkan kualitas diagnosis dan penanganan Arthritis Gout melalui pemanfaatan teknologi data mining.

2. LANDASAN TEORI

2.1. Arthritis Gout

Arthritis Gout disebabkan oleh penumpukan kristal urat dalam sendi akibat tingginya kadar asam urat dalam darah. Faktor risiko termasuk genetik, diet tinggi purin, obesitas, dan kondisi seperti hipertensi (Nofia et al., 2021). Gejala utama meliputi serangan nyeri sendi akut, terutama pada jempol kaki, serta kemerahan dan pembengkakan. Diagnosis melibatkan analisis cairan sendi dan tes darah. Pengobatan meliputi penggunaan OAINS, kolkisin, dan obat penurun asam urat, disertai dengan perubahan gaya hidup. Pencegahan melibatkan diet sehat dan monitoring kadar asam urat secara teratur untuk mencegah serangan.

2.2. Data Mining

Data Mining merupakan salah satu disiplin ilmu yang berkaitan dengan proses (Handayani et al., 2024) mengeksplorasi dan menganalisis data besar untuk menemukan pola-pola yang bermanfaat dan pengetahuan yang tersembunyi. Tujuan utama dari data mining adalah untuk mengidentifikasi hubungan yang dapat digunakan untuk pengambilan keputusan yang lebih baik. Metode data mining mencakup berbagai teknik seperti klasifikasi, klustering, regresi, asosiasi, dan penggalian pola urutan (Cahyanti et al., 2023). Data mining diterapkan dalam berbagai industri dan bidang, termasuk bisnis, kedokteran, keuangan, pemasaran, dan ilmu pengetahuan lainnya untuk mendukung pengambilan keputusan berdasarkan analisis data yang kuat dan relevan.

2.3. KNN

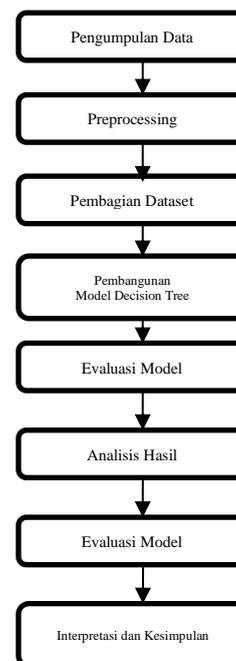
K-Nearest Neighbors (KNN) adalah algoritma dalam data mining dan machine learning yang digunakan untuk klasifikasi dan regresi. Algoritma ini menentukan kategori atau nilai target dari data baru berdasarkan mayoritas tetangga terdekatnya dalam ruang fitur (Sumarlinda & Lestari, 2022). KNN bekerja dengan menghitung jarak antara data yang akan diklasifikasikan dengan data pelatihan, lalu memilih kategori mayoritas dari tetangga terdekat (dengan nilai K yang ditentukan). Algoritma ini populer karena kesederhanaannya dan fleksibilitasnya dalam menangani berbagai jenis data.

2.4. Dataset

Penelitian ini bertujuan untuk menganalisis dan mengelompokkan data dari pasien dengan penyakit arthritis gout menggunakan sebuah dataset berisi 500 entri. Data tersebut mencakup informasi seperti gejala klinis, faktor

risiko, riwayat medis, dan respons terhadap pengobatan. Penelitian ini memiliki dua tujuan utama: pertama, mengidentifikasi pola-pola tersembunyi dalam data menggunakan algoritma pohon keputusan; kedua, mengembangkan model prediktif untuk mendukung diagnosa awal dan manajemen penyakit arthritis gout. Selama penelitian, dataset arthritis gout dipelajari secara menyeluruh untuk memahami distribusi fitur utama dan interaksi di antara mereka. Pra pemrosesan data dilakukan untuk membersihkan data dari nilai yang hilang atau tidak valid dan mengubah data kategorikal ke format yang sesuai untuk analisis. Dataset dibagi menjadi subset data latih dan uji untuk membangun serta mengevaluasi model. Algoritma KNN diterapkan pada dataset latih untuk membangun model klasifikasi yang mengenali pola-pola penting terkait diagnosis arthritis gout. Evaluasi dilakukan menggunakan dataset uji terpisah untuk memverifikasi keandalan dan ketepatan prediksi model. Hasil penelitian diharapkan memberikan wawasan baru tentang faktor faktor yang memengaruhi arthritis gout dan mendukung pengembangan model prediktif untuk penanganan yang lebih efektif.

3. METODOLOGI PENELITIAN



Gambar 1. Alur Penelitian

Penelitian ini menggunakan langkah-langkah sistematis untuk menganalisis dataset penyakit gout dan membuat model prediktif menggunakan algoritma KNN. Ini adalah tahapan utama dari metodologi penelitian:

1. Pengumpulan Data: 500 dataset pasien arthritis gout dikumpulkan dari berbagai sumber, termasuk survei, catatan medis, dan database kesehatan yang relevan. Data ini mencakup berbagai informasi seperti gejala klinis, faktor risiko, riwayat medis, dan respons pasien terhadap pengobatan.

2. Preprocessing Data: Langkah ini melibatkan pengolahan data kategorikal menjadi format yang sesuai untuk analisis dan pembersihan nilai-nilai yang hilang atau tidak valid. Mengubah format data dan menangani

nilai-nilai yang hilang atau tidak valid membantu mempersiapkan dataset untuk analisis lebih lanjut.

3. Pembagian Dataset: Data tentang penyakit gout dibagi menjadi dua kelompok: data latih dan data uji. Data latih digunakan untuk membuat model prediktif, sedangkan data uji digunakan untuk mengevaluasi dan memvalidasi model yang telah dikembangkan.

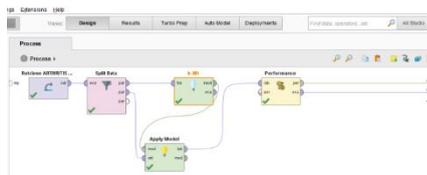
4. Membangun Model KNN: Algoritma KNN digunakan pada dataset latih untuk membangun model klasifikasi yang dapat menemukan pola penting yang terkait dengan diagnosis artritis gout. Proses ini melibatkan validasi silang untuk mengoptimalkan kinerja model dan penyesuaian parameter algoritma KNN.

5. Evaluasi Model: Evaluasi dilakukan menggunakan dataset uji yang berbeda setelah model KNN dibangun. Untuk memastikan keakuratan dan keandalan prediksi model, evaluasi melibatkan pengukuran akurasi, presisi, recall pada performance vector.

6. Analisis Hasil: Hasil evaluasi model digunakan untuk menganalisis kinerja model prediktif dan menemukan faktor-faktor yang penting untuk diagnosis artritis gout. Analisis ini membantu memahami peran variabel tertentu dalam prediksi diagnosis dan penanganan penyakit gout.

7. Interpretasi dan Kesimpulan: Hasil analisis memberikan interpretasi pola-pola yang ditemukan dalam data dan implikasinya terhadap diagnosis dan penanganan artritis gout. Penelitian ini mencakup hasil penting, masalah, dan rekomendasi untuk penelitian lebih lanjut. Penelitian ini bertujuan untuk mendapatkan pemahaman yang lebih baik tentang faktor-faktor yang mempengaruhi penyakit gout dan untuk membuat model prediktif yang dapat digunakan untuk membantu praktik klinis mengelola penyakit ini secara lebih baik.

3. HASIL DAN PEMBAHASAN



Gambar 2. Penghubungan port KNN , apply model dan performance

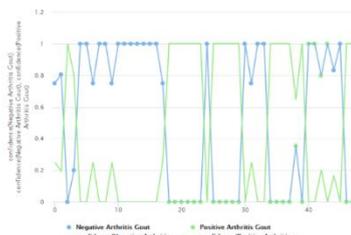
Pertama, input atau recovery data dilakukan, kemudian pembagian data dilakukan. Tahap berikutnya, seperti yang ditunjukkan pada gambar di atas, adalah menghubungkan port-port dari operator penerapan model, operator penerapan, dan operator kinerja. Untuk melihat hasilnya, klik icon run di toolbar. Untuk nilai akurasi dari Algoritma KNN.

	True Negative Arthritis Gout	True Positive Arthritis Gout	Class Precision
pred Negative Arthritis Gout	24	4	85.71%
pred Positive Arthritis Gout	2	19	90.48%
class recall	82.31%	89.51%	

Gambar 3. Hasil Akurasi Klasifikasi Penyakit Arthritis Gout Menggunakan Algoritma KNN

Nilai akurasi sistem sebesar 89,80% diperoleh melalui pengolahan data menggunakan aplikasi RapidMiner. Gambar menunjukkan bahwa prediksi tidak adalah 19, dan prediksi iya adalah 25, dan nilai precision sebesar 86,21% dan nilai recall sebesar 96,15%.

Dengan pengolahan data menggunakan aplikasi RapidMiner didapat nilai akurasi sistem sebesar 89.80%. Dari gambar di jelaskan bahwa prediksi tidak adalah 19 dan prediksi iya adalah 25 dengan nilai precision sebesar 86.21% dan nilai recall sebesar 96.15%.



Gambar 4. Visualisasi Arthritis Gout

Langkah berikutnya adalah menguji hasil analisis dengan perhitungan manual setelah data diklasifikasikan dan dianalisis menggunakan Algoritma KNN. Proses klasifikasi Arthritis Gout dilakukan dengan menggunakan Aplikasi RapidMiner.

KESIMPULAN

Algoritma KNN pada data mining dapat digunakan untuk mendapatkan informasi tentang klasifikasi penyakit Arthritis Gout. Sebuah Algoritma KNN dapat dibangun dari lima ratus contoh data pelatihan dan menghasilkan aturan untuk mengklasifikasikan penyakit Arthritis Gout pada nilai $K=5$, Akurasi 87,76%, recall 92,31%, dan precision 90,48% ditunjukkan oleh tes.

DAFTAR PUSTAKA

- Cahyanti, F. L. D., Sarasati, F., Astuti, W., & Firasari, E. (2023). Klasifikasi Data Mining Dengan Algoritma Machine Learning Untuk Prediksi Penyakit Liver. *Technologia : Jurnal Ilmiah*, 14(2), 134. <https://doi.org/10.31602/tji.v14i2.10093>
- Dwi Fasnuari, H. A., Yuana, H., & Chulkamdi, M. T. (2022). Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Penyakit Diabetes Melitus. *Antivirus : Jurnal Ilmiah Teknik Informatika*, 16(2), 133–142. <https://doi.org/10.35457/antivirus.v16i2.2445>
- Handayani, Y., Hidayat, T., & Putro, S. P. B. (2024). Pemodelan Data Mining Sebagai Klasifikasi Penyakit Arthritis Gout Dengan Teknik Decision Tree. *Jurnal Teknik Informatika Dan Desain Komunikasi Visual*, 3(1), 114–118. <https://www.jfik.uniss.ac.id/index.php/journals/article/view/83>
- Nofia, V. R., Apriyeni, E., & Prigawuni, F. (2021). Pendidikan Kesehatan Tentang Arthritis Gout Di Puskesmas Dadok Tunggul Hitam Padang. *Jurnal*

- Abdimas Saintika*, 3(1), 130.
<https://doi.org/10.30633/jas.v3i1.1108>
- Sumarlinda, S., & Lestari, W. (2022). Aplikasi K-Nearest Neighbor (KNN) untuk Klasifikasi Penyakit Kardiovaskuler. *Sumarlinda, Sri Lestari, Wiji*, 55, 259–262.
<http://ojs.uib.ac.id/index.php/Senatib/article/download/1897/1487>
- Tabalujan, J. A., Palandeng, H. M. F., & Ottay, R. I. (2023). Arthritis gout dan perilaku dokter keluarga di Kota Manado. *J Kedokt Kom Tropik*, 11(2), 463–466.