

BAB 2

LANDASAN TEORI

2.1 Kolesterol Tinggi

Peningkatan kadar kolesterol dalam sirkulasi juga dapat mengakibatkan penyumbatan pada pembuluh darah. Obstruksi ini dapat menyebabkan peningkatan tekanan darah sistolik. Penumpukan kolesterol juga dapat mengakibatkan pembentukan plak dan aterosklerosis. Kadar kolesterol yang tinggi merujuk pada tingkat kolesterol dalam darah yang melebihi batas normal, terutama LDL atau "kolesterol jahat," yang dapat menyebabkan penumpukan plak pada dinding arteri (Husen dkk., 2022).. Hal ini meningkatkan risiko aterosklerosis, penyakit jantung koroner, stroke, dan masalah pembuluh darah perifer. Pemantauan rutin dan konsultasi dengan profesional kesehatan diperlukan untuk mengelola risiko dan mencegah komplikasi yang dapat timbul akibat kolesterol tinggi. Faktor-faktor yang menyebabkan kolesterol tinggi adalah :

1. Pola Makan:

Pola makan memainkan peran kunci dalam regulasi kolesterol darah. Makanan tinggi lemak jenuh dan kolesterol, seperti daging merah, makanan berlemak, dan produk olahan, dapat meningkatkan kadar kolesterol LDL (low-density lipoprotein) atau kolesterol jahat. Lemak trans, yang ditemukan dalam makanan olahan, juga dapat meningkatkan LDL dan menurunkan kadar HDL (high-density lipoprotein) atau kolesterol baik. Sebaliknya, makanan tinggi serat, seperti buah-buahan, sayuran, dan sereal gandum utuh, dapat membantu menurunkan kolesterol total.

2. Aktivitas Fisik:

Aktivitas fisik teratur memiliki dampak positif pada profil lipid. Olahraga aerobik, seperti berlari atau berenang, dapat meningkatkan kadar HDL, yang membantu menghilangkan kolesterol dari arteri. Sebaliknya, kurangnya aktivitas fisik dapat mengurangi efisiensi sistem *kardiovaskular* dalam mengelola kolesterol, meningkatkan risiko penumpukan plak di arteri.

3. Faktor Genetik:

Faktor genetik dapat memainkan peran signifikan dalam menentukan tingkat kolesterol seseorang. Jika ada riwayat keluarga dengan kolesterol tinggi atau masalah kesehatan *kardiovaskular*, individu tersebut mungkin memiliki kecenderungan genetik untuk memiliki kadar kolesterol yang lebih tinggi. Pengawasan kesehatan secara rutin dan tindak lanjut medis mungkin diperlukan untuk mengelola risiko yang terkait.

4. Berat Badan dan Obesitas:

Kelebihan berat badan atau obesitas seringkali terkait erat dengan tingkat kolesterol yang tidak sehat. Adiposit, sel lemak, dapat memproduksi kolesterol, terutama jenis LDL. Kelebihan berat badan dapat meningkatkan produksi LDL dan menurunkan kadar HDL, yang dapat menyebabkan peningkatan risiko penyakit jantung.

5. Kondisi Kesehatan:

Beberapa kondisi kesehatan, seperti diabetes dan hipotiroidisme, dapat mempengaruhi kadar kolesterol. Diabetes dapat merusak pembuluh darah, meningkatkan risiko penyumbatan arteri. Hipotiroidisme, produksi hormon tiroid yang rendah, dapat memperlambat metabolisme, mempengaruhi cara tubuh mengelola lemak dan kolesterol.

6. Faktor Hormonal:

Perubahan hormonal, terutama pada wanita selama menopause, dapat mempengaruhi tingkat kolesterol. Menopause seringkali dikaitkan dengan peningkatan kadar LDL dan penurunan kadar HDL. Hormon seks seperti estrogen memiliki peran dalam regulasi metabolisme lipid, dan fluktuasi hormon dapat mengubah distribusi kolesterol dalam tubuh.

7. Perilaku Merokok dan Konsumsi Alkohol:

Merokok dan konsumsi alkohol dapat mempengaruhi kesehatan *kardiovaskular* dan tingkat kolesterol. Zat-zat kimia dalam rokok dapat merusak dinding arteri dan meningkatkan produksi LDL. Alkohol dalam jumlah moderat dapat meningkatkan kadar HDL, tetapi konsumsi berlebihan dapat menyebabkan peningkatan risiko penyakit jantung.

8. Faktor Usia dan Jenis Kelamin:

Risiko kolesterol tinggi cenderung meningkat seiring bertambahnya usia. Proses penuaan dapat mempengaruhi metabolisme lipid dan distribusi kolesterol dalam tubuh. Selain itu, jenis kelamin juga berperan; wanita sebelum menopause memiliki kecenderungan memiliki kadar kolesterol yang lebih rendah dibandingkan dengan pria, tetapi setelah menopause, perbedaan ini dapat menyusut. Estrogen, hormon seks wanita, memiliki efek protektif pada kesehatan jantung, dan penurunannya selama menopause dapat berkontribusi pada peningkatan risiko kolesterol tinggi.

2.2 *Machine Learning*

Machine Learning dapat didefinisikan sebagai aplikasi komputer dan algoritme matematika yang diadopsi dengan cara pembelajaran dari data untuk menghasilkan prediksi di masa yang akan datang. Adapun proses pembelajaran melibatkan dua tahap, yaitu latihan (training) dan pengujian (testing). Bidang *Machine Learning* terkait dengan pertanyaan tentang bagaimana membangun program komputer agar dapat meningkat secara otomatis berdasarkan pengalaman. Penelitian terkini mengungkapkan bahwa *Machine Learning* terbagi menjadi tiga kategori: Supervised Learning, Unsupervised Learning, dan Reinforcement Learning (Homepage et al., 2019).

2.3 *Data Mining*

Data mining adalah ilmu yang mempelajari alur kerja untuk menemukan pengetahuan atau pola dalam kumpulan data besar dengan menganalisis data atribut yang saling berhubungan (Afif, 2020). Ini merupakan bagian dari kecerdasan buatan yang mencakup pengetahuan dan pengalaman, yang dapat digunakan oleh siapa pun untuk memecahkan masalah dan membuat keputusan (Hasanah et al., 2023). Proses data mining melibatkan ekstraksi informasi dari kumpulan data besar dengan menggunakan algoritme dan teknik ekstraksi dari bidang statistik, pembelajaran mesin, dan sistem manajemen basis data. Proses ini melibatkan analisis data dari berbagai perspektif untuk mengekstrak

informasi penting yang dapat digunakan untuk meningkatkan keuntungan, mengurangi biaya, atau keduanya. Secara teknis, data mining dapat diartikan sebagai proses menemukan korelasi atau pola dari ratusan atau ribuan bidang dalam sebuah basis data relasional yang besar. Intinya, data mining adalah pengumpulan data dalam jumlah besar untuk diekstraksi menjadi informasi yang dapat digunakan. Dalam konteks fungsional, data mining dapat dibagi menjadi beberapa kelompok, yaitu:

1. Prediksi, atau juga dikenal sebagai prediksi data, merupakan proses mengidentifikasi pola khusus dalam data yang dapat ditarik dari variabel yang ada dalam dataset tersebut. Tujuannya adalah memproyeksikan nilai variabel lain yang saat ini masih tidak diketahui. (Primadona,2023)
2. Deskripsi, atau sering disebut sebagai penjelasan, memberikan pemahaman yang lebih mendalam mengenai data yang diamati dan mengidentifikasi ciri-cirinya.(Primadona,2023)
3. Klasifikasi, yang juga dikenal sebagai pengelompokan, adalah proses pengolahan data untuk menemukan fungsi atau model tertentu dari konsep data. Proses ini digunakan untuk membagi setiap data ke dalam kelompok-kelompok yang berbeda.(Primadona,2023)
4. Fungsi asosiasi melibatkan penemuan kombinasi atau aturan asosiatif dari data yang akan diproses untuk mendapatkan informasi mengenai hubungan antar variabel (Primadona,2023).
5. Tujuan dari pengelompokan lebih spesifik, yaitu mengorganisasikan data ke dalam kelas-kelas dengan karakteristik yang serupa (Primadona,2023).

2.4 Prediksi

Prediksi adalah suatu proses sistematis untuk memperkirakan kejadian yang paling mungkin terjadi di masa depan. Hal ini didasarkan pada informasi dari masa lalu dan saat ini dengan tujuan untuk meminimalkan kesalahan, yaitu selisih antara hasil prediksi dengan kenyataan yang terjadi(Kafil, 2019). Algoritme seperti regresi, pohon keputusan, dan jaringan saraf dalam pembelajaran mesin digunakan untuk mengembangkan model prediktif.

Keberhasilan prediksi tergantung pada kualitas data, model yang dipilih, dan pemahaman domain yang mendalam. Hasil prediksi memberikan wawasan berharga untuk perencanaan strategis, pengambilan keputusan bisnis, dan formulasi kebijakan. Meskipun prediksi tidak dapat menjamin akurasi absolut, mereka memberikan pandangan yang berharga dalam menghadapi ketidakpastian, membantu organisasi dan individu merencanakan tindakan, dan mengoptimalkan strategi di berbagai sektor ekonomi, keuangan, kesehatan, dan lainnya.

2.5 *Naïve Bayes*

Naïve bayes, juga dikenal sebagai Simple Bayes dan Independent Bayes (Fitriyani, 2021), merupakan teknik prediksi berbasis probabilitas sederhana yang mengadopsi teori Bayes (atau aturan Bayes). Metode ini berdasarkan pada asumsi independensi yang kuat, atau sederhana (naif). Pengklasifikasi *Naïve Bayes* termasuk yang paling sederhana dan paling umum digunakan dalam praktek (Yulita dkk., 2021).Keunggulan utama dari *Naïve bayes* mencakup efisiensi, komputasi yang cepat, dan kebutuhan data yang relatif sedikit (Wardana & Prima Sari, 2023). Prediksi *Naïve bayes* didasarkan pada teori Bayes dengan formula umum. Dalam definisi lain, *Naïve bayes* dijelaskan sebagai metode klasifikasi yang menggunakan pendekatan probabilitas dan statistik yang pertama kali diperkenalkan oleh ilmuwan Inggris, Thomas Bayes. Metode ini digunakan untuk memprediksi peluang kejadian di masa depan berdasarkan pengalaman di masa sebelumnya.

$$P(H|X) = \frac{P(H)P(X)}{P(X)} \dots\dots\dots(1)$$

Keterangan :

1. X adalah data sampel dengan klas (label) yang belum diketahui.
2. H merupakan hipotesa X adalah data dengan kelas (label).
3. P(H) adalah peluang dari hipotesa H.
4. P(X) adalah peluang data sampel yang diamati.

5. $P(X|H)$ adalah peluang data sampel X , bila diasumsikan bahwa hipotesa benar

(valid).

Metode *Naive bayes* memerlukan beberapa petunjuk atau informasi sebagai penentu kelas yang sesuai untuk contoh yang sedang dianalisis. Oleh karena itu, metode ini disesuaikan sebagai berikut:

$$P(C|F_1..F_n) = \frac{P(C)P(F_1..F_n|C)}{P(F_1..F_n)} \dots\dots\dots (2)$$

2.6 Improved Naive Bayes

Improved Naive bayes adalah evolusi penting dalam dunia algoritme klasifikasi, menghadirkan peningkatan yang substansial terhadap pendekatan klasik *Naive bayes*. *Naive bayes*, sementara efisien, sering kali terhambat oleh asumsi ketergantungan fitur yang mungkin tidak sepenuhnya mencerminkan kompleksitas dunia nyata. Dalam menanggapi tantangan ini, *Improved Naive bayes* menyelipkan sejumlah strategi peningkatan yang mencakup teknik *smoothing*, seleksi fitur, dan pengenalan varian khusus.

Teknik *smoothing* menjadi kunci dalam meningkatkan ketepatan model *Naive bayes*. Ketika suatu fitur tidak muncul dalam kelas tertentu pada data pelatihan, probabilitas nol dapat menyebabkan ketidakstabilan. Dengan menerapkan teknik *smoothing* seperti Laplacian atau Lidstone *smoothing*, teknik ini berhasil menangani masalah probabilitas nol, menjaga stabilitas perhitungan, dan menghasilkan estimasi probabilitas yang lebih realistis.

Seiring dengan teknik *smoothing*, pendekatan seleksi fitur digunakan untuk mengoptimalkan akurasi model. Dengan memilih subset fitur yang paling informatif dan relevan, *Improved Naive bayes* dapat mengurangi dimensi data, mengurangi *noise*, dan meningkatkan keakuratan prediksi. Berbagai metode seleksi fitur, seperti Information Gain atau Recursive Feature Elimination, memberikan fleksibilitas dan keunggulan dalam menyesuaikan model dengan karakteristik unik dari beragam dataset.

Tidak hanya memperbaiki probabilitas dan mengoptimalkan fitur, *Improved Naive bayes* juga memperkenalkan varian khusus untuk menanggulangi tantangan tertentu. *Complement Naive bayes*, sebagai contoh, diarahkan untuk menangani ketidakseimbangan dataset, di mana frekuensi kelas-kelas berbeda secara signifikan. Dengan memberikan bobot lebih besar pada kelas minoritas, *Complement Naive bayes* berhasil meningkatkan akurasi model dalam kondisi ketidakseimbangan data.

Dalam konteks data kontinu, kemunculan *Naive bayes Gaussian* menandakan adaptabilitas *Improved Naive bayes* terhadap variasi jenis dataset. Dengan menggantikan asumsi distribusi multinomial dengan distribusi Gaussian, *Naive bayes Gaussian* lebih sesuai untuk dataset dengan fitur kontinu, seperti data sensor atau hasil pengukuran.

Melalui integrasi berbagai strategi ini, *Improved Naive bayes* tidak hanya menciptakan model klasifikasi yang lebih canggih dan adaptif, tetapi juga memperluas aplikasinya dalam berbagai konteks analisis data. Penerapan teknik smoothing, seleksi fitur, dan varian *Naive bayes* yang lebih spesifik membuka peluang baru untuk pengembangan teknik-teknik klasifikasi yang lebih unggul, menjadikan *Improved Naive bayes* sebagai pilihan utama dalam pemrosesan dan analisis data di berbagai sektor, dari ilmu pengetahuan hingga industri. Dengan demikian, *Improved Naive bayes* berfungsi sebagai landasan yang kuat untuk peningkatan kemampuan analisis data, membawa dampak positif dalam era ketidakpastian informasi yang terus berkembang.

2.7 Laplacian Correction

Analisis data dilakukan dengan menghitung probabilitas variabel pada masing-masing kelas. Jika terdapat kondisi dimana nilai probabilitas hasil adalah nol, maka koreksi Laplacian diterapkan (Gede Widyawati dkk., 2022). *Laplacian correction* yang juga dikenal sebagai *Lidstone smoothing*, adalah teknik *smoothing* yang digunakan dalam probabilitas dan statistika untuk menangani masalah probabilitas nol (Rizki dkk., 2021). Teknik ini berguna ketika memiliki data yang jarang atau ukuran sampel yang kecil, yang dapat

menyebabkan probabilitas nol untuk beberapa kejadian. Teknik Laplacian smoothing, yang juga dikenal sebagai metode add-one smoothing, melibatkan penambahan nilai 1 pada setiap frekuensi token yang ditemukan (Ningsih et al., 2023).

Rumus umum untuk *Laplacian correction* pada suatu probabilitas P dari suatu kejadian A dalam konteks klasifikasi adalah:

$$P(A) = \frac{m+1}{N+c} \dots\dots\dots (3)$$

Di mana:

- m adalah jumlah kejadian A yang terjadi dalam data.
- N adalah jumlah total kejadian dalam data.
- c adalah jumlah kategori atau kelas yang mungkin.

Dalam konteks klasifikasi dengan *Naive bayes*, seringkali *Laplacian correction* digunakan pada perhitungan probabilitas kondisional. Misalnya, probabilitas $P(X_i|C)$, di mana X_i adalah suatu fitur dan C adalah suatu kelas, dapat dihitung dengan menggunakan *Laplacian correction*:

$$P(X_i|C) = \frac{m+1}{N+c} \dots\dots\dots (4)$$

Di sini:

- m adalah jumlah observasi di kelas C di mana fitur X_i muncul.
- N adalah jumlah total observasi di kelas C .
- c adalah jumlah nilai yang mungkin untuk fitur X_i .

Penambahan nilai 1 pada pembilang dan c pada penyebut mencegah probabilitas nol dan memberikan bobot kecil pada setiap kemungkinan nilai, sehingga memastikan bahwa setiap kemungkinan kombinasi mendapatkan probabilitas yang lebih besar dari nol.

Laplacian correction membantu mengatasi masalah probabilitas nol, yang dapat menghambat kinerja model klasifikasi, terutama dalam kasus data yang langka atau ukuran sampel yang kecil. Meskipun cukup sederhana, teknik

ini membuktikan kegunaannya dalam menjaga kestabilan dan realisme probabilitas

2.8 Penelitian Terkait

Sebagai penguat topik penelitian, dilakukan beberapa analisis dari penelitian yang ada sebelumnya yang berkaitan dengan topik penelitian.

Berikut tabel 2.1 merupakan beberapa hasil dari penelitian Terkait :

Tabel 2.1 Daftar Penelitian Terkait

no	Nama Penulis dan tahun	Judul	Hasil Penelitian
1	Gilang Atala Panharsi, Indra Gita Anugrah. Tahun 2022	Klasifikasi Waktu Penyelesaian Skripsi Mahasiswa Menggunakan Metode Weighted Naïve Bayes (Studi Kasus: Program Studi Teknik Informatika Universitas Muhammadiyah Gresik)	Hasil klasifikasi skripsi mahasiswa Teknik Informatika di Universitas Muhammadiyah Gresik menunjukkan probabilitas penyelesaian dalam waktu 1 semester sebesar 0.14286, dalam waktu 2 semester sebesar 0.53968, dan lebih dari 2 semester sebesar 0.31746. Evaluasi model menunjukkan akurasi sebesar 77,5%, dengan presisi kelas 1S sebesar 100%, 2S sebesar 72,3%, dan 3S sebesar 82,4%. <i>Recall</i> kelas 1S mencapai 11,1%, 2S sebesar 100%, dan 3S sebesar 93,3%. Spesifisitas untuk kelas 1S sebesar 100%, 2S sebesar 75%, dan 3S sebesar 88%. Meskipun akurasi tinggi, perlu diperhatikan bahwa <i>recall</i> kelas 1S rendah,

			menunjukkan kurangnya sensitivitas model dalam mengidentifikasi mahasiswa yang dapat menyelesaikan skripsi dalam 1 semester. Evaluasi ini memberikan wawasan kritis yang dapat digunakan untuk perbaikan model (Atala Panhares et al., 2022).
Alyaa Nadira1, Nanang Yudi Setiawan, Welly Purnomo. Tahun 2023	Analisis Sentimen Pada Ulasan Aplikasi Mobile Banking Menggunakan Metode Naïve Bayes Dengan Kamus Inset	Penelitian ini menyimpulkan bahwa sistem klasifikasi Naïve Bayes dalam analisis sentimen ulasan aplikasi <i>Victoria Mobile Banking</i> memiliki performa tinggi, mencapai 95% tingkat ketepatan klasifikasi ulasan. Mayoritas ulasan positif menyoroti kemudahan penggunaan dan kepuasan transaksi, meskipun beberapa pengguna mengemukakan kebingungan terkait proses login, update, dan aktivasi akun. Secara keseluruhan, pengguna memberikan penilaian positif terhadap aplikasi. Feedback dari perusahaan menunjukkan bahwa visualisasi hasil analisis sentimen efektif dalam mengidentifikasi keluhan pengguna, memungkinkan	

			penyesuaian dan pengembangan aplikasi untuk meningkatkan kualitas layanan(5138-205-17885-2-10-20230517, n.d.).
3	Nilu Gede Widyawati, Maliyatul Khasanah, Muttaqin, Errissya Rasywir. Tahun 2022	Prediksi Tingkat Stress Pada Mahasiswa Universitas Dinamika Bangsa Jambi Dalam Melakukan Perkuliahan Metode Hybrid Menggunakan Algoritme <i>Naive Bayes</i>	Berdasarkan hasil penelitian, dapat disimpulkan bahwa penerapan perhitungan probabilitas <i>Naive bayes</i> pada data testing menghasilkan prediksi klasifikasi sebagai Klasifikasi Stres Tingkat Sedang, dengan nilai probabilitas sebesar 0,003840878. Meskipun demikian, algoritme <i>Naive bayes</i> menunjukkan tingkat akurasi yang masih rendah, yakni sebesar 68%, setelah dilakukan perbaikan dengan menggunakan teknik Laplace Correction. Rendahnya akurasi ini disebabkan oleh kurangnya variasi nilai pada berbagai variabel dalam data training.(Gede Widyawati et al., 2022)

4	Naomi Chatrina Siregar, Riki Ruli A. Siregar, M. Yoga Distra Sudirman. Tahun 2020	Implementasi Metode Naïve Bayes <i>Classifier</i> (Nbc) Pada Komentar Warga Sekolah Mengenai Pelaksanaan Pembelajaran Jarak Jauh (Pjj)	Berdasarkan penelitian yang dilakukan, dapat disimpulkan bahwa sistem yang dibangun mampu memberikan analisis komentar dengan tingkat akurasi yang lebih baik dibandingkan dengan proses manual. Proses klasifikasi komentar terkait pelaksanaan pembelajaran jarak jauh dimulai dengan tahap preprocessing teks, yang melibatkan transformasi teks menjadi kata tunggal, penghilangan kata tidak penting berdasarkan kamus data, dan penghapusan tanda baca. Selanjutnya, dilakukan perhitungan pembobotan TF-IDF dan penggunaan Naïve Bayes <i>Classifier</i> . Hasil pengujian pada 50 data komentar menggunakan algoritme Naïve Bayes <i>Classifier</i> menunjukkan tingkat akurasi sebesar 68%. Meskipun demikian, sistem ini tetap memberikan kontribusi signifikan dalam meningkatkan efisiensi analisis komentar terkait pembelajaran jarak jauh.(Hariyanto et al., 2021)
---	---	--	---

5	Heliyanti Susana, Nana Suarna, Fathurrohman, Kaslani.tahun 2022	Penerapan Model Klasifikasi Metode <i>Naive</i> <i>Bayes</i> Terhadap Penggunaan Akses Internet	Dari hasil penelitian ini, dapat disimpulkan bahwa penerapan model klasifikasi metode <i>Naive bayes</i> melibatkan beberapa langkah, seperti menginput data latih pada <i>RapidMiner</i> , melakukan <i>Data Selection & Cleaning</i> , melakukan <i>Data Transformation</i> , mengolah <i>Data Mining</i> , dan melakukan Interpretasi. Akurasi model yang diperoleh mencapai 89.83%, dengan hasil prediksi Ya dan ternyata Ya sebanyak 34, Ya dan ternyata Tidak sebanyak 6, Tidak dan ternyata Ya sebanyak 0, dan Tidak dan ternyata Tidak sebanyak 19. Pengujian dengan 59 data baru menunjukkan hasil Ya sebanyak 40 siswa dan Tidak sebanyak 19 siswa, mengkonfirmasi bahwa model <i>Naive bayes</i> efektif dalam prediksi dengan tingkat akurasi yang tinggi.(Susana & Suarna, 2022)
---	--	---	---

6	Ahmad Afif. Tahun 2020	Penerapan Algoritme Naïve Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus Di Rumah Sakit Aisyiah	Penelitian <i>Early Stage</i> Diabetes <i>Risk Prediction</i> menggunakan Algoritme Naïve Bayes dengan 16 atribut, termasuk <i>Age, Gender,</i> <i>Polyuria, Polydipsia, sudden</i> <i>weight loss, weakness, Polyphagia,</i> <i>Genital thrush, visual blurring,</i> <i>Itching, Irritability, delayed</i> <i>healing, partial paresis, muscle</i> <i>stiffness, Alopecia, Obesity,</i> dan <i>class.</i> Dari awalnya 520 data yang kotor, setelah pembersihan data dan seleksi atribut, didapatkan 500 data bersih. Model Naïve Bayes mencapai akurasi sebesar 90,20%, dengan Class Precision 82,35% dan Class <i>Recall</i> 87,50%. Evaluasi klasifikasi menunjukkan kinerja yang baik, namun disarankan untuk meningkatkan akurasi menggunakan metode seperti <i>Assemble.</i> Kurva ROC menunjukkan nilai AUC sebesar 0.955, menandakan klasifikasi yang sangat bagus.(Afif, 2020)
---	---------------------------	--	--

7	Winda Yulita, Eko Dwi Nugroho, Muhammad Habib Algifari. Tahun 2021	Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritme Naïve Bayes <i>Classifier</i>	Kesimpulan dari penelitian ini mengindikasikan bahwa sebagian besar masyarakat Indonesia, terutama pengguna media sosial Twitter, menunjukkan respon positif terhadap kebijakan vaksinasi Covid-19 di Indonesia. Persentase respon positif mencapai 60,3% (2278 respon). Penerapan algoritme <i>Naive bayes Classifier</i> dalam melakukan analisis sentimen telah berhasil, seiring dengan tingkat akurasi yang tinggi sebesar 93%.(Yulita et al., 2021)
8	Fitriyani. Tahun 2021	Prediksi Diabetes Menggunakan Algoritme Naïve Bayes Dan Greedy Forward Selection	Dalam penelitian ini, algoritme <i>Machine Learning</i> yang dipilih adalah Naïve Bayes karena kesederhanaannya. Eksperimen menunjukkan bahwa nilai akurasi terbaik diperoleh pada model Naïve Bayes dan Greedy Forward Selection (NB+GFS) sebesar 91,73%, sedangkan model Naïve Bayes sendiri mendapatkan nilai akurasi sebesar 87,69%. Dari hasil eksperimen ini, dapat disimpulkan bahwa kombinasi algoritme Naïve Bayes dan Greedy Forward Selection mampu memprediksi diabetes dengan sangat baik.(Fitriyani, 2021)

9	Muhammad Rizki, Muhammad Arhami, dan Huzeni. Tahun 2021	Perbaikan Algoritme <i>Naive Bayes Classifier</i> Menggunakan Teknik Laplacian Correction	Algoritme <i>Naive Bayes Classifier</i> mencapai tingkat akurasi yang tinggi, yaitu sebesar 94,44%, setelah mengalami perbaikan dengan menggunakan teknik Laplacian Correction. Penerapan teknik ini terbukti efektif dalam memperbaiki dan menghasilkan nilai probabilitas akhir yang semula bernilai 0 pada tahap awal.(Rizki et al., 2021)
10	Ramaulvi Muhammad Akhyar, Leonard Peter Gelu, Agustinus Ardiyanto Nena, Antonius Siki, Siriakus Lalang Lelang AyaQ. Tahun 2023	Penerapan Aloritma <i>Naive Bayes</i> Untuk Membantu Mahasiswa Dalam Memprediksi Pengambilan Mata Kuliah (Studi Kasus: Prodi Teknologi Informasi Unimor Angkatan Tahun 2020)	<i>Naive bayes</i> dapat efektif digunakan untuk memprediksi pengambilan mata kuliah oleh mahasiswa. Dalam pengujian menggunakan dataset mahasiswa tahun 2022 semester ganjil, <i>Naive bayes</i> menunjukkan tingkat akurasi yang cukup baik. Sebagai contoh, pada mata kuliah "Keamanan Sistem dan Data", nilai confusion matrix menunjukkan akurasi sebesar 21,21%, sedangkan pada mata kuliah "Audit Sistem Informasi", akurasi mencapai 78,26%. Dengan demikian, <i>Naive bayes</i> memberikan solusi yang efisien untuk permasalahan penjadwalan mata kuliah dan prediksi pilihan mata kuliah mahasiswa.(Akhyar, 2023)

11	Yuni Wardana, Devni Prima Sari. Tahun 2023	Prediksi Penyakit Diabetes Dengan Naïve Bayes	<p>Model Naïve Bayes yang digunakan dalam prediksi diabetes memberikan penjelasan tentang hubungan antara diabetes dan gejalanya. Menurut model ini, peluang seseorang terjangkit diabetes adalah sebesar 28%, sementara peluang untuk tidak terjangkit diabetes mencapai 72%.</p> <p>Hasil penelitian menunjukkan tingkat akurasi sebesar 94%, presisi 95%, dan recall 98%, menandakan bahwa model ini sangat efektif dalam mengidentifikasi kasus diabetes. (Wardana & Prima Sari, 2023)</p>
12	Ade Riani, Yessy Susianto, Nur Rahman. Tahun 2019	Implementasi Data Mining Untuk Memprediksi Penyakit Jantung Menggunakan Metode <i>Naive Bayes</i>	<p>Diagnosis penyakit jantung menghasilkan empat kemungkinan hasil: true positive (TP), true negative (TN), false positive (FP), dan false negative (FN). True positive dan true negative adalah hasil diagnosis yang benar, sementara false positive terjadi ketika pasien sehat tetapi diprediksi sakit, dan false negative terjadi ketika pasien sakit tetapi diprediksi sehat. Akurasi diagnosis dihitung dengan rumus $(TP + TN) / (TP + FP + TN + FN)$. Pada</p>

			dataset dengan total 303 hasil, akurasi mencapai 86%.(Riani et al., 2019)
13	Uswatun Hasanah, Asbon Hendra Azhar. Tahun 2023	Data Mining Menentukan Balita Yang Mendapatkan Makanan Tambahan Dari Kemenkes Dengan Metode <i>Naive Bayes</i> (Studi Kasus: Puskesmas Pembantu Bantan)	Hasil klasifikasi balita yang mendapatkan makanan tambahan di Puskesmas Pembantu Bantan menggunakan algoritme <i>Naive bayes</i> menunjukkan probabilitas hipotesis untuk kelas "Dapat" sebesar 0,073 dan kelas "Tidak Dapat" sebesar 0,002. Validasi confusion matrix menghasilkan Precision kelas "Dapat" sebesar 1, Precision kelas "Tidak Dapat" sebesar 0,964, Recall kelas "Dapat" 0,964, dan Recall kelas "Tidak Dapat" 1. Akurasi mencapai 98,1818%, dengan persentase <i>Correctly Classified Instance</i> sebesar 98,1818%, dan <i>Incorrectly Classified Instance</i> sebesar 1,8182%. Dari 55 data balita, 54 berhasil diklasifikasikan dengan benar, dan 1 data balita tidak berhasil diklasifikasikan dengan benar.(Hasanah et al., 2023)

14	Bambang Tri Rahmat Doni, Sari Susanti, Ade Mubarak. Tahun 2021	Penerapan Data Mining Untuk Klasifikasi Penyakit Hepatocellular Carcinoma Menggunakan Algoritme Naïve Bayes	<p>Dalam penelitian ini, algoritme Naïve Bayes diterapkan untuk mengklasifikasikan tingkat kemungkinan hidup pasien yang didiagnosis dengan <i>Hepatocellular Carcinoma</i> (HCC) berdasarkan probabilitas gejala penyakit yang diperoleh dari catatan medis.</p> <p>Performa algoritme dievaluasi dengan <i>Confusion Matrix</i> menggunakan HCC Survival Data Set, menghasilkan Akurasi sebesar 70,30%, Presisi 73,53%, dan Recall 77,32% dengan metode validasi <i>10-Fold Cross Validation</i>.</p> <p>Selain itu, hasil evaluasi menggunakan kurva ROC menunjukkan nilai AUC sebesar 0,783, termasuk dalam kategori <i>Fair Classification</i> atau Klasifikasi Cukup.(Tri et al., 2021)</p>
----	--	---	---

15	Firza Novaldy, Asti Herliana. Tahun 2021	Penerapan Pso Pada Naïve Bayes Untuk Prediksi Harapan Hidup Pasien Gagal Jantung	Hasil perhitungan menggunakan algoritme <i>Naive bayes</i> pada Heart <i>Failure Clinical Records</i> Dataset menunjukkan nilai akurasi pada confusion matrix sebesar 75,00%, dan nilai AUC sebesar 0,847. Setelah menerapkan Particle Swarm Optimization untuk seleksi fitur pada dataset, akurasi meningkat menjadi 91,67%, dan nilai AUC mencapai 0,908. Uji coba ini mengidentifikasi beberapa atribut yang berpengaruh terhadap bobot atribut, seperti age, anaemia, diabetes, ejection fraction, high blood pressure, serum sodium, sex, smoking, dan time. Dengan demikian, dapat disimpulkan bahwa penggunaan Particle Swarm Optimization efektif dalam meningkatkan akurasi algoritme <i>Naive bayes</i> pada Heart Failure Clinical Records Dataset. (Novaldy & Herliana, 2021)
----	---	--	---

Hasil studi literatur menunjukkan bahwa hingga saat ini, belum ada aplikasi prediksi kolesterol tinggi menggunakan metode *Improved Naive bayes* dengan Laplace Correction. Oleh karena itu, penelitian ini memberikan kontribusi signifikan dalam mengisi celah pengetahuan tersebut. Dengan menerapkan metode *Improved Naïve Bayes*, diharapkan penelitian ini dapat memberikan pemahaman yang lebih mendalam dan efektif dalam memprediksi

tingkat kolesterol tinggi, menyediakan landasan yang lebih kuat untuk identifikasi risiko dini, dan pada akhirnya mendukung upaya pencegahan penyakit *kardiovaskular*.

