



Metode Analisis Kesehatan Dengan Menggunakan Machine Learning Atau Artificial Inteligenci Atau Data Mining *Literature Review*

Rizki Agam Syahputra¹, Maulia Rahmi Hanifah¹

Program Studi Teknik Industri, Fakultas Teknik, Universitas Teuku Umar

*Email Korespondensi: rizkiagamsyahputra@utu.ac.id

Abstract - Kesehatan merupakan aspek penting dalam masyarakat dan menjadi pondasi bagi kemajuan suatu negara. Negara yang berhasil memperkuat sistem kesehatannya akan memiliki masyarakat yang sejahtera. Pemerintah seringkali memprioritaskan kesehatan dan kesejahteraan masyarakat dalam agenda pembangunannya. Salah satu upaya yang dilakukan adalah pembangunan dan peningkatan pelayanan serta fasilitas kesehatan guna memastikan kesehatan bangsa dan tercapainya negara yang sehat. Penelitian ini bertujuan untuk menyelidiki berbagai metode analisis kesehatan yang dapat dipergunakan untuk memprediksi atau menyelesaikan masalah kesehatan sesuai kebutuhan. Metode penelitian ini melibatkan kajian literatur dari berbagai artikel yang diterbitkan dalam beberapa tahun terakhir. Paper ini fokus pada analisis terhadap berbagai metode analisis kesehatan yang digunakan khususnya dalam penggunaan big data dan juga artificial intelligence. Berdasarkan analisis terhadap berbagai artikel jurnal, metode analisis kesehatan yang umum digunakan meliputi data mining, clustering, algoritma k-mean, naïve bayes, decision tree, support vector machine, neural network, k-nearest neighbor, rapidminer, algoritma C4.5, multiple linear regression, artificial intelligence, tensorflow, ssdmobilenet berbasis python, machine learning, classifier, backpropagation, prediksi, random forest, cross validation, dan confusion matrix. Meskipun beberapa metode tersebut umum digunakan, terdapat pula metode atau pendekatan lain yang dapat memberikan hasil yang lebih optimal dalam analisis kesehatan.

Katakunci: prediksi penyakit, analisis Kesehatan

Abstract - Health is an important factor in society and is one of the foundations of a country. A strong country has a healthy and prosperous society. Public health and welfare is one of the main focuses for government. Continuous development to improve health services and facilities is one of the government's efforts to build national health and create a healthy country. The aim of this research is to find out various health analysis methods that can be used to predict or solve other health problems for every need. This research uses a systematic method of literature review of various articles that have been published based on the year and method analyzed to focus on various methods of health analysis. From the analysis of various journal articles, this research shows that health analysis methods are generally used, such as data mining, clustering, k-mean algorithm, naïve Baye, decision tree, support vector machine, neural network, k-neuarets neighbor, rapidminer, C4 algorithm. 5, multiple linear regression, intuitive intelligence, tensorflow, ssdmobilenet based on python, machine learning, classifier, backpropogion, prediction, random forests, cross validation and confusion matrix. Some of these methods are common, but they also use other methods or approaches so that the results obtained are more optimal.

Keywords: disease prediction, health analysis

PENDAHULUAN

Kualitas hidup merupakan persepsi individu terhadap posisinya dalam kehidupan, yang dipengaruhi oleh budaya, sistem nilai, dan hubungannya dengan tujuan hidup, harapan, standar, dan aspek terkait lainnya. Isu yang berkaitan dengan kualitas hidup mencakup beragam aspek seperti masalah kesehatan fisik, status psikologis, tingkat kebebasan, serta hubungan sosial dan lingkungan tempat

individu berada (Jacob & Sanjaya, 2018).

Kesehatan merupakan hal yang sangat diinginkan oleh semua makhluk hidup di planet ini. Kondisi tubuh yang sehat sangat penting karena kondisi sakit dapat membuat seseorang tidak produktif dan meningkatkan risiko kematian (KKRI, 2024). Menurut para ahli, seperti yang diungkapkan oleh White (1977), seseorang dianggap sehat ketika tidak memiliki keluhan atau tanda-tanda penyakit saat diperiksa. Perkembangan teknologi telah memberikan dampak positif dalam kemajuan bidang kesehatan. Teknologi membantu tenaga medis dalam melakukan tindakan dan perawatan terhadap pasien, termasuk dalam penyediaan layanan informasi kesehatan. Hal ini memberikan kesempatan bagi masyarakat untuk lebih mandiri dan percaya diri terhadap kesehatan mereka. Kesehatan memiliki peran penting dalam masyarakat dan merupakan pondasi utama bagi kemajuan suatu negara. Pemerintah seringkali memprioritaskan kesehatan dan kesejahteraan masyarakat dalam pembangunan negara. Upaya terus dilakukan untuk meningkatkan pelayanan dan fasilitas kesehatan guna memastikan kesehatan masyarakat dan terwujudnya negara yang sehat. Fasilitas pelayanan kesehatan dianggap sebagai hak dasar masyarakat yang harus dipenuhi dalam upaya pembangunan kesehatan.

Machine learning adalah suatu metode yang digunakan untuk mengembangkan program yang dapat belajar dari data. Berbeda dengan program komputer konvensional yang statis, program machine learning dirancang agar dapat belajar sendiri (Wardhana dkk (2023)). Teknologi ini memiliki potensi besar untuk digunakan dalam berbagai bidang, termasuk di industri kesehatan. Penggunaan machine learning dalam perawatan kesehatan dapat meningkatkan kesejahteraan manusia dengan memberikan layanan yang lebih efektif dan berkualitas tinggi. Data mining adalah proses analisis data dari berbagai perspektif dengan tujuan untuk menghasilkan informasi yang berguna. Informasi yang dihasilkan dari data mining dapat digunakan untuk meningkatkan pendapatan dan mengurangi biaya, serta membantu dalam pengambilan keputusan. Teknik data mining ini saat ini banyak digunakan oleh perusahaan untuk mendukung aktivitas mereka. Untuk itu, paper ini bertujuan untuk menganalisa perkembangan keimuan, dan teknologi dari machine learning dan artificial intelligence terhadap aplikasinya dalam dunia kesehatan.

METODE PENELITIAN

Penelitian ini menggunakan metode literature review jurnal. Data yang digunakan dalam penelitian adalah data sekunder yang bersumber dari hasil pencarian mengenai kesehatan dengan menggunakan machine learning atau *artificial intelligence* atau *data mining*. Artikel dikumpulkan berasal dari penelitian terdahulu berdasarkan index sinta dan scopus yang diterbitkan dari tahun 2015 hingga tahun 2023 yang berkaitan dengan metode analisis Kesehatan dengan menggunakan machine learning atau artificial intelligence atau data mining.

HASIL DAN PEMBAHASAN

Adapun pada pembahasan ini, peneliti bertujuan untuk menganalisis review artikel-artikel jurnal terdahulu diperoleh 30 artikel yang telah dikumpulkan, namun yang sesuai topik penelitian yaitu sebanyak 25 artikel. Adapun hasil pengumpulan artikel ditampilkan pada [Tabel berikut](#):

Tabel 1. Klasifikasi artikel berdasarkan tahun penelitian, metode dan hasil dari penelitian

No	Pengarang	Metode	Hasil penelitian
1	Mei Pilidia Tam`bunan, 2021	<i>Data mining dan clustering algoritma K-Mean</i>	Setelah dilakukan pengelompokan data pemakaian obat menggunakan metode k-mean clustering terbentuk tiga cluster yaitu cluster 1 dengan jumlah 1 items, cluster 2 dengan jumlah 18 items, dan cluster 3 dengan jumlah 20 items.

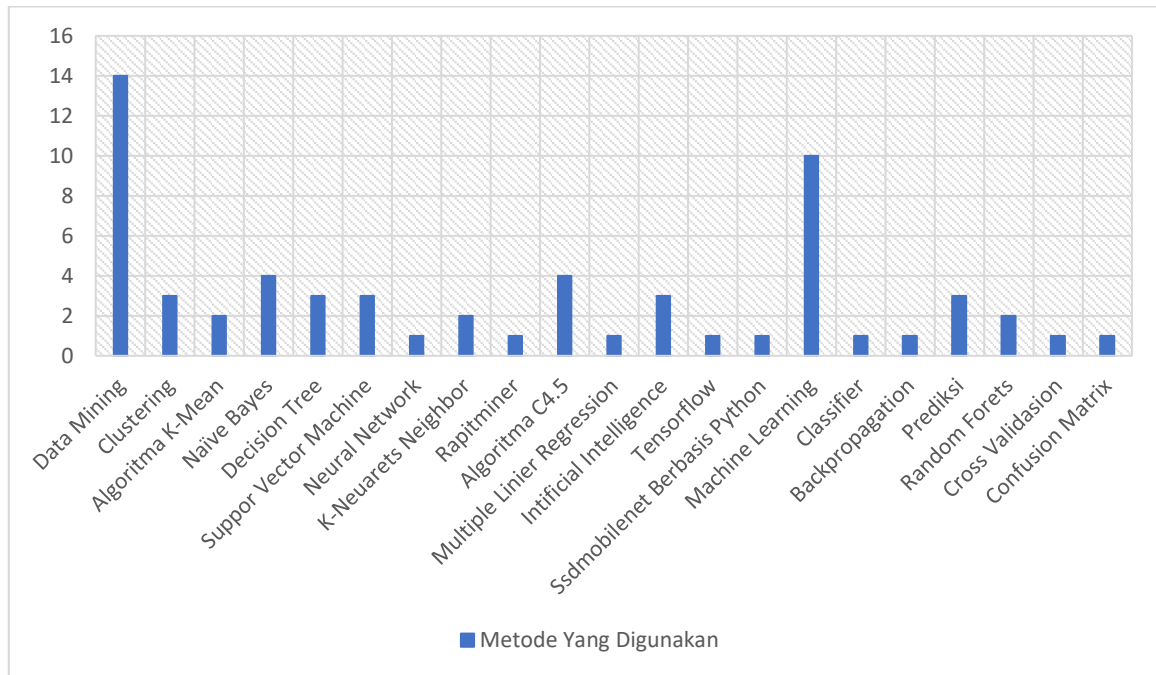
2	Ade Riani, Yessi Susianto & Nur Rahman, 2019	<i>Naive Bayes dan data mining</i>	Diagnosis terhadap penyakit jantung ini memiliki 4 (empat) buah kemungkinan hasil yang dapat dihasilkan yaitu true positive (TP), true negative (TN), false positive (FP), dan false negative (FN). Hasil akhir dari dataset sebanyak 303 memiliki akurasi 86%.
3	Annas Gading Pertiwi & Utomo Pujianto, 2020	<i>Decision tree dan naïve Bayesian dan support vector machine dan neural network dan K-neuarets neighbor dan data mining</i>	Meskipun algoritma C4.5 mempunyai kelemahan yaitu rentan terhadap noise atau outliers, tetap dapat digunakan karena pada data kesehatan, tidak terlalu banyak outliers, selain itu terdapat berbagai cara untuk mengatasi kelemahan tersebut.
4	Kiki Fatmawati & Agus Perdana Windarto, 2018	<i>Rapidminer dan clustering data dan algoritma K-means dan data mining</i>	Dari 34 data jumlah kasus yang terjangkit Demam Berdarah Dengue (DBD) berdasarkan provinsi dapat diketahui, 4 provinsi cluster tingkat tinggi untuk penderita kasus DBD yakni Sumatera Utara, , Jawa Barat, Jawa Tengah dan Jawa Timur, 13 provinsi cluster tingkat sedang dan 17 provinsi lainnya termasuk cluster tingkat rendah.
5	Eka Wulansari Frida Yanthie, 2015	<i>Naïve bayes dan support vector machine dan data mining</i>	Yang terkena penyakit hepatitis adalah pengujian support vector machines didapat nilai accuracy adalah 68.38 % dan nilai AUC adalah 0.726. Sedangkan pengujian dengan menggunakan support vector machines berbasis Naïve Bayes didapatkan nilai accuracy 83.71 % dengan nilai dan nilai AUC adalah 0.812.
6	Rian Ordila, Refni Wahyuni Yuda Irawan & Maulita Tulia Sari, 2020	<i>Algoritma clustering dan data mining</i>	terdapat jenis penyakit Lain-lain yang di maksud adalah jenis penyakit yang jarang di Poli Klinik PT.Inecda seperti Batuk, Demam, Sakit Tenggorokan, Kontrol Pasca OP, Cek Colesterol, Pasien jatuh Dari Motor Dan lain-lain, Dalam jenis penyakit ini dikelompokkan berdasarkan karakterisk umur yaitu Balita 0 pasien, Anak-anak 12 pasien, Dan Dewasa 2124 pasien, dengan jumlah keseluruhan pasien adalah 2136 pasien.
7	Suhastomo Widyanto, 2022	<i>Algoritma C4.5 dan data mining</i>	Hasil pengujian pada saat penulis menggunakan teknik data mining ini tingkat kesalahan dalam memprediksi kesehatan lapangan migas tersebut stabil atau tidaknya dapat dikurangi dengan tingkat kesalahan 1%. sehingga ke depannya lapangan yang tidak stabil akan terus menghasilkan hasil produksi migas yang baik untuk PT Pertamina Hulu Indonesia.
8	Mohammad Yazdi Pusadan Dan Ari Ghifari & Yusuf Anshori, 2023	<i>Algoritma naive bayes dan data mining</i>	Hasil pengujian akurasi untuk memprediksi status proses persalinan pada ibu hamil berdasarkan kondisi ibu, kondisi janin, hasil tes hematologi dan penyakit yang diderita oleh ibu, didapatkan hasil nilai precision

9	Anastasia Carolina Rumahorbo & Kemal Ade Sekarwati, 2020	<i>Algoritma C4.5 dan data mining</i>	<p>sebesar 82.4%, nilai recall sebesar 94%, nilai f-measure sebesar 88.7% dan nilai accuracy sebesar 92%.</p> <p>Hasil implementasi dari data fasilitas kesehatan Indonesia tahun 2017 dengan cross validation terdapat 2 provinsi yang berbeda dengan peringkat IPM, yaitu provinsi Lampung diprediksi dari peringkat Sedang menjadi peringkat Tinggi dan provinsi Sulawesi Utara diprediksi dari peringkat Tinggi menjadi peringkat Sedang.</p>
10	Wisti Dwi Septiani, 2017	<i>Algoritma C4.5 dan naive bayes dan data mining</i>	<p>Dari hasil penelitian yang telah dilakukan pada data pasien penderita penyakit hepatitis maka dapat disimpulkan bahwa metode klasifikasi data mining Algoritma C4.5 menghasilkan akurasi 77,29% dan nilai AUC 0,846 yang termasuk dalam Good Classification. Naive Bayes menghasilkan akurasi 83,71% dan nilai AUC 0,812.</p>
11	Muhammad Zunaidi, Asyahri Hadi Nasuha & Sinta Mega Sinaga, 2020	<i>Multiple linier regression dan data mining</i>	<p>Maka untuk memprediksi pertumbuhan jumlah penderita HIV untuk bulan berikutnya yaitu bulan januari berdasarkan data yang sebelumnya Diketahui 0,140048501, sekarang akan dikalikan dengan 1000 untuk mendapatkan hasil yang akurat. Maka diprediksikan bahwa pertumbuhan jumlah penderita HIV untuk bulan berikutnya adalah sebanyak 140 Penderita.</p>
12	Fitri Andria Astuti, 2021	<i>Machine learning dan artificial intelligence</i>	<p>teknologi artificial intelligence mempunyai potensi besar dalam penguatan sistem kesehatan dan pemulihan ekonomi nasional. Dengan adanya teknologi artificial intelligence bisa dimanfaatkan dalam rangka penguatan sistem kesehatan nasional terutama dalam kasus pengendalian COVID-19.</p>
13	Risnaldi Fatwa Muharram & Ahmad Suryadi, 2022	<i>Tensorflow dan ssdmobilenet berbasis python dan artificial intelligence</i>	<p>salah satu protokol kesehatan yaitu memakai masker. Intensitas cahaya dan jarak dari kamera ke objek mempengaruhi performa serta akurasi model untuk mendeteksi objek. Akurasi tertinggi yang diperoleh mencapai 100% dan akurasi terendah 97% dengan asumsi model mendeteksi masker sesuai data dengan intensitas cahaya lingkungan yang sedang. Hasil Analisa menunjukkan bahwa 3 (tiga) variabel terpenting/paling berpengaruh dalam klasifikasi resiko kejadian BBLR adalah jarak dan urutan kelahiran, pemeriksaan kehamilan, umur ibu. Selanjutnya, 3 (tiga) variabel yang paling</p>
14	Istiqomatul Fajriyah Yulianti & Pardomuan Robinson Sihombing, 2021	<i>Machine learning</i>	

			rendah pengaruhnya terhadap klasifikasi resiko kejadian BBLR adalah sumber air minum, kebiasaan merokok, dan kelahiran kembar.
15	Andry Melyani & Edi Surya Negara, 2022	<i>Naïve bayes classifier dan machine learning</i>	hasil penelitian ini memberikan pendekatan yang efektif dan efisien untuk mendeteksi Kesehatan secara cepat dan mudah. Dengan bantuan aplikasi deteksi Kesehatan, seorang individu dapat mengetahui dan mengatasi pada tahap awal, sehingga dapat mencegah situasi menjadi lebih buruk.
16	Dodi Budiman Margana, 2023	<i>Machine learning</i>	Untuk mengikuti protocol Kesehatan maka diwajibkan menggunakan masker. Keseluruhan percobaan dan pengujian, sistem yang dibuat dapat berhasil mendeteksi jarak, suhu, dan penggunaan masker menggunakan metode machine learning. Sensor jarak yang digunakan mampu mendeteksi jarak dengan error terbesar kurang dari 4 cm dan sensor suhu mampu mendeteksi suhu dengan rata-rata error kurang dari 1,3°C.
17	Ari Purno Wahyu, Muhammad Benny Chaniago, Murnawan, R.A.E Virgana & Cecep Kurnia, 2020	<i>Machine learning</i>	hasil pengujian kesehatan lobster bisa dikenali oleh aplikasi, algoritma computer vision mengenali jenis lobster. Uji coba tersebut diatas terlihat tiga jenis species lobster yaitu jenis lobster bambu, lobster pakistan dan lobster pasir dengan tingkat akurasi masing-masing antara 80% dan 90%.
18	Anjar Wanto & Agus Perdana Windarto, 2017	<i>Backpropagation dan artificial intellingence dan machine learning</i>	Dengan model arsitektur 12-70-1, dapat melakukan prediksi Indeks Harga Konsumen berdasarkan kelompok kesehatan dengan Akurasi 92%. Dengan melihat hasil pengujian, dapat diambil kesimpulan bahwa terjadi kecepatan dan hasil akurasi yang bervariasi pada 8 percobaan di setiap pengujian yang dilakukan.
19	Rifqi Hammad, Julia Kurniasih, Nur Fitrianiingsih Hasan, Christin Nandari Dengen & Kusriani, 2019	<i>machine learning</i>	Hasil Analisa prognosis penyakit demensia terlihat pada tampilan proses kerja dari machine learning prognosis demensia yang dimulai dengan pemilihan jenis pemeriksaan yang akan dilakukan oleh pasien dan kemudian memasukkan data hasil pemeriksaannya. Data hasil pemeriksaan akan diproses berdasarkan rule yang telah didefinisikan dan disimpan di dalam database.
20	Amin Fahri & Yudi Ramdhani, 2023	<i>Decision tree dan machine learning</i>	algoritma Decision Tree Regressor memiliki rmse yang baik yaitu 57, sehingga dapat digunakan oleh para ahli di bidang kesehatan dalam memprediksi COVID-19

21	<u>Agung Supoyo & Putri Tagwa Prasetyanimgrum, 2022</u>	<i>Data mining dan prediksi dan random forests dan k-nearest neighbour</i>	<p>dan bagi seorang peneliti dapat menjadi bahan rujukan metode yang akan diimplementasikan pada saat membuat penelitian mengenai prediksi COVID-19. Hasil Analisa didapatkan performance model terbaik sebesar 4,817. Namun performance model tersebut belum cukup baik untuk digunakan dalam memprediksi lama perawatan pasien Covid-19 di DIY karena jika dibandingkan dengan rata-rata lama perawatan sebesar 15.339 hari maka diperoleh nilai NRMSE yang masih terlalu besar yakni 31,40%.</p>
22	<u>Rakha Gusti Wardhana, Gunawan Wang, Farida & Sibuea, 2023.</u>	<i>Metode machine learning dan prediksi dan decision tree</i>	<p>Dari hasil seluruh proses pemodelan terhadap enam jenis penyakit, proses evaluasi terhadap seluruh pemodelan diatas menyimpulkan bahwa pemodelan dengan teknik decision tree dalam memprediksi klasifikasi tingkat kasus penyakit berdasarkan tingkat pelayanan kesehatan, serta tingkat sosio-demografi penduduk dari masing-masing provinsi.</p>
23	<u>Lusy Octoria Sitanggang & Nurdin Bahtiar, 2018</u>	<i>Metode data mining dan support vector machine (SVM) dan cross validasion dan confusion matrix</i>	<p>Dari hasil penelitian maka didapatkan dataset yang berjumlah 579 ini akan dilakukan proses evaluasi dengan K = 10.menggunakan pengujian pada 10 atribut. data Pasien Hati India memiliki rata-rata nilai akurasi sebesar 87.65%.</p>
24	<u>Baiq Andriska Candra Permana & Muhammad Djamaluddin, 2023</u>	<i>Metode data mining dan algoritma C4.5 dan prediksi</i>	<p>Berdasarkan penelitian yang dilakukan sudah layak untuk dikembangkan dalam dunia kesehatan sebagai model prediksi terhadap penyakit kangker paru karena akurasi yang baik yaitu 87%. Dengan tingkat akurasi tersebut besar kemungkinan prediksi algoritma terhadap suatu gejala dapat menentukan seseorang dinyatakan menderita kangker paru atau tidak.</p>
25	<u>Purwono, Anggit Wirasto2 & Khoirun Nisa, 2021</u>	<i>Algoritma machine learning dan support vector machine dan random forest dan decission tree.</i>	<p>Algoritma SVM memiliki nilai akurasi model sebesar 94.7% namun masih lebih rendah daripada akurasi model algoritma random forest dan decission tree dengan nilai 98,2%. Hasil ini rupanya memiliki nilai akurasi yang lebih tinggi dari penelitian yang sudah dilakukan sebelumnya yaitu oleh Lauv dengan nilai akurasi 90% dengan neural network. Penelitian ini terkesan mendapatkan data akurasi yang sangat bagus yaitu mampu menembus angka 98,2% untuk algoritma decission tree dan random forest.</p>

Literatur artikel riviw memiliki atribut yang berbeda-beda, atribut tersebut akan memberikan pemahaman dalam pemanfaatan bagi pengguna. Artikel terkandung dalam tinjauan literatur sebagai atribut yang memberikan manfaat untuk metode analisis kesehatan dengan menggunakan mechine learning atau artificial inteligenci atau data mining literature review bagi rumah sakit maupun di bagian medis negara. Artikel menurut tahun dan metode risiko yang digunakan dijabarkan pada Gambar dibawah ini.



Gambar 2. Terbitan Artikel Berdasarkan Pada Metode Analisis Kesehatan

Berdasarkan dapat kita lihat jumlah metode yang paling banyak digunakan untuk menganalisis metode analisis Kesehatan itu pada metode data mining. Hal itu dikarenakan karena data mining adalah metode yang paling mudah dipelajari. Metode yang jarang digunakan adalah metode neural network, studi literatur riviw, tensorflow, ssdmobilenet berbasis python, classifier, backpropagation, cross validasion dan confusion matrik. Berdasarkan hal tersebut, beberapa

1. *Data mining* adalah bidang yang sepenuhnya menggunakan apa yang dihasilkan oleh data were house; bersama dengan bidang yang menandai masalah pelaporan dan manajemen data. Sementara, data ware house sendiri bertugas untuk menarik/meng-quary data dari basis data mentah untuk memberikan hasil data yang nantinya digunakan oleh bidang yang menangani manajemen, pelaporan dan data mining (Mei, 2021).
2. *Pengelompokan (clustering)* merupakan bagian dari ilmu data mining yang bersifat tanpa arahan (unsupervised). Clustering adalah proses pembagian data ke dalam kelas atau cluster berdasarkan tingkat kesamaannya. Dalam clustering, data yang memiliki kesamaan dimasukkan ke dalam cluster yang sama, sedangkan data yang tidak memiliki kesamaan dimasukkan dalam cluster yang berbeda (Mei, 2021)
3. *Algoritma K-means* merupakan algoritma yang membutuhkan parameter input sebanyak k dan membagi sekumpulan n objek kedalam k cluster sehingga tingkat kemiripan antar anggota dalam suatu cluster tinggi sedangkan tingkat kemiripan dengan anggota pada cluster lain sangat rendah. Kemiripan anggota terhadap cluster diukur dengan kedekatan objek terhadap nilai mean pada cluster atau dapat disebut sebagai centroid cluster atau pusat massa (Mei, 2021).
4. *Naïve bayes* adalah metode klasifikasi statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Naïve bayes didasarkan pada Teorema Bayes yang memiliki kemampuan klasifikasi. Teorema Bayes adalah memprediksiprobabilitas di masa depan

- berdasarkan pengalaman di masa sebelumnya. Teorema tersebut dikombinasikan dengan naïve yang diasumsikan kondisi antar atribut saling bebas ([Istiqomatu, dkk, 2021](#)).
5. *Decision tree* adalah sebuah model yang direpresentasikan sebagai sebuah pohon, dimana nodes-nya merupakan atribut dari data dan rules-nya dituliskan pada cabang-cabangnya, sehingga terlihat seperti flowchart. Sebuah binary decision tree memisahkan data (parent node) menjadi dua subset (child nodes) dengan menghitung feature terbaik yang ditentukan oleh sebuah kriteria yang dipilih. Dua subset yang terpilih itu menjadi parent nodes yang baru dan dibagi lagi menjadi dua child nodes. ([Annas, dkk, 2020](#)).
 6. *Support Vector Machine (SVM)* adalah metode untuk klasifikasi untuk data linear dan nonlinear. Metode SVM bekerja dengan pemetaan nonlinear untuk mengubah training menjadi dimensi yang lebih tinggi, kemudian dicari linear optimal separating hyperplane. Selanjutnya, data dari dua kelas tersebut dapat dipisahkan oleh sebuah hyperplane, yang dicari menggunakan support vectors dan margins [26]. Support vector machines diprediksikan adalah algoritma paling baik dalam akurasi [32]. Support vector machines dinyatakan sebagai algoritma yang kuat karena strukturnya yang sederhana dan memerlukan jumlah fitur yang lebih sedikit. ([Annas, dkk, 2020](#)).
 7. *Neural Network* adalah bagian utama dari sistem saraf yang mengatur dan mengontrol fungsi tubuh dan aktifitas. Jaringan saraf memiliki dua sistem utama, yaitu Sistem Saraf Pusat yang terdiri dari otak dan sumsum tulang belakang, lalu yang kedua adalah Sistem Saraf Tepi Artificial Neural Network ([Risnaldy, dkk, 2022](#)).
 8. *Metode klasifikasi K-Nearest Neighbour (K-NN)* merupakan salah satu metode dari algoritma yang paling sederhana yang menemukan data yang tidak teridentifikasi menggunakan data points yang diketahui (nearest neighbor) dan mengklasifikasikan data dengan sistem voting [31]. Jika ditambahkan sebuah data yang tidak diketahui labelnya, K-NN mengklasifikasikan data tersebut dengan menghitung jarak terdekat majority dengan data yang sudah ada. Untuk menghitung jarak terdekat tersebut, digunakan banyak metode, namun yang paling banyak digunakan adalah Euclidean distance. ([Annas, dkk, 2020](#)).
 9. RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. Hasil pengelompokan data daerah terjangkit Demam Berdarah Dengue (DBD) berdasarkan provinsi dengan RapidMiner. (Kiki, dkk, 2018) ([Sulistianingtyas dkk, 2021](#)).
 10. *Algoritma C4.5* merupakan pengembangan dari algoritma ID3 (Iterative Dichotomiser 3) yang dikembangkan oleh J. Ross Quinlan. Idanya adalah membuat pohon dengan percabangan awal merupakan atribut yang paling signifikan, dan seterusnya dibagi menjadi percabangan-percabangan sampai aturan terpenuhi. Pohon keputusan itu sendiri diartikan sebagai suatu cara untuk membagi sekumpulan data menjadi himpunan-himpunan yang lebih kecil dengan menerapkan serangkaian rule atau aturan keputusan. Algoritma IDS ini pada dasarnya hanya mengulang langkah pamtisian sampai diperoleh keadaan dimana semua sampel pada sebuah simpul tergolong ke dalam kelas yang sama. ([Ifan, 2019](#)).
 11. *Metode Multiple Linier Regression* adalah nalisis Regresi adalah salah satu teknik analisis yang dipergunakan untuk mencari hubungan dari dua atau lebih variabel yang bersifat sebab akibat. Metode Multiple Linier Regression digunakan untuk membantu memprediksi pertumbuhan jumlah penderita HIV, agar persediaan penanganan yang diperlukan bisa disesuaikan dengan jumlah penderita HIV yang ada. ([Muhammad, dkk, 2020](#))
 12. *Artificial Intelligence (AI)* yang dalam Bahasa Indonesia artinya Kecerdasan Buatan adalah kecerdasan yang diperlihatkan oleh mesin, tidak seperti kecerdasan alami yang diperlihatkan oleh manusia dan hewan yang melibatkan kesadaran dan emosi. Kecerdasan dibuat lalu dimasukkan ke dalam mesin (komputer) dengan harapan supaya mesin (komputer) dapat melakukan pekerjaan sesuai dengan perintah yang sudah diberikan. ([Risnaldy, dkk, 2022](#))
 13. *Tensorflow* menggabungkan banyak model, algoritma machine learning dan algoritma deep learning (jaringan syaraf). Tensorflow menggunakan Python untuk menyediakan API front-end untuk membangun aplikasi dengan framework, sekaligus menjalankan aplikasi tersebut dengan performa tinggi. ([Risnaldy, dkk, 2022](#))
 14. *Ssdmobilenet berbasis phyton* model deteksi objek yang menghitung kotak pembatas keluaran dan kelas objek dari gambar masukan. Python adalah Intrepeter yang memiliki kemampuan exeption handling serta agar sintaksnya yang mudah dipahami. Python menggunakan indentasi untuk membedakan blok kode, tidak seperti bahasa pemograman lain yang menggunakan simbol sebagai pembeda blok kode. ([Risnaldy, dkk, 2022](#))

15. *Machine learning* merupakan ilmu dan seni tentang pemrograman komputer yang bisa belajar dari data. Machine learning secara umum dibagi menjadi 4, yaitu supervised learning, unsupervised learning, semi unsupervised learning, dan reinforcement learning. Sebuah model machine learning tidak mampu untuk langsung mengolah data yang ditemukan dari berbagai sumber. Ada istilah Garbage In - Garbage Out yang berarti hasil dari machine learning akan buruk jika input yang dimasukkan juga buruk. (Fitri, 2021)
16. *Naive Bayes classifier* merupakan salah satu metode pembelajaran mesin yang memanfaatkan perhitungan probabilitas. Dalam mengimplementasikan aplikasi m-health prediksi kesehatan, sebuah algoritma telah dikembangkan dengan mengintegrasikan Naïve Bayes Classifier untuk mengklasifikasikan pengguna dengan tingkat skor kesehatan 0,1,2, dan 3. Indikator yang menjadi tolak ukur adalah suhu tubuh, detak jantung, respirasi, tekanan darah sistolik dan diastolik, dan saturasi oksigen. (Andry, dkk, 2022)
17. Backpropagation pendekatan jaringan saraf tiruan dapat meniru perilaku yang kompleks dan non-linear melalui neuron, dan telah banyak digunakan dalam prediksi. Model yang paling banyak digunakan pada kecerdasan buatan adalah model backpropagation. (Anjar, dkk, 2017)
18. *Prediksi* (peramalan) adalah usaha menduga atau memperkirakan sesuatu yang akan terjadi di waktu mendatang dengan memanfaatkan berbagai informasi yang relevan pada waktu-waktu sebelumnya (historis) melalui suatu metode ilmiah. (Anjar, dkk, 2017)
19. *Random Forest* merupakan salah satu metode dalam decision tree atau pohon keputusan. Metode ini digunakan untuk membangun pohon keputusan yang terdiri dari root node, internal node, dan leaf node dengan mengambil atribut dan data secara acak sesuai ketentuan yang diberlakukan. Root node digunakan untuk mengumpulkan data, sebuah inner node yang berada pada root node berisi pertanyaan tentang data, dan sebuah leaf node digunakan untuk memecahkan masalah serta membuat keputusan. (Istiqomatu, dkk, 2021)
20. *Cross-validation (CV)* adalah metode statistik yang dapat digunakan untuk mengevaluasi kinerja model atau algoritma dimana data dipisahkan menjadi dua subset yaitu data proses pembelajaran dan data validasi / evaluasi. K-Fold Cross Validation melakukan pembagian data dengan membagi dataset secara terstruktur menjadi k himpunan bagian (subset). (Lusy, dkk, 2018)
21. *Confusion matrix* memberikan keputusan yang diperoleh dalam training dan testing, confusion matrix memberikan penilaian performance klasifikasi berdasarkan objek dengan benar atau salah. (Eka, 2015)

Dari literature review mengenai metode analisis kesehatan menggunakan machine learning, artificial intelligence (AI), atau data mining, dapat disimpulkan bahwa teknologi ini memberikan kontribusi besar dalam transformasi cara kita memahami dan mengelola kesehatan. Berikut adalah beberapa poin penting yang dapat diambil:

Pertama, penggunaan machine learning, AI, dan data mining telah membuka pintu untuk analisis yang lebih mendalam terhadap data kesehatan. Algoritma-algoritma yang kompleks mampu mengolah dan menganalisis data medis dalam skala yang besar dengan kecepatan yang lebih tinggi daripada manusia. Hal ini memungkinkan identifikasi pola-pola yang mungkin tidak terdeteksi oleh metode tradisional. Kedua, aplikasi utama dari teknologi ini adalah dalam prediksi dan diagnosis penyakit. Berbagai model prediktif telah dikembangkan untuk memperkirakan risiko penyakit tertentu berdasarkan faktor-faktor tertentu seperti riwayat kesehatan, pola hidup, dan genetik. Selain itu, sistem diagnosis yang menggunakan machine learning mampu memberikan rekomendasi diagnosis yang lebih akurat dan cepat berdasarkan gejala-gejala yang dipresentasikan oleh pasien.

Ketiga, analisis kesehatan menggunakan teknologi ini juga membantu dalam mengidentifikasi pola-pola epidemiologi dan tren kesehatan populasi. Dengan menganalisis data yang berasal dari berbagai sumber seperti catatan kesehatan elektronik, sensor wearable, dan data genomik, kita dapat memahami faktor-faktor yang mempengaruhi kesehatan suatu populasi dan merencanakan intervensi yang lebih efektif. Keempat, pengembangan model-model prediktif dan klasifikasi memerlukan evaluasi yang cermat terhadap keandalan dan akurasi model tersebut. Metode seperti cross-validation dan confusion matrix digunakan untuk mengukur performa model dalam mengklasifikasikan penyakit atau memprediksi hasil kesehatan. Evaluasi yang baik sangat penting untuk memastikan bahwa model yang dikembangkan dapat dipercaya dan digunakan dalam praktik klinis.

Kelima, peran software dan platform seperti RapidMiner dan TensorFlow sangat penting dalam pengembangan dan implementasi model analisis kesehatan. Mereka menyediakan alat-alat yang diperlukan untuk pengolahan data, pembangunan model, serta evaluasi kinerja model. Penggunaan platform-platform ini mempermudah para peneliti dan praktisi kesehatan dalam memanfaatkan teknologi machine learning dan AI untuk meningkatkan pemahaman dan pengelolaan kesehatan. Secara keseluruhan, teknologi machine learning, artificial intelligence, dan data mining memiliki potensi besar dalam meningkatkan pemahaman, diagnosis, dan pengelolaan kesehatan. Namun, tantangan yang dihadapi termasuk interpretasi yang tepat dari hasil analisis, privasi dan keamanan data, serta integrasi dengan praktik klinis yang ada. Dengan terus melakukan penelitian dan pengembangan, kita dapat memanfaatkan potensi teknologi ini untuk meningkatkan kesehatan dan kesejahteraan masyarakat secara keseluruhan.

KESIMPULAN

kondisi tubuh yang sakit, akan membuat seseorang menjadi tidak produktif dan bisa mendapatkan risiko kematian. Seiring dengan perkembangan teknologi, turut serta memberikan dampak dalam kemajuan bidang kesehatan. Teknologi membantu tenaga medis untuk melakukan tindakan dan perawatan kepada pasien. Penerapan metode analisis kesehatan dapat membantu rumah sakit ataupun lainnya menyangkut tentang Kesehatan dapat berguna untuk meminimalisir permasalahan yang sulit ditemukan. Hasil Analisa jurnal artikel riviw maka ditarik kesimpulan bahwa terdapat beberapa metode analisis Kesehatan yang umum digunakan seperti data mining, clustering, algoritma k-mean, naïve baye, decision tree, suppor vector machine, neural network, k-neuarets neighbor, rapitminer, algoritma C4.5, multiple linier regression, intifical intelligence, tensorflow, ssdmobilenet berbasis python, machine learning, classifier, backpropogion, prediksi, random forets, cross validasion dan confusion matrix. Beberapa metode tersebut umumnya namun didalamnya juga memakai metode atau pendekatan lainnya agar hasil yang didapatkan lebih optimal.

DAFTAR PUSTAKA

- [Annas Gading Pertiwi & Utomo Pujianto, \(2020\). Metode-metode data mining untuk penyelesaian masalah kehamilan dan persalinan. Sains Aplikasi, Komputasi dan Teknologi Informasi \(Vol 2, No 1, 11-17\)](#)
- [Agung Supoyo & Putri Taqwa Prasetyanimgrum, 2022 Analisis Data Mining Untuk Memprediksi Lama Perawatan Pasien Covid-19 Di DIY. Bianqlala Informatika \(Vol. 10 No. 1\).](#)
- [Anastasia Carolina Rumahorbo & Kemal Ade Sekarwati, 2020 Penerapan Data Mining dengan Menggunakan Algoritma C4.5 pada Klasifikasi Fasilitas Kesehatan Provinsi di Indonesia. Jurnal Ilmiah KOMPUTASI, \(Vol. 19 No : 1\)](#)
- [Andry Melyani & Edi Surya Negara, 2022. Aplikasi Prediksi Kesehatan Menggunakan Machine Learning. Jurnal Jupiter,\(Vol. 14 No. 2, 208 – 215\)](#)
- [Kiki Fatmawati & Agus Perdana Windarto, 2018 DATA MINING: PENERAPAN RAPIDMINER DENGAN K-MEANS CLUSTER PADA DAERAH TERJANGKIT DEMAM BERDARAH DENGUE \(DBD\) BERDASARKAN PROVINSI. CESS \(Journal of Computer Engineering System and Science\) \(Vol. 3 No. 2\)](#)
- [Anjar Wanto & Agus Perdana Windarto, 2017. Analisis Prediksi Indeks Harga Konsumen Berdasarkan Kelompok Kesehatan Dengan Menggunakan Metode Backpropagation. Jurnal & Penelitian Teknik Informatika \(Vol. 2 No. 2\).](#)
- [Amin Fahri & Yudi Ramdhani, 2023. Visualisasi Data dan Penerapan Machine Learning Menggunakan Decision Tree Untuk Keputusan Layanan Kesehatan COVID-19. Jurnal TEKNO KOMPAK, \(Vol. 17, No. 2, 50-60\)](#)
- [Ari Purno Wahyu, Muhammad Benny Chaniago, Murnawan, R.A.E Virgana & Cecep Kurnia, 2020. SISTEM PEMANTAUAN KESEHATAN LOBSTER \(LHMS\) DENGAN MACHINE LEARNING. JURSISTEKNI \(Jurnal Sistem Informasi dan Teknologi Informasi\) \(Vol 2, No.1, 1- 9\).](#)
- [Ade Riani, Yessi Susianto & Nur Rahman, 2019. Implementasi Data Mining Untuk Memprediksi Penyakit Jantung Menggunakan Metode Naive Bayes. Journal of Innovation Information Technology and Application \(JINITA\) \(Vol.1, No.01\).](#)

- Adinda Inez Sang, Edi Sutoyo 2021. ANALISIS DATA MINING UNTUK KLASIFIKASI DATA KUALITAS UDARA DKI JAKARTA MENGGUNAKAN ALGORITMA DECISION TREE DAN SUPPORT VECTOR MACHINE. e-Proceeding of Engineering (Vol.8, No.5).
- Baiq Andriskha Candra Permana & Muhammad Djamaluddin, 2023. Penerapan Python Dalam Data Mining Untuk Prediksi Kanker Paru. Infotek : Jurnal Informatika dan Teknologi (Vol. 6 No. 2,470-477)
- Dodi Budiman Margana, 2023. Pendeteksian jarak, suhu, dan penggunaan masker otomatis pada pelayanan publik dengan metode machine learning. JITEL (Jurnal Ilmiah Telekomunikasi, Elektronika, dan Listrik Tenaga) (Vol. 3, No. 1, 65-70).
- Eka Wulansari Frida Yanthie. (2015). ANALISA DATA MINING UNTUK PREDIKSI PENYAKIT HEPATITIS DENGAN MENGGUNAKAN METODE NAIVE BAYES DAN SUPPORT VECTOR MACHINE. JURNAL KHATULISTIWA INFORMATIKA, (Vol. 3, No. 1).
- Fitri Andria Astuti. (2021). Pemanfaatan Teknologi Artificial Intelligence untuk Penguatan Kesehatan dan Pemulihan Ekonomi Nasional. Jurnal Sistem Cerdas (2021) (Vol 04 - No 01, 25 – 34).
- Istiqomatul Fajriyah Yulianti & Pardomuan Robinson Sihombing,2021. Penerapan Metode Machine Learning dalam Klasifikasi Risiko Kejadian Berat Badan Lahir Rendah di Indonesia. Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer (Vol. 20, No. 2, pp. 417~426)
- Imca Pero Hasfera1. (2023). ARTIFICIAL INTELEGENCY UNTUK LAYANAN KESEHATAN DI PERPUSTAKAAN UNBRAH. Jurnal Ilmu Perpustakaan dan Informasi Islam
- Jacob Delwin Esther, Sajaya. (2018). Faktor Faktor Yang Mempengaruhi Kualitas Hidup Masyarakat Karubaga District Sub District Tolikara Propinsi Papua. Jurnal Nasional Ilmu Kesehatan (JNIK), 1, 1-16
- Kementerian Kesehatan Republik Indonesia. "Cetak Biru Strategi Transformasi Digital Kesehatan 2024", 2021
- Kiki Fatmawati & Agus Perdana Windarto, (2018) DATA MINING: PENERAPAN RAPIDMINER DENGAN K-MEANS CLUSTER PADA DAERAH TERJANGKIT DEMAM BERDARAH DENGUE (DBD) BERDASARKAN PROVINSI. CESS (Journal of Computer Engineering System and Science) (Vol. 3 No. 2)
- Lusy Octoria Sitanggang & Nurdin Bahtiar, 2018. Aplikasi Data Mining Untuk Mendeteksi Penyakit Hepatitis Menggunakan Metode Support Vector Machine (SVM) Polynomial (Studi Kasus : Data Pasien Hati India). Jurnal Masyarakat Informatika, (Vol. 10, No. 1),
- Muhammad Zunaidi, Asyahri Hadi Nasyuha & Sinta Mega Sinaga, 2020. Penerapan Data Mining Untuk Memprediksi Pertumbuhan Jumlah Penderita Human Immunodeficiency Virus (HIV) Menggunakan Metode Multiple Linier Regression (Studi Kasus Dinas Kesehatan Provinsi Sumatera Utara). Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD (Vol.3, No.1, pp.137-147)
- Mohammad Yazdi Pusadan Dan Ari Ghifari & Yusuf Anshori, 2023. Implementasi Data Mining untuk Prediksi Status Proses Persalinan pada Ibu Hamil Menggunakan Algoritma Naive Bayes. Technomedia Journal (TMJ) (Vol. 8 No. 1 Juni 2023).
- Mei Pilidia Tambunan, 2021. Penerapan Data Mining Dalam Analisa Data Pemakaian Obat Dengan Menerapkan Algoritma K-Means. Jurnal Informasi dan Teknologi Ilmiah (INTI)ISSN 2301-9425 (Media Cetak) (Vol 8, No 3,109-113).
- Purwono, Anggit Wirasto2 & Khoirun Nisa, 2021. Komparasi Algoritma Machine Learning Untuk Klasifikasi Kelompok Obat Comparison of Machine Learning Algorithms for Classification of Drug Groups. JURNAL SISFOTENIKA (Vol. 11 No. 2).
- Rian Ordila Refni Wahyuni Yuda Irawan & Maulita Tulia Sari, 2020. Penerapan Data Mining Untuk Pengelompokan Data Rekam Medis Pasien Berdasarkan Jenis Penyakit Dengan Algoritma Clustering (Studi Kasus : Poli Klinik Pt.Inecda). JURNAL ILMU KOMPUTER (COMPUTER SCIENCE JOURNAL). 9 (2):148 – 153.

- Risnaldy Fatwa Muharram & Ahmad Suryadi, 2022. Implementasi Artificial Intelligence Untuk Deteksi Masker Secara Realtime Dengan Tensorflow Dan Ssdmobilenet Berbasis Python. *Jurnal Widya*. 3(2),281-290.
- Rakha Gusti Wardhana, Gunawan Wang, Farida & Sibuea, 2023. PENERAPAN MACHINE LEARNING DALAM PREDIKSI TINGKAT KASUS PENYAKIT DI INDONESIA. *Journal of Information System Management (JOISM)*. 5(1)
- Suhastomo Widyanto, 2022. Implementasi Data Mining untuk Menentukan Kesehatan Lapangan Migas dengan Metode Algoritma C4.5 (Studi Kasus PT Pertamina Hulu Indonesia). *Jurnal Ilmu Komputer dan Teknologi Informasi (Vol. 7 No: 1, April 2022)*.
- Tri Wahyudi,2023. Studi Kasus Pengembangan dan Penggunaan Artificial Intelligence (AI) Sebagai Penunjang Kegiatan Masyarakat Indonesia. *Indonesian Journal on Software Engineering (IJSE)* 9(1)
- Wardhana Gusti Rakha, Wang Gunawan, Sibuea Fariba. (2023). Penerapan Machine Learning Dalam Prediksi Tingkat Kasus Penyakit Di Indonesia. *Journal of Information System Management (JOISM)*. 5(1),
- Wisti Dwi Septiani, 2017. KOMPARASI METODE KLASIFIKASI DATA MINING ALGORITMA C4.5 DAN NAIVE BAYES UNTUK PREDIKSI PENYAKIT HEPATITIS. *Jurnal Pilar Nusa Mandiri (Vol13 No.1)*