

Perbandingan Algoritma Naive Bayes dan Decission Tree untuk Prediksi Penyakit Kanker Paru-Paru

M.Fadillah¹, Mulia Gusti Firmansyah², M.Khairuddin³, Lusiana Efrizoni⁴
Rahmaddeni⁵

^{1,2,3,4,5}Jurusan Teknik Informatika, Universitas Sains dan Teknologi Indonesia

e-mail: ¹ 2110031802107@sar.ac.id, ² 2110031802102@sar.ac.id,

³ 2110031802110@sar.ac.id, ⁴ lusiana@stmik-amik-riau.ac.id,

⁵rahmaddeni@sar.ac.id

Intisari

Dalam studi ini, kami membandingkan kinerja dua algoritma pembelajaran mesin, *Naïve Bayes* dan *Decission Tree*, untuk mendiagnosis penyakit paru-paru menggunakan kumpulan data kesehatan pasien. Tujuan utama dari studi ini adalah untuk mengevaluasi akurasi, presisi, recall, dan skor F1 dari kedua algoritma tersebut untuk menentukan metode yang lebih efektif dalam memprediksi penyakit paru-paru. Hasilnya menunjukkan bahwa algoritma klasifikasi pohon mengungguli *Naïve Bayes* dalam hal akurasi, mencapai 95% dalam pembagian 80:20, dibandingkan dengan akurasi 78% yang dicapai oleh *Naïve Bayes* pada data yang sama. Analisis lebih lanjut menunjukkan bahwa sebagian besar pasien dalam kumpulan data ini memiliki risiko tinggi dengan 365 pasien, diikuti oleh risiko dengan 332 pasien, dan risiko rendah dengan 303 pasien. Struktur pohon keputusan terbukti lebih efektif dalam menangani kompleksitas data dan menghasilkan prediksi yang lebih akurat. Hasil-hasil ini menunjukkan bahwa pohon keputusan merupakan metode yang lebih baik daripada *Naïve Bayes* untuk mendiagnosis penyakit paru-paru dan dapat memberikan dasar yang kuat untuk mengembangkan model pembelajaran mesin yang akurat untuk penelitian kesehatan di masa mendatang.

Kata kunci Klasifikasi, *Naive Bayes*, *Decission Tree*, Kanker Paru-paru, Pembelajaran Mesin

Abstract

In this study, we compared the performance of two machine learning algorithms, Naïve Bayes and Decission Tree, for diagnosing lung diseases using patient health datasets. The main objective of this study is to evaluate the accuracy, precision, recall, and F1 score of the two algorithms to determine which method is more effective in predicting lung diseases. The results showed

Perbandingan Algoritma Naive Bayes dan Decision Tree untuk Prediksi Penyakit Kanker Paru-Paru
(M.Fadillah, Mulia Gusti Firmansyah, M.Khairuddin, Lusiana Efrizoni, Rahmaddeni)

that the tree classification algorithm outperformed Naïve Bayes in terms of accuracy, reaching 95% in an 80:20 split, compared to the 78% accuracy achieved by Naïve Bayes on the same data. Further analysis showed that most patients in this dataset were high risk with 365 patients, followed by risk with 332 patients, and low risk with 303 patients. The decision tree structure proved to be more effective in handling the complexity of the data and produced more accurate predictions, improving efficiency by creating a new "Risk_Score". These results show that decision trees are a better method than Naïve Bayes for diagnosing lung diseases and can provide a solid foundation for developing accurate machine learning models for future health research.

Keywords Classification, Naive Bayes, Decision Tree, Lung Cancer, Machine Learning

PENDAHULUAN

Kanker paru paru merupakan salah satu jenis kanker yang paling mematikan di dunia, dengan angka kematian yang tinggi. Menurut laporan dari Organisasi Kesehatan Dunia (WHO), kanker paru merupakan penyebab kematian akibat kanker terbanyak di seluruh dunia, yang menewaskan lebih dari 1,8 juta orang setiap tahunnya (WHO, p. et al.2023). Situasi ini diperburuk oleh kenyataan bahwa banyak kanker paru-paru baru terdiagnosis pada stadium lanjut, sehingga peluang pengobatannya kecil. Kurangnya deteksi dini dan kesadaran akan gejala kanker paru sering kali menyebabkan pasien baru mencari pengobatan ketika kanker sudah mencapai stadium lanjut.

Untuk meningkatkan kemungkinan deteksi dini dan pengobatan yang lebih efektif, diperlukan penelitian dan metode prediksi yang lebih baik untuk mengetahui risiko kanker paru-paru. Pendekatan yang menjanjikan adalah dengan menerapkan teknik pembelajaran mesin pada analisis data medis untuk mengembangkan model prediktif yang dapat mendukung deteksi dini dan pencegahan kanker paru-paru. Pada banyak penelitian sebelumnya, regresi linier digunakan untuk memprediksi risiko kanker paru dan menunjukkan hasil yang akurat (Wahid, Nugroho, & Anshor, 2023) (Jurnal FKPT). Selain itu, jika

dibandingkan dengan algoritma XGBoost dan SVM untuk diagnosis kanker payudara, hasilnya menunjukkan bahwa XGBoost memiliki akurasi yang lebih tinggi (Andryan, Fajri, & Sulistyowati, 2022)(Jurnal E AKAKOM).

Dalam konteks tersebut, penelitian ini berupaya mengevaluasi kinerja dua algoritma mesin, yaitu *Naive Bayes* dan *Decission Tree*, dalam memprediksi risiko kanker paru-paru. Dengan membandingkan hasil kedua algoritma ini, diharapkan dapat ditemukan cara yang lebih efektif dalam mengetahui risiko kanker paru serta meningkatkan metode pencegahan dan deteksi dini.

LANDASAN TEORI

1. Kanker Paru-Paru

Kanker paru-paru adalah jenis kanker yang bermula di paru-paru dan berhubungan dengan kebiasaan merokok, meskipun dapat juga terjadi pada orang yang bukan perokok. Gejala kanker paru-paru biasanya tidak muncul sampai penyakit tersebut berada pada stadium lanjut. Oleh karena itu, deteksi dini sangat penting untuk meningkatkan peluang kesembuhan. Risiko lainnya termasuk paparan radon, asbes, dan polusi udara.

2. *Naïve Bayes*

Naive Bayes adalah algoritma klasifikasi yang didasarkan pada metode Bayes dan asumsi independensi antar fitur. Algoritma ini bekerja dengan baik untuk dataset yang besar dan dapat menangani ketidakpastian dengan baik. Keuntungan utama *Naive Bayes* adalah kesederhanaannya dan kemampuannya untuk bekerja dengan baik meskipun asumsi independensinya sering salah dalam praktik

3. *Decission Tree*

Decision Tree adalah algoritma yang menggunakan struktur pohon untuk membuat keputusan berdasarkan aturan yang berasal dari karakteristik suatu set data . Algoritma ini mudah dipahami dan dijelaskan serta banyak digunakan dalam berbagai aplikasi klasifikasi. Keuntungan utama dari *decision tree* adalah kemampuannya untuk menangani data dengan nilai numerik dan kategoris serta kemampuannya untuk menangani nilai null.

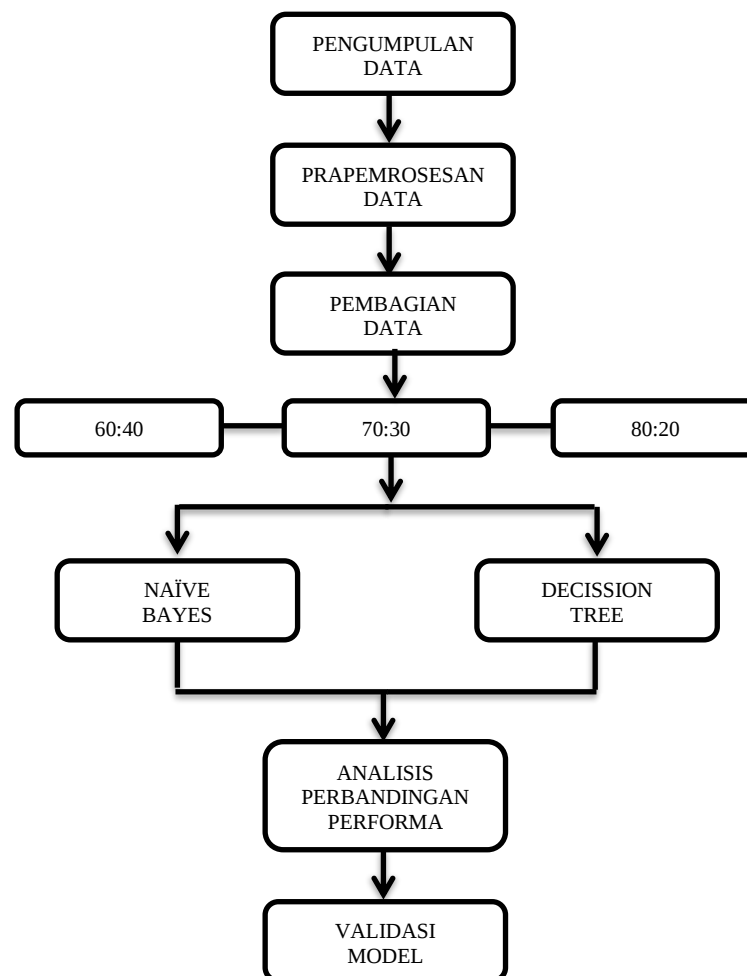
4. Prapemrosesan Data

Perbandingan Algoritma Naive Bayes dan Decision Tree untuk Prediksi Penyakit Kanker Paru-Paru
(M.Fadillah, Mulia Gusti Firmansyah, M.Khairuddin, Lusiana Efrizoni, Rahmaddeni)

Prapemrosesan data merupakan tahap pertama dari proses analisis data yang bertujuan untuk menyiapkan data mentah menjadi data yang siap digunakan untuk pengolahan atau analisis lebih lanjut. Langkah ini sangat penting karena kualitas data akan mempengaruhi hasil analisis dan keakuratan model yang dibangun.

METODE PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan yang menggambarkan urutan logis untuk memperoleh keluaran yang sesuai dengan harapan dan gambaran sistem. Dalam penelitian ini dilakukan nya 3 pemisahan data yaitu 60:40, 70:30, 80:20. Tahapan-tahapan tersebut dapat dilihat sebagai berikut :



Gambar 1.Metode Penelitian

1. Pengumpulan Data

Data dikumpulkan dari dataset yang tersedia di Kaggle, Data ini memiliki potensi besar untuk analisis prediktif kanker paru-paru menggunakan metode seperti Naive Bayes atau Decision Tree untuk mengidentifikasi faktor risiko atau penanda yang berkontribusi terhadap prediksi penyakit. Dataset ini bisa dilihat pada tabel 1 berikut :

Tabel 1.Dataset

index	Patient Id	Age	Gender	Air Pollution	Alcohol use	Genetic Risk	Smoking	...	Level
0	P1	33	1	2	4	3	3	...	Low
1	P10	17	1	3	1	4	2	...	Medium
2	P100	35	1	4	5	5	2	...	High
3	P1000	37	1	7	7	6	7	...	High
4	P101	46	1	6	8	7	8	...	High
...	...	...	...	...	...	...	...	...	...
997	P997	25	2	4	5	5	2	...	High
998	P998	18	2	6	8	7	8	...	High
999	P999	47	1	6	5	5	2	...	High

2. Prapemrosesan Data

Setelah data terkumpul, selanjutnya dilakukan proses preprocessing data untuk memastikan data siap digunakan dalam analisis. Tahapan ini meliputi :

- Penanganan missing values : Identifikasi dan isi nilai yang kosong menggunakan metode seperti ekspresi rata-rata atau ekspresi median.
- Normalisasi Data : Sesuaikan fitur numerik untuk memastikan bahwa semua fitur berada pada skala yang sama, menggunakan metode seperti penskalaan Min-Max atau kalibrasi.
- Visualisasi Distribusi Data : Distribusi data yaitu proses normalisasi yang menunjukkan bagaimana data ditransformasikan hingga memiliki persyaratan tertentu, seperti mean 0 dan standar deviasi 1. Normalisasi bertujuan untuk menjadikan data lebih baik dan membantu mempelajari algoritma dan mengotomatisasi agar bekerja secara efisien.

3. Pembagian Data

Pembagian data set dilakukan dengan menentukan nilai data latih dan data uji, dengan dilakukan nya pemisahan data 60:40, 70:30, 80:20.

4. Implementasi Model

- *Naïve Bayes*

Pada tahap ini, kami menggunakan algoritma *Naïve Bayes* untuk membangun model prediktif. Langkah pertama adalah mengimpor data yang terkumpul. Setelah itu, dilakukan pengolahan data , yang meliputi penanganan nilai null dan penyesuaian karakteristik numerik. Kumpulan data dibagi menjadi data latih dan data analisis dengan berbagai skenario split, seperti 60:40, 70:30, dan 80:20. Model *Naïve Bayes* dilatih menggunakan data latih dan kinerja model dievaluasi menggunakan data uji untuk memperoleh metrik seperti akurasi, presisi, recall, dan skor F1.

- *Decission Tree*

Selain *Naïve Bayes*, algoritma pohon keputusan juga diterapkan untuk membangun model prediktif. Proses implementasi dimulai dengan set data yang sama seperti pada *Naïve Bayes*. Setelah itu, set data dibagi menjadi data pelatihan dan data analisis dengan kondisi klasifikasi yang sama. Model pohon keputusan dilatih menggunakan data pelatihan dan kinerja model dievaluasi menggunakan data uji untuk mencapai hasil yang sama seperti *Naïve Bayes*.

5. Analisis Perbandingan Performa

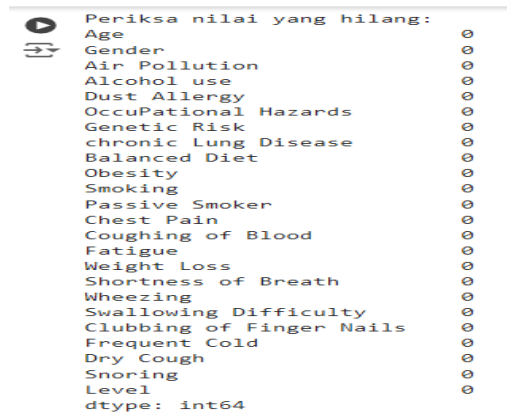
Analisis perbandingan kinerja dilakukan untuk mengevaluasi kekuatan dan kelemahan masing-masing algoritma (*Naïve Bayes* dan *Decision Tree*). Perbandingan dilakukan berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan skor F1 yang diperoleh dari data uji.

6. Validasi Model

Pengujian model dilakukan untuk mengukur kinerja dan efektivitas model yang telah dibangun. Dalam penelitian ini, evaluasi model mencakup beberapa metode dan metrik evaluasi penting, seperti akurasi, presisi, *recall*, dan skor F1.

HASIL DAN PEMBAHASAN

Setelah dilakukan nya pengumpulan data maka dilanjutkan dengan tahap prapemrosesan data, hasil dari prapemrosesan data disajikan pada gambar berikut:



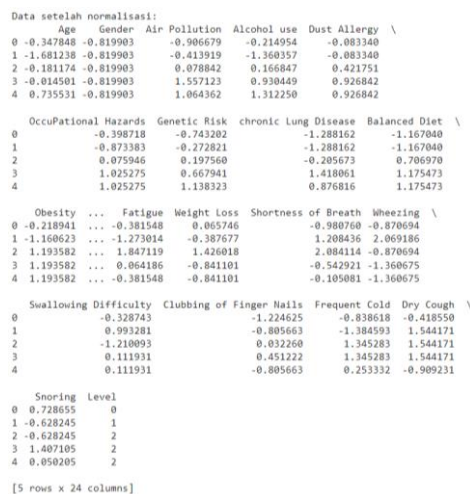
```

Periksa nilai yang hilang:
Age 0
Gender 0
Air Pollution 0
Alcohol use 0
Dust Allergy 0
OccuPational Hazards 0
Genetic Risk 0
chronic Lung Disease 0
Balanced Diet 0
Obesity 0
Smoking 0
Passive Smoker 0
Chest Pain 0
Coughing of Blood 0
Fatigue 0
Weight Loss 0
Shortness of Breath 0
Wheezing 0
Swallowing Difficulty 0
Clubbing of Finger Nails 0
Frequent Cold 0
Dry Cough 0
Snoring 0
Level 0
dtype: int64

```

Gambar 2.*Missing Value*

Hasil ini menunjukkan bahwa tidak ada nilai null pada kolom mana pun di dataset. Selanjutnya dilakukan tahap Normalisasi data. hasil dari normalisasi data disajikan pada gambar berikut :



```

Data setelah normalisasi:
Age Gender Air Pollution Alcohol use Dust Allergy \
0 -0.347848 -0.819903 -0.906679 -0.214954 -0.083340
1 -1.681238 -0.819903 -0.413919 -1.360357 -0.083340
2 -0.181174 -0.819903 0.078842 0.166847 0.421751
3 -0.014501 -0.819903 1.557123 0.930449 0.926842
4 0.735531 -0.819903 1.064362 1.312250 0.926842

OccuPational Hazards Genetic Risk chronic Lung Disease Balanced Diet \
0 -0.398718 -0.743202 -1.288162 -1.167040
1 -0.873383 -0.272821 -1.288162 -1.167040
2 0.075946 0.197560 -0.205673 0.706970
3 1.025275 0.667941 1.418061 1.175473
4 1.025275 1.138323 0.876816 1.175473

Obesity ... Fatigue Weight Loss Shortness of Breath Wheezing \
0 -0.218941 ... -0.381548 0.065746 -0.980768 -0.870694
1 -1.160623 ... -1.273014 -0.387677 1.208436 2.069186
2 1.193582 ... 1.847119 1.426018 2.084114 -0.870694
3 1.193582 ... 0.064186 -0.841101 -0.542921 -1.360675
4 1.193582 ... -0.381548 -0.841101 -0.105081 -1.360675

Swallowing Difficulty Clubbing of Finger Nails Frequent Cold Dry Cough \
0 -0.328743 -1.224625 -0.838618 -0.418550
1 0.993281 -0.805663 -1.384593 1.544171
2 -1.210093 0.032260 1.345283 1.544171
3 0.111931 0.451222 1.345283 1.544171
4 0.111931 -0.805663 0.253332 -0.909231

Snoring Level
0 0.728655 0
1 -0.628245 1
2 -0.628245 2
3 1.407105 2
4 0.050205 2

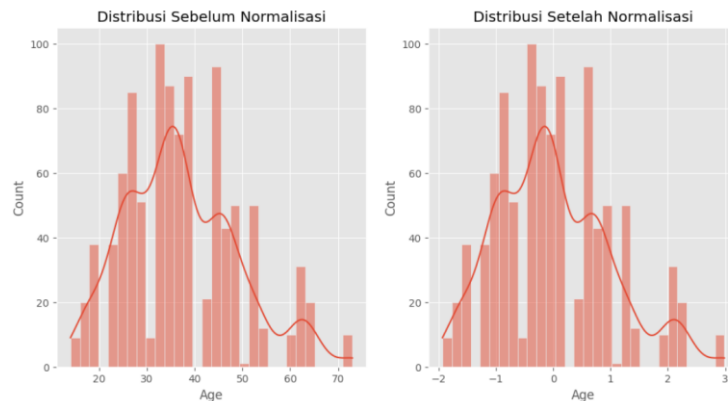
[5 rows x 24 columns]

```

Gambar 3.*Normalisasi Data*

Setelah dilakukan normalisasi, nilai setiap kolom (pengukuran) diubah sehingga mean kolom 0 dan simpangan baku menjadi 1. Nilai fitur Usia yang pertama adalah -0,347848 yang menunjukkan bahwa usia seseorang lebih rendah dari rata-rata usia dataset. Tahap selanjutnya ialah melakukan proses distribusi data, bisa dilihat pada gambar dibawah ini :

Perbandingan Algoritma Naive Bayes dan Decision Tree untuk Prediksi Penyakit Kanker Paru-Paru
(M.Fadillah, Mulia Gusti Firmansyah, M.Khairuddin, Lusiana Efrizoni, Rahmaddeni)



Gambar 4. Visualisasi Distribusi Data

Pada hasil diatas dapat dilihat perbedaan dari fitur *Age*, sebelum disesuaikan, data usia mempunyai sebaran yang mungkin tidak normal. Setelah pemrosesan yang benar, distribusi data usia diubah sehingga rata-rata berada di sekitar 0 dan standar deviasi sekitar 1. Maka tahap selanjutnya menerapkan data latih dan data uji dengan menggunakan beberapa splitting data diantaranya 60:40, 70:30, dan juga 80:20, berikut ditampilkan hasil nya :

```
Ukuran Training set (60:40): (600, 23)
Ukuran Test set (60:40): (400, 23)
Ukuran Training set (70:30): (700, 23)
Ukuran Test set (70:30): (300, 23)
Ukuran Training set (80:20): (800, 23)
Ukuran Test set (80:20): (200, 23)
```

Gambar 5. Data Latih & Data Uji

Berdasarkan hasil dari data latih dan data uji dengan menggunakan splitting data 60:40, 70:30, dan juga 80:20, langkah selanjutnya dilakukan penerapan model *Naïve Bayes* dan juga *Decision Tree* berikut disajikan hasil dari implementasi *Naïve Bayes* dan juga *Decision Tree* :

- *Naïve Bayes* :

```
# Menghitung jumlah penyakit kanker paru dengan tingkat Low, Medium, dan High
low_count = df[df["Level"] == "Low"].shape[0]
medium_count = df[df["Level"] == "Medium"].shape[0]
high_count = df[df["Level"] == "High"].shape[0]

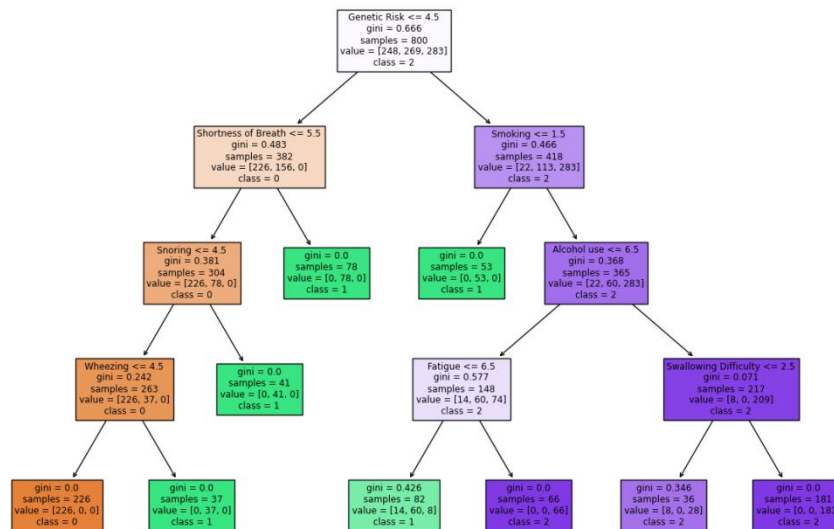
print("Jumlah pasien dengan tingkat kanker paru Low: ", low_count)
print("Jumlah pasien dengan tingkat kanker paru Medium: ", medium_count)
print("Jumlah pasien dengan tingkat kanker paru High: ", high_count)
```

Jumlah pasien dengan tingkat kanker paru Low: 303
Jumlah pasien dengan tingkat kanker paru Medium: 332
Jumlah pasien dengan tingkat kanker paru High: 365

Gambar 6. Hasil *Naïve Bayes*

Dari hasil tersebut, dapat disimpulkan bahwa mayoritas pasien dalam kumpulan data ini memiliki risiko tinggi sebanyak 365 pasien, diikuti oleh tingkat risiko medium sebanyak 332 pasien, dan tingkat risiko lemah sebanyak 303.

- *Decission Tree*



Gambar 7.Hasil *Decission Tree*

Pohon keputusan menunjukkan bahwa *Shortness of Breath*, *Genetic Risk*, dan *Smoking* merupakan prediktor utama tingkat kanker paru-paru. Pasien dengan energi rendah masuk ke kelas bawah, sedangkan pasien dengan risiko genetik tinggi dan merokok masuk ke kelas atas. Fitur tambahan seperti mendengkur, minum, mengi, dan kelelahan juga memiliki pengaruh, tetapi dengan dampak yang lebih kecil. Model ini efektif dalam mengidentifikasi pola kesehatan yang terkait dengan tingkat keparahan kanker paru-paru, sehingga memberikan informasi untuk prediksi dan penanganan penyakit. Berikut juga ditampilkan untuk melihat klasifikasi dari *Naïve Bayes* dan *Decission Tree* :

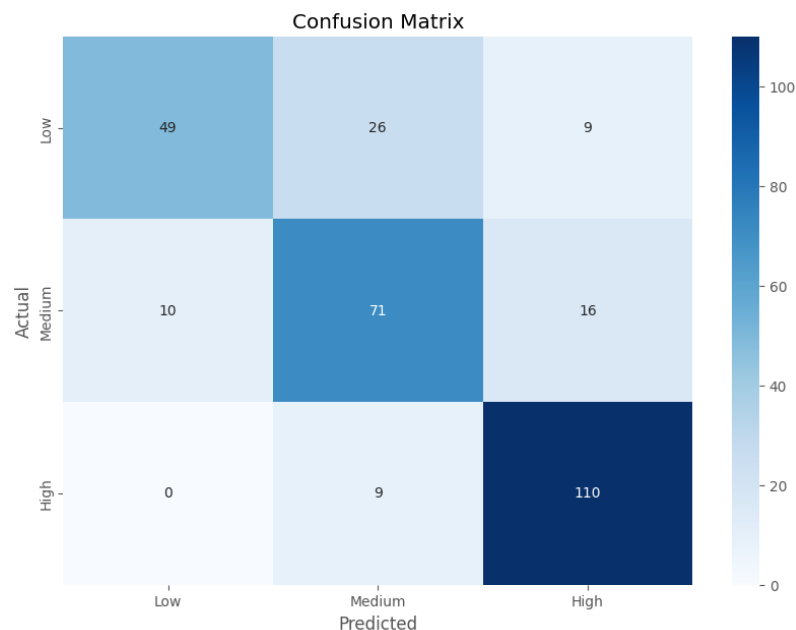
Perbandingan Algoritma Naive Bayes dan Decision Tree untuk Prediksi Penyakit Kanker Paru-Paru
(M.Fadillah, Mulia Gusti Firmansyah, M.Khairuddin,Lusiana Efrizoni, Rahmaddeni)

Tabel 2.*Classification Naïve Bayes*

Accuracy: 0.7667
Classification Report:

	precision	recall	f1-score	support
0	0.83	0.58	0.69	84
1	0.67	0.73	0.70	97
2	0.81	0.92	0.87	119
accuracy			0.77	300
macro avg	0.77	0.75	0.75	300
weighted avg	0.77	0.77	0.76	300

Model *Naive Bayes* memiliki kinerja yang baik dengan akurasi keseluruhan sebesar 76,67%. Model ini sangat efektif dalam mengklasifikasikan kelas "0" dengan presisi sebesar 0,83 dan recall sebesar 0,58. Untuk kelas "1" dan "2", presisi dan recall juga baik dengan presisi masing-masing sebesar 0,67 dan 0,81 serta recall sebesar 0,73 dan 0,9. Untuk hasil dari confusion matrix bisa dilihat pada gambar berikut :



Gambar 8.Implementasi Naïve bayes

Pada hasil matriks menggunakan 3 kelas, Low, Medium dan High, ditemukan bahwa model Naive Bayes menunjukkan kinerja yang baik pada klasifikasi kelas High dengan 110 prediksi benar dari 119. Model ini juga menunjukkan kinerja Low memiliki 49 prediksi benar dari 84.

- *Decission Tree* :

Hasil untuk klasifikasi pada model *Decision Tree* disajikan pada tabel berikut :

Tabel 3.*Classification Decission Tree*
Classification Report:

	precision	recall	f1-score	support
0	1.00	0.85	0.92	55
1	0.89	1.00	0.94	63
2	0.98	0.98	0.98	82
accuracy			0.95	200
macro avg	0.95	0.94	0.95	200
weighted avg	0.95	0.95	0.95	200

Pemodelan pohon keputusan menunjukkan kinerja yang sangat baik dengan akurasi 95%. Nilai presisi, recall, dan f1 untuk setiap kelas, yaitu rendah, sedang, dan tinggi, semuanya lebih besar dari 0,85. Kelas rendah memiliki presisi 1,00 dan recall 0,85 , kelas sedang memiliki presisi 0,89 dan recall 1,00 , dan kelas tinggi memiliki presisi dan recall 0,98 . Setelah memperoleh hasil dari klasifikasi Naïve Bayes dan Decission Tree selanjutnya di proses untuk perbandingan setiap hasil splitting data yang sudah ditentukan,hasil dari perbandingan setiap splitting data nya disajikan pada tabel berikut :

Tabel 4.Perbandingan *Naïve Bayes* dengan *Decission Tree*

Splitting Data	Akurasi		Precission		Recal		F1 Score	
	Naïve Bayes	Decission Tree	Naïve Bayes	Decission Tree	Naïve Bayes	Decission Tree	Naïve Bayes	Decission Tree
60:40	74,50%	97.75%	0.84	0.95	0.61	0.98	0.71	0.97
70:30	76,00%	96.33%	0.82	0.91	0.65	0.97	0.73	0.95
80:20	78,00%	98.5	0.81	0.95	0.69	0.99	0.75	0.98

Berdasarkan hasil perbandingan diatas , *Decission Tree* sering kali mengungguli Naïve Bayes dalam hal akurasi, presisi, perolehan kembali, dan skor F1 dalam distribusi data yang berbeda. Hal ini menunjukkan bahwa *Decission Tree* merupakan model klasifikasi yang lebih efektif untuk kumpulan data ini

KESIMPULAN

Dalam studi ini, kami membandingkan kinerja dua algoritma pembelajaran mesin, *Naïve Bayes* dan *Decission Tree*, pada kumpulan data yang terkait dengan diagnosis penyakit paru-paru. Hasilnya menunjukkan bahwa klasifikasi decision tree mengungguli klasifikasi Naïve Bayes dalam hal akurasi, presisi, recall, dan skor F1 di akhir pengujian pelatihan. Pengklasifikasi Decision Tree mencapai akurasi 95% dan klasifikasi 80:20, sedangkan pengklasifikasi Naïve Bayes mencapai akurasi 78% dalam klasifikasi yang sama. Desain decision tree terbukti efektif dalam menangani kompleksitas data dan membuat prediksi yang akurat. Proses identifikasi, termasuk pembuatan “Risk_Score” baru, juga meningkatkan efisiensi struktur decision tree.

Dari hasil ini, dapat disimpulkan bahwa mayoritas pasien dalam kumpulan data ini berisiko tinggi dengan 365 pasien, diikuti oleh risiko sedang dengan 332 pasien, dan risiko rendah dengan 303 pasien. Secara keseluruhan, studi ini menunjukkan keunggulan prosesor decision tree dibandingkan Naïve Bayes untuk diagnosis penyakit paru-paru. Hasil penelitian ini mungkin berguna dalam mengembangkan model pembelajaran mesin yang akurat untuk diagnosis penyakit paru-paru dan membantu meningkatkan hasil kesehatan.

SARAN

Untuk penerapan di dunia nyata, Decission Tree memberikan prediktabilitas yang sangat baik pada kanker paru-paru, namun bisa menjadi rumit untuk dipahami oleh orang non-teknis. Sebaliknya, Naive Bayes memberikan cara yang sederhana dan mudah untuk menafsirkan solusi, meskipun ukurannya sedikit lebih kecil. Oleh karena itu, pemilihan model harus disesuaikan dengan kebutuhan pengaplikasiannya, apakah mengarah pada presisi atau detail. Penelitian di masa depan mungkin mengeksplorasi algoritma lain seperti Random Forest atau Gradient Boosting Machines untuk membandingkan kinerja dengan Naive Bayes dan Decision Trees, dan meningkatkan model melalui iterasi hyper parameter dan strategi rekayasa untuk hasil yang lebih baik..

UCAPAN TERIMA KASIH

Kami mengucapkan terima kasih yang sebesar-besarnya kepada teman-teman, yang telah memberikan dukungan dan bantuan yang berharga selama proses penelitian ini. Saran, komentar dan dorongan Anda sangat berarti bagi kami. Kami ingin mengucapkan terima kasih sebesar-besarnya kepada Bpk. Rahmadeni, M. KOM, dan Ibu Lusiana Efrizoni, M.KOM pembimbing saya di Universitas Sains dan Teknologi Indonesia, atas bimbingan, arahan, dan dukungannya selama penulisan penelitian ini. Pengetahuan dan saran mereka membantu saya menyelesaikan penelitian ini. Tak lupa saya ucapkan terima kasih kepada seluruh sivitas akademika Universitas Sains dan Teknologi Indonesia yang telah menyediakan lingkungan yang baik untuk pembelajaran dan pengembangan. Kami berharap kerja keras dan upaya yang dilakukan dalam penelitian ini dapat membawa manfaat dan kontribusi bagi ilmu pengetahuan dan sivitas akademika.

Perbandingan Algoritma Naive Bayes dan Decision Tree untuk Prediksi Penyakit Kanker Paru-Paru
(M.Fadillah, Mulia Gusti Firmansyah, M.Khairuddin, Lusiana Efrizoni, Rahmaddeni)

DAFTAR PUSTAKA

- (Andryan and Fajri - 2022 - KOMPARASI KINERJA ALGORITMA XGBOOST DAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) UNTUK DIAGNOSA PENYAKIT KANKER PAYUDARA, Vol. 6, no. 1, Febuari 2022, hal. 1–5)
- (Dina Rachmawaty - 2022 - PENERAPAN METODE KLASIFIKASI DECISION TREE UNTUK MEMPREDIKSI KELULUSAN TEPAT WAKTU, Vol. 2, No. 1, Desember 2021, pp. 1-14)
- (Fauziningrum, M.Pd and Sulistyaningsih - 2021 - PENERAPAN DATA MINING METODE DECISION TREE UNTUK MENGUKUR PENGUASAAN BAHASA INGGRIS MARITIM (STUDI KASUS DI UNIVERSITAS MARITIM AMNI, Volume 22 Nomor 1, Bulan September Tahun 2021)
- (Hayuningtyas - 2020 - Penerapan Algoritma Naïve Bayes untuk Rekomendasi Pakaian Wanita, Vol. 6 No. 1 April 2020, pp. 18~22)
- (Heliyanti Susana - 2022 - PENERAPAN MODEL KLASIFIKASI METODE NAIVE BAYES TERHADAP PENGGUNAAN AKSES INTERNET, Vol 4, No.1, January 2022: Hal 1– 8)
- (Devina Larassati¹, Ati Zaidiah, Sarika Afrizal - 2022 SISTEM PREDIKSI PENYAKIT JANTUNG KORONER MENGGUNAKAN METODE NAÏVE BAYES, Volume 07, Nomor 02, Juni 2022 : 533–546)
- (Marzuq et al. - Prediksi Kanker Paru-Paru menggunakan Algoritme Random Forest Decision Tree, Vol. 7, No. 7, Juli 2023, hlm. 3448-3456)
- (Nurussakinah and Faisal - 2023 - Klasifikasi Penyakit Diabetes Menggunakan Algoritma Decision Tree, Vol. 10 No. 2 Oktober 2023, Halaman 143-149)
- (Pakarti - 2022 - Sistem Prediksi Lama Studi Kuliah Menggunakan Metode Naïve Bayes, Vol-2, Issue-1, 2022 (JIKOBIS))
- (Rahman Wahid et al. - 2023 - Prediksi Penyakit Kanker Paru-Paru Dengan Algoritma Regresi Linear, Vol 4, No 1, Maret 2023, Hal 63 - 74)
- (Rifai et al. - 2019 - Penerapan Algoritma Naïve Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist (MOS), Vol. 12, No. 2, September 2019)
- (Roihan et al. - 2020 - Pemanfaatan Machine Learning dalam Berbagai Bidang, IJCIT (Indonesian Journal on Computer and Information Technology) 5 (1) (2019) 75-8)

(Meza Silvana , Ricky Akbar , Alfi Syahnumc - 2020 - *Pemanfaatan Metode Naïve Bayes dalam Implementasi Sistem Pakar Untuk Menganalisis Gangguan Perkembangan Anak*, VOL. 06 NO. 02 (2020) 074-081)

(Sutoyo - 2018 - *IMPLEMENTASI ALGORITMA DECISION TREE UNTUK KLASIFIKASI DATA PESERTA DIDIK*, Vol. 14, No. 2 September 2018)

(Wardhana et al. - 2023 - *PENERAPAN MACHINE LEARNING DALAM PREDIKSI TINGKAT KASUS PENYAKIT DI INDONESIA*, Vol. 5, No. 1 (2023))

(Zudyanti Dwi - Penerapan Data Mining Untuk Prediksi Penyakit Diabetes Menggunakan Algoritma C4.5, Volume 4, Nomor 1, April 2024,)