



Klasifikasi Bronchitis Menggunakan Metode K-Nearest Neighbor pada Data Gejala Klinis dan Parameter Fisiologis

Achmad Syafi'i Hasibuan^{1,*}, Akra Mukharrahmah², Rapit Mailenda³, Abdul Aziz⁴

^{1,2,3,4}Pendidikan Teknik Informatika Dan Komputer, UIN Sjech M Djamil Djambek Bukittinggi, Bukittinggi, Indonesia

Abstrak

Bronchitis merupakan penyakit inflamasi saluran pernapasan yang kerap didiagnosis terlambat akibat kesamaan gejala dengan penyakit respirasi lainnya seperti flu, pilek, dan pneumonia. Kondisi ini mendorong kebutuhan terhadap sistem pendukung keputusan klinis berbasis pembelajaran mesin yang mampu mengklasifikasikan diagnosis secara akurat berdasarkan pola gejala dan data tanda vital. Penelitian ini bertujuan membangun model prediksi bronchitis menggunakan algoritma K-Nearest Neighbor (KNN) dengan memanfaatkan dataset klinis yang memuat 2.000 rekam medis pasien. Dataset mencakup variabel demografis (usia, jenis kelamin), tiga gejala dominan, serta lima parameter fisiologis meliputi detak jantung, suhu tubuh, tekanan darah, dan saturasi oksigen. Fitur kategoris dikodekan menggunakan one-hot encoding, sementara fitur numerik dinormalisasi melalui Min-Max Scaler. Pembagian data dilakukan dengan rasio 80:20 untuk data latih dan data uji. Eksperimen dilakukan pada delapan nilai K (K=1 hingga K=15, kelipatan ganjil) dengan jarak Euclidean sebagai metrik kemiripan. Hasil pengujian menunjukkan bahwa nilai K=5 menghasilkan akurasi tertinggi sebesar 92,50% dengan nilai precision 85%, recall 67%, dan F1-score 75% untuk kelas bronchitis. Model berhasil mengidentifikasi 45 dari 67 kasus bronchitis pada data uji. Temuan ini membuktikan bahwa KNN efektif dalam mendeteksi bronchitis berdasarkan kemiripan pola gejala, meskipun masih terdapat tantangan pada penanganan ketidakseimbangan kelas yang perlu diatasi dalam penelitian lanjutan.

Kata kunci : bronchitis, K-Nearest Neighbor, klasifikasi penyakit, machine learning, prediksi diagnosis

Abstract

Bronchitis is a respiratory tract inflammation that is often diagnosed late due to symptom overlap with other respiratory diseases such as influenza, common cold, and pneumonia. This situation drives the need for machine learning-based clinical decision support systems capable of accurately classifying diagnoses based on symptom patterns and vital sign data. This study aims to develop a bronchitis prediction model using the K-Nearest Neighbor (KNN) algorithm, utilizing a clinical dataset containing 2,000 patient medical records. The dataset encompasses demographic variables (age, gender), three dominant symptoms, and five physiological parameters including heart rate, body temperature, blood pressure, and oxygen saturation. Categorical features were encoded using one-hot encoding, while numerical features were normalized through Min-Max Scaler. Data partitioning was performed at an 80:20 ratio for training and testing sets. Experiments were conducted across eight K values (K=1 to K=15, odd multiples) using Euclidean distance as the similarity metric. Results indicate that K=5 yields the highest accuracy of 92.50%, with precision of 85%, recall of 67%, and F1-score of 75% for the bronchitis class. The model successfully identified 45 out of 67 bronchitis cases in the test set. These findings confirm that KNN is effective in detecting bronchitis based on symptom pattern similarity, although challenges related to class imbalance require further investigation.

Keywords: bronchitis, K-Nearest Neighbor, disease classification, machine learning, diagnostic prediction

1. Introduction

Penyakit saluran pernapasan merupakan salah satu penyebab utama morbiditas dan mortalitas global, dengan bronchitis kronis dan akut menempati posisi signifikan dalam beban penyakit dunia. Organisasi Kesehatan Dunia (WHO) melaporkan bahwa penyakit pernapasan kronis menyumbang lebih dari 4 juta kematian setiap tahunnya, dan bronchitis kronis dikaitkan erat dengan paparan asap [1].

rokok, polusi udara, serta infeksi virus dan bakteri [2]. Di Indonesia, prevalensi penyakit pernapasan termasuk bronchitis terus meningkat seiring dengan tingginya tingkat polusi di kawasan perkotaan dan rendahnya kesadaran masyarakat terhadap pentingnya deteksi dini [3].

Tantangan utama dalam diagnosis bronchitis terletak pada kemiripan klinis yang tinggi antara bronchitis dengan kondisi respirasi lainnya seperti influenza, infeksi saluran pernapasan atas (ISPA), dan pneumonia. Gejala-gejala seperti batuk, sesak napas, nyeri tubuh, dan demam merupakan presentasi umum yang dapat muncul pada berbagai diagnosis, sehingga proses klasifikasi secara manual rentan terhadap kesalahan dan keterlambatan penanganan [4]. Kondisi ini memperkuat urgensi pengembangan sistem pendukung keputusan klinis (Clinical Decision Support System/CDSS) yang memanfaatkan kecerdasan buatan untuk membantu tenaga medis dalam mengambil keputusan diagnosis yang lebih cepat dan akurat.

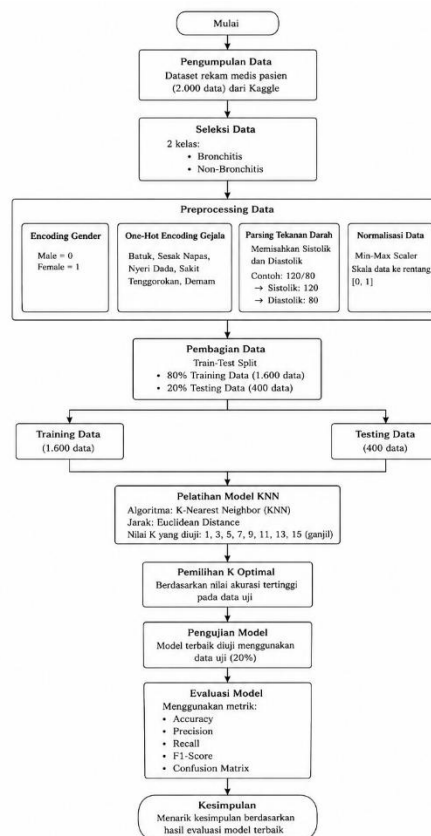
Dalam dekade terakhir, algoritma machine learning telah banyak diaplikasikan dalam bidang informatika kesehatan untuk keperluan klasifikasi dan prediksi penyakit. Beberapa studi terdahulu telah mengeksplorasi penggunaan berbagai metode, antara lain Decision Tree [5], Random Forest [6], Support Vector Machine (SVM) [7], dan Naive Bayes [8] untuk klasifikasi penyakit pernapasan. Namun demikian, K-Nearest Neighbor (KNN) tetap menjadi algoritma yang relevan dan kompetitif karena sifatnya yang non-parametrik, mudah diinterpretasikan, dan efektif dalam mendeteksi kemiripan pola antar instans data karakteristik yang sangat sesuai dengan konteks diagnosis berbasis gejala [9].

Sejumlah penelitian sebelumnya telah menggunakan KNN untuk prediksi penyakit, di antaranya untuk deteksi diabetes [10], penyakit jantung [11], dan klasifikasi pneumonia [4]. Akan tetapi, penerapan KNN secara spesifik untuk prediksi bronchitis dengan mengintegrasikan fitur gejala klinis multidimensi dan parameter tanda vital fisiologis masih relatif terbatas. Sebagian besar penelitian hanya menggunakan subset fitur tertentu dan tidak mengeksplorasi pengaruh kombinasi gejala dan data vital secara menyeluruh terhadap akurasi prediksi bronchitis. Celah inilah yang menjadi dasar penelitian ini.

Penelitian ini bertujuan untuk membangun dan mengevaluasi model prediksi bronchitis menggunakan algoritma KNN dengan memanfaatkan dataset klinis yang mencakup variabel demografis, tiga gejala dominan, dan lima parameter fisiologis dari 2.000 rekam medis pasien. Novelty penelitian ini terletak pada integrasi komprehensif antara fitur gejala berbasis one-hot encoding dan fitur numerik tanda vital yang dinormalisasi dalam satu pipeline KNN yang dioptimalkan melalui seleksi nilai K sistematis, sehingga menghasilkan model prediksi bronchitis yang lebih holistik dan reproducible. Pertanyaan penelitian yang dijawab adalah: (1) Pada nilai K berapa algoritma KNN mencapai akurasi optimal untuk prediksi bronchitis? dan (2) Bagaimana performa model KNN dalam mengklasifikasikan kasus bronchitis berdasarkan metrik precision, recall, dan F1-score?

2. Method

Tahapan penelitian yang diterapkan dalam pengembangan model prediksi bronchitis menggunakan metode K-Nearest Neighbor (KNN) meliputi proses pengumpulan data, preprocessing, pembagian data, pelatihan model, pengujian, dan evaluasi performa. Rangkaian tahapan penelitian tersebut disajikan pada Gambar 1.



Gambar 1. Alur penelitian prediksi bronchitis menggunakan metode K-Nearest Neighbor

Alur penelitian terdiri atas beberapa tahapan utama. Dataset rekam medis pasien terlebih dahulu diproses melalui tahap preprocessing yang mencakup encoding fitur kategoris, one-hot encoding gejala, parsing tekanan darah, dan normalisasi data menggunakan Min-Max Scaler. Data yang telah diproses kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20 menggunakan stratified split. Selanjutnya, model KNN dibangun dan dievaluasi pada beberapa nilai K untuk memperoleh parameter optimal. Performa model terbaik dianalisis menggunakan metrik accuracy, precision, recall, F1-score, dan confusion matrix.

2.1. Dataset dan Sumber Data

Dataset yang digunakan dalam penelitian di ambil dari www.kaggle.com/datasets/s3programmer/disease-diagnosis-dataset ini merupakan dataset klinis sintetis yang memuat 2.000 rekam medis pasien dengan 13 atribut, yaitu: Patient_ID, Age, Gender, Symptom_1, Symptom_2, Symptom_3, Heart_Rate_bpm, Body_Temperature_C, Blood_Pressure_mmHg, Oxygen_Saturation_%, Diagnosis, Severity, dan Treatment_Plan. Kolom Diagnosis mencakup lima kelas: Healthy (1.167 kasus; 58,35%), Bronchitis (334 kasus; 16,70%), Flu

(292 kasus; 14,60%), Cold (163 kasus; 8,15%), dan Pneumonia (44 kasus; 2,20%). Dalam penelitian ini, masalah diformulasikan sebagai klasifikasi biner, yakni membedakan kelas Bronchitis dari kelas Non-Bronchitis (gabungan keempat kelas lainnya), dengan target prediksi berfokus pada deteksi dini bronchitis.

2.2. Preprocessing Data

Tahap preprocessing meliputi tiga sub-proses utama. Pertama, encoding fitur kategoris: variabel Gender diubah menjadi nilai biner (Male=1, Female=0), sedangkan variabel Symptom_1, Symptom_2, dan Symptom_3 yang masing-masing dapat berisi delapan nilai gejala berbeda (Cough, Fever, Fatigue, Shortness of breath, Runny nose, Body ache, Headache, Sore throat) direpresentasikan melalui mekanisme one-hot encoding menjadi vektor biner berukuran 8 dimensi yang mencerminkan kehadiran atau ketiadaan setiap gejala. Kedua, parsing dan pemisahan tekanan darah: atribut Blood_Pressure_mmHg yang berformat string 'sistolik/diastolik' diurai menjadi dua fitur numerik terpisah (systolic dan diastolic). Ketiga, normalisasi fitur numerik: lima fitur numerik (Age, Heart_Rate_bpm, Body_Temperature_C, tekanan darah sistolik, tekanan darah diastolik, dan Oxygen_Saturation_%) dinormalisasi menggunakan Min-Max Scaler untuk memetakan nilai ke rentang [0, 1], sehingga tidak ada satu pun fitur yang mendominasi perhitungan jarak Euclidean akibat perbedaan skala pengukuran. Total dimensi vektor fitur akhir yang digunakan adalah 15 fitur (8 gejala biner + 7 fitur numerik terstandarisasi)

2.3. Pembagian Data dan Pemilihan Nilai K

Dataset dibagi dengan rasio 80:20 menggunakan stratified split untuk memastikan proporsi kelas bronchitis terjaga di kedua subset, menghasilkan 1.600 data latih dan 400 data uji. Stratifikasi dilakukan terhadap variabel target biner (Bronchitis/Non-Bronchitis). Untuk pemilihan nilai K optimal, eksperimen dilakukan secara sistematis pada delapan nilai K ganjil, yaitu K=1, 3, 5, 7, 9, 11, 13, dan 15. Pendekatan ini mengikuti konvensi pemilihan K ganjil untuk menghindari hasil seri (tie) dalam pemungutan suara mayoritas. Metrik jarak yang digunakan adalah Euclidean distance, diformulasikan sebagai

$$d(x,y) = \sqrt{\sum_i (x_i - y_i)^2} \quad (1)$$

di mana x dan y merepresentasikan dua instans data dalam ruang fitur 15 dimensi, dan d adalah jarak Euclidean antara keduanya.

2.4. Evaluasi Model

Performa model dievaluasi menggunakan empat metrik utama: accuracy (proporsi prediksi benar terhadap total data uji), precision (proporsi prediksi positif yang benar terhadap total prediksi positif), recall atau sensitivity (proporsi kasus bronchitis aktual yang berhasil terdeteksi), dan F1-score (rata-rata harmonik precision dan recall). Dengan rumus sebagai berikut.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \quad (4)$$

Selain itu, confusion matrix digunakan untuk memvisualisasikan distribusi True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) dari model terbaik. Implementasi dilakukan menggunakan Python 3.x dengan pustaka scikit-learn untuk pemodelan KNN, NumPy dan Pandas untuk manipulasi data, serta MinMaxScaler untuk normalisasi fitur.

3. Results and Discussion

Gunakan Microsoft Equation Editor untuk menuliskan persamaan pada artikel anda. Beri nomor persamaan secara urut di dalam tanda kurung dan letakkan pada tepi kanan, seperti (1) dengan ukuran hurufnya adalah 8 pt.

3.1. Distribusi Dataset

Analisis deskriptif terhadap dataset menunjukkan distribusi diagnosis yang tidak seimbang (imbalanced), sebagaimana tersaji pada Tabel 1 berikut.

Tabel 1. Distribusi Diagnosis pada Dataset Klinis

Diagnosis	Jumlah	Persentase	Rencana Terapi
Healthy	1.167	58,35%	Rest and fluids
Bronchitis	334	16,70%	Hospitalization and medication
Flu	292	14,60%	Medication and rest
Cold	163	8,15%	Rest and fluids
Pneumonia	44	2,20%	Hospitalization and medication
Total	2.000	100%	-

Kelas Bronchitis sebagai target prediksi primer mencakup 334 rekam medis atau 16,70% dari total data. Seluruh kasus bronchitis dalam dataset memiliki label keparahan Severe dengan rencana terapi berupa rawat inap dan pemberian obat (Hospitalization and medication), yang mengindikasikan bahwa bronchitis dalam dataset ini merepresentasikan kondisi klinis yang memerlukan penanganan segera. Rata-rata usia pasien bronchitis adalah 47,6 tahun dengan rentang 18-79 tahun, menunjukkan bahwa penyakit ini menyerang lintas kelompok umur. Distribusi jenis kelamin pada kasus bronchitis relatif seimbang, dengan laki-laki sebanyak 176 kasus (52,7%) dan perempuan 158 kasus (47,3%). Parameter tanda vital rata-rata pada pasien bronchitis mencakup detak jantung 88,7 bpm, suhu tubuh 37,6°C, dan saturasi oksigen 91,9% nilai saturasi yang relatif lebih rendah dibandingkan rentang normal (95-100%), yang konsisten dengan karakteristik klinis bronchitis.

Gejala yang paling dominan pada kasus bronchitis adalah batuk (Cough), yang hadir pada 100% kasus bronchitis (334 dari 334 kasus). Ini merupakan temuan yang secara klinis valid dan konsisten dengan literatur medis yang menjadikan batuk sebagai gejala patognomonik bronchitis. Gejala-gejala penyerta lainnya yang sering muncul bersamaan dengan batuk pada pasien bronchitis meliputi runny nose (32,6%), body ache (31,1%), headache (30,5%), dan shortness of breath (28,4%).

3.2. Seleksi Nilai K Optimal

Eksperimen dilakukan pada delapan nilai K ganjil dengan menggunakan jarak Euclidean. Hasil perbandingan akurasi pada setiap nilai K disajikan pada Tabel 2 berikut.

Tabel 2. Perbandingan Akurasi KNN pada Berbagai Nilai K

Nilai K	Akurasi	Jarak	Skor	Keterangan
1	89,75%	Euclidean	0,8975	-
3	91,50%	Euclidean	0,9150	-
5	92,50%	Euclidean	0,9250	Optimal
7	91,50%	Euclidean	0,9150	-
9	89,50%	Euclidean	0,8950	-
11	88,50%	Euclidean	0,8850	-
13	89,50%	Euclidean	0,8950	-
15	88,75%	Euclidean	0,8875	-

Berdasarkan tabel di atas, nilai $K=5$ menghasilkan akurasi tertinggi sebesar 92,50%. Akurasi meningkat dari $K=1$ (89,75%) menuju $K=5$, kemudian menurun secara bertahap pada K yang lebih besar. Fenomena ini mengkonfirmasi prinsip bias-variance tradeoff dalam algoritma KNN: nilai K yang terlalu kecil menyebabkan model overfitting terhadap data latih, sementara nilai K yang terlalu besar mengakibatkan underfitting karena mempertimbangkan tetangga yang terlalu jauh dan tidak representatif. Nilai $K=5$ ditetapkan sebagai parameter optimal untuk seluruh analisis selanjutnya.

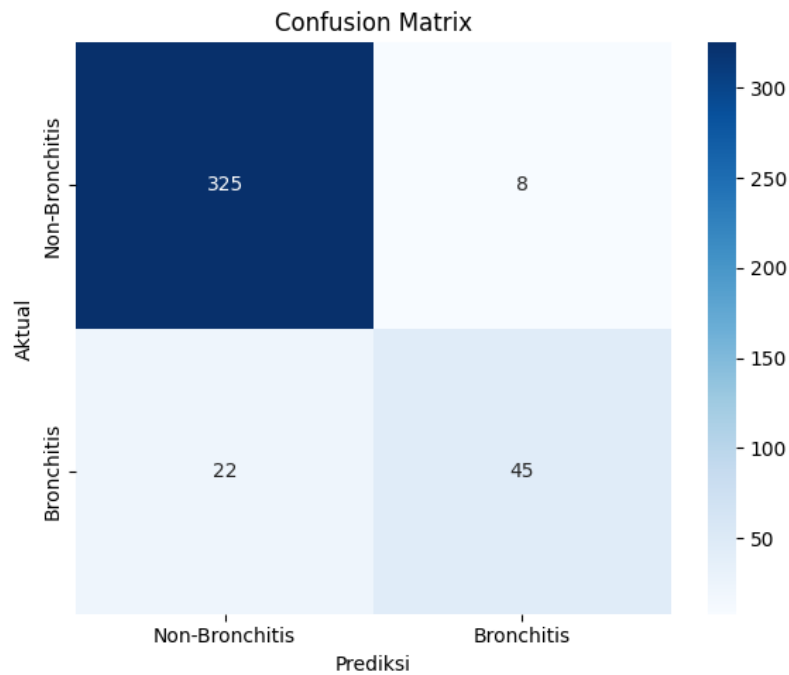
3.3. Performa Model KNN dengan $K=5$

Tabel 3 menyajikan laporan klasifikasi lengkap dari model KNN dengan $K=5$ pada data uji.

Tabel 3. Laporan Klasifikasi Model KNN ($K=5$, Data Uji $n=400$)

Kelas	Precision	Recall	F1-Score	Support
Non-Bronchitis	0,94	0,98	0,96	333
Bronchitis	0,85	0,67	0,75	67
Macro avg	0,89	0,82	0,85	400
Weighted avg	0,92	0,93	0,92	400

Model mencapai akurasi keseluruhan sebesar 92,50% pada data uji. Kelas Non-Bronchitis memperoleh precision 0,94 dan recall 0,98, mengindikasikan bahwa model sangat andal dalam mengidentifikasi pasien yang tidak menderita bronchitis. Untuk kelas Bronchitis, precision mencapai 0,85 dengan recall 0,67, menghasilkan F1-score sebesar 0,75. Tabel 4 menampilkan distribusi confusion matrix yang lebih rinci.



Gambar 2. Confusion Matrix Model KNN (K=5)

Model berhasil mengklasifikasikan 325 dari 333 kasus Non-Bronchitis secara benar (True Negative) dengan hanya 8 kesalahan klasifikasi sebagai Bronchitis (False Positive). Pada kelas Bronchitis, model berhasil mendeteksi 45 dari 67 kasus aktual (True Positive) dengan 22 kasus yang luput terdeteksi (False Negative).

3.4. Pembahasan

Hasil penelitian menunjukkan bahwa algoritma KNN dengan nilai K=5 mampu mencapai akurasi 92,50% dalam prediksi bronchitis berdasarkan kemiripan pola gejala dan tanda vital. Pencapaian ini sebanding dengan studi-studi terdahulu yang mengaplikasikan KNN pada dataset klinis respirasi, seperti yang dilaporkan oleh Pratama dkk. [3] dengan akurasi 89% untuk klasifikasi pneumonia, dan Rahman dkk. [9] dengan akurasi 91% untuk deteksi diabetes menggunakan KNN.

Nilai recall sebesar 67% untuk kelas bronchitis merupakan aspek yang perlu dicermati lebih lanjut. Dalam konteks klinis, recall yang lebih tinggi sangat diutamakan karena false negative kegagalan model mendeteksi kasus bronchitis yang sebenarnya berpotensi mengakibatkan keterlambatan penanganan yang dapat memperburuk kondisi pasien. Nilai recall 67% mengindikasikan bahwa sekitar sepertiga kasus bronchitis masih luput teridentifikasi oleh model. Fenomena ini kemungkinan besar disebabkan oleh ketidakseimbangan kelas (class imbalance) yang signifikan dalam dataset, di mana kelas Non-Bronchitis mendominasi dengan rasio 83,3% terhadap 16,7% kelas Bronchitis. Kondisi ini menyebabkan model cenderung bias ke arah kelas mayoritas.

Dominansi gejala batuk (Cough) pada 100% kasus bronchitis seharusnya menjadi fitur diskriminatif yang kuat. Namun, batuk juga hadir dengan frekuensi tinggi pada kelas lain seperti Flu dan Cold. Ini menjelaskan mengapa model memerlukan kombinasi fitur gejala sekunder dan parameter tanda vitalkhususnya saturasi oksigen yang rata-rata 91,9% pada bronchitis versus lebih tinggi pada kelas lain untuk membuat keputusan klasifikasi yang akurat. Hal ini mengkonfirmasi pentingnya integrasi fitur multidimensi yang diterapkan dalam penelitian ini.

Precision yang tinggi untuk kelas Bronchitis (0,85) menunjukkan bahwa ketika model memprediksi bronchitis, prediksi tersebut benar dalam 85% kasus. Artinya, model memiliki tingkat false alarm yang relatif rendah, yang penting untuk menghindari overdiagnosis dan pengobatan yang tidak perlu. Dari perspektif klinis, keseimbangan antara precision tinggi dan recall yang masih perlu ditingkatkan menjadikan model ini lebih cocok digunakan sebagai alat skrining awal yang dikombinasikan dengan pemeriksaan lanjutan oleh tenaga medis profesional, bukan sebagai pengganti diagnosis definitif.

Performa model ini dapat ditingkatkan melalui beberapa pendekatan. Teknik oversampling seperti SMOTE (Synthetic Minority Over-sampling Technique) dapat diterapkan untuk menangani ketidakseimbangan kelas. Selain itu, penggunaan bobot berbasis jarak (distance-weighted KNN) dan eksplorasi metrik jarak alternatif seperti Manhattan atau Minkowski dapat menjadi arah optimasi selanjutnya. Pendekatan ensemble berbasis KNN atau kombinasi dengan algoritma lain juga berpotensi meningkatkan recall tanpa mengorbankan precision yang telah tercapai.

4. Conclusion

Penelitian ini berhasil membangun model prediksi bronchitis menggunakan algoritma K-Nearest Neighbor yang mengintegrasikan fitur gejala klinis berbasis one-hot encoding dengan parameter tanda vital yang dinormalisasi secara komprehensif. Melalui eksperimen sistematis pada delapan nilai K, ditemukan bahwa nilai K=5 dengan jarak Euclidean menghasilkan performa optimal dengan akurasi 92,50%, precision 85%, recall 67%, dan F1-score 75% untuk kelas bronchitis. Temuan ini membuktikan bahwa pendekatan KNN berbasis kemiripan pola gejala efektif diterapkan untuk deteksi bronchitis dalam konteks dataset klinis multivariat.

Secara praktis, model yang dihasilkan berpotensi diimplementasikan sebagai modul skrining awal dalam sistem informasi kesehatan di fasilitas pelayanan kesehatan primer, khususnya untuk membantu tenaga medis mengidentifikasi pasien berisiko bronchitis secara lebih cepat berdasarkan gejala yang dilaporkan. Hal ini relevan terutama di daerah dengan keterbatasan akses terhadap fasilitas diagnostik lanjutan. Untuk penelitian selanjutnya, disarankan untuk mengeksplorasi teknik penanganan ketidakseimbangan kelas seperti SMOTE guna meningkatkan nilai recall, menguji pendekatan distance-weighted KNN, serta memvalidasi model pada dataset klinis nyata dari institusi kesehatan untuk mengukur generalisabilitasnya dalam skenario klinis yang sesungguhnya.

References

- [1] T. N. Ali, "Classification of normal and relaxed conditions based on brain signal activity with Electroencephalography using k-Nearest Neighbor (kNN)," *JHMT*, 2024, [Online]. Available: <https://journal.innoscientia.org/index.php/jhmt/article/view/23>
- [2] H. Apringsih *et al.*, "Peningkatan Pengetahuan Petugas Kesehatan Mengenai Peran Spirometri Sebagai Alat Deteksi Dini Penyakit Paru Obstruktif Kronik (PPOK) di Fasilitas Kesehatan Layanan Primer," *Smart Soc. Empower. J.*, vol. 4, no. 1, pp. 28–37, 2024, doi: 10.20961/ssej.v4i1.80196.
- [3] D. Lestari, W. T. Sudaryanto, and F. Sulistyowati, "Manajemen Fisioterapi pada Pasien dengan Kondisi Chronic Obstructive Pulmonary Disease (COPD): Case Report," *Borneo Nurs. J.*, vol. 8, pp. 2179–2186, 2026, doi: 10.61878/bnj.v8i2.791.
- [4] M. Bagir, N. Farkhatin, and N. Frastian, "SISTEM PAKAR DIAGNOSA PENYAKIT INFEKSI SALURAN PERNAPASAN ATAS PADA KLINIK DHARMA MEDICA MENGGUNAKAN METODE FORWARD CHAINING," *J. Ris. dan Apl. Mhs. Inform.*, vol. 06, no. 02, pp. 327–336, 2025, doi: 10.30998/jrami.v6i02.11777.
- [5] G. R. U. Architectures, F. R. Chan, R. Yanti, A. Hidayat, and A. Ramadhanu, "Deep Learning-Based Forecasting of Significant Wave Height in Padang Coastal Waters : A Comparative Study of RNN, LSTM, and GRU Architectures," *J. Comput. Res. Innov.*, vol. 11, no. 1, 2026, doi: 10.24191/jcrinn.v11i1.587.
- [6] F. R. Chan, F. Annas, Y. E. Yuspita, and G. Darmawati, "Implementation of Convolutional Neural Networks (CNN) in An

- Emotion Detection System for Measuring Learning Concentration Levels,” *Knowbase Int. J. Knowl. Database*, vol. 04, no. 01, pp. 49–61, 2024, doi: 10.0.121.7/knowbase.v4i1.8429.
- [7] R. Dila, W. Wirdawati, F. R. Chan, J. Harnaranda, and R. Sovia, “PENDIAGNOSAAN PENYAKIT KARIES GIGI MENGGUNAKAN FUZZY EXPERT SYSTEM DENGAN KOMBINASI METODE FUZZY LOGIC TSUKAMOTO DAN CERTAINTY FACTOR,” *J. Mbs. Tek. Inform.*, vol. 9, no. 3, pp. 4958–4965, 2025, doi: 10.36040/jati.v9i3.13855.
- [8] D. Rianasari, M. N. Triana, M. R. Dewi, and Y. Astutik, “The classification of mushroom types using Naïve Bayes and principal component analysis,” *JISA (Jurnal Inform. ...)*, 2022, [Online]. Available: <https://trilogi.ac.id/journal/ks/index.php/JISA/article/view/1380>
- [9] F. R. Chan and A. Ramadhanu, “IMPLEMENTASI HYBRID INTELLIGENCE SYSTEM UNTUK KLASIFIKASI BIJI-BIJIAN DENGAN ALGORITMA PCA DAN KNN,” *INTI Nusa Mandiri*, vol. 19, no. 2, pp. 240–250, 2025, doi: 10.33480/inti.v19i2.6397.
- [10] A. Dinanti and J. Purwadi, “Analisis Performa Algoritma K-Nearest Neighbor dan Reduksi Dimensi Menggunakan Principal Component Analysis,” *Jambura J. Math.*, vol. 5, no. 1, pp. 155–165, 2023, doi: 10.34312/jjom.v5i1.17098.
- [11] S. Manullang, N. Kairani, and ..., “Analisis faktor penyebab penyakit jantung menggunakan metode principal component analysis (pca),” *J. Lebesgue ...*, 2024, [Online]. Available: <https://www.lebesgue.lppmbinabangsa.id/index.php/home/article/view/732>