

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

2.1.1 Distribusi Tenaga Medis

Distribusi tenaga medis yang merata sangat penting untuk menjamin akses layanan kesehatan yang adil. Ketimpangan distribusi tenaga kesehatan dapat memperburuk disparitas kesehatan antarwilayah. Di Indonesia, sebagai negara kepulauan, kondisi geografis yang beragam secara signifikan memengaruhi penyebaran tenaga medis. Ini pada gilirannya berdampak besar pada ketersediaan dan kualitas layanan kesehatan di berbagai daerah, seperti yang diungkapkan oleh Handayani dan Sitokdana (2024) dalam penelitian mereka.

Ketidakmerataan ini menyebabkan beberapa wilayah, khususnya yang terpencil atau kurang berkembang, seringkali mengalami kekurangan jumlah dan jenis tenaga medis yang diperlukan. Sebaliknya, pusat-pusat perkotaan cenderung memiliki konsentrasi tenaga medis yang lebih tinggi. Situasi ini menciptakan perbedaan yang jelas dalam layanan kesehatan di seluruh provinsi (Handayani & Sitokdana, 2024).

Oleh karena itu, memahami pola distribusi tenaga medis menjadi sangat krusial untuk mengidentifikasi area yang membutuhkan intervensi kebijakan dan untuk merencanakan alokasi sumber daya kesehatan secara lebih efektif. Dengan menganalisis data distribusi tenaga kesehatan, pemerintah dan pembuat kebijakan

dapat mengambil keputusan yang lebih tepat untuk memastikan setiap warga negara memiliki akses yang setara terhadap pelayanan kesehatan yang berkualitas.

2.1.2 Faktor Penentu Prioritas Penambahan Tenaga Medis

Identifikasi faktor-faktor penentu prioritas penambahan tenaga medis memerlukan analisis multidimensional yang komprehensif. Pertama, urgensi penyakit menjadi indikator utama, yang dapat dinilai dari prevalensi penyakit menular dan tidak menular, angka morbiditas dan mortalitas yang tinggi, serta beban penyakit endemik di suatu wilayah (Ramadhan & Amalia, 2023). Kedua, sebaran fasilitas kesehatan harus dievaluasi, termasuk jumlah dan jenis fasilitas primer (puskesmas, klinik) serta fasilitas sekunder (rumah sakit), serta aksesibilitas geografisnya bagi penduduk. Faktor demografi seperti kepadatan penduduk, rasio tenaga medis per kapita, tingkat pertumbuhan penduduk, dan proyeksi kebutuhan tenaga medis di masa mendatang juga sangat memengaruhi penilaian prioritas. Kombinasi faktor-faktor ini akan memberikan gambaran holistik mengenai kebutuhan riil di setiap wilayah, memungkinkan pengambilan keputusan yang berbasis bukti (Febriyanto et al., 2022).

2.1.3 Data Mining

Data mining adalah sebuah disiplin ilmu yang mengeksplorasi kumpulan data berukuran besar dengan tujuan menemukan pola-pola tersembunyi, hubungan yang tidak terduga, dan informasi yang bermakna yang dapat digunakan untuk pengambilan keputusan (Han et al., 2012). Dalam konteks penelitian ini, teknik data mining akan menjadi tulang punggung dalam mengolah dan menganalisis dataset kompleks yang

mencakup informasi kesehatan masyarakat, data demografi, dan sebaran fasilitas kesehatan di Kota Bekasi. Proses ini melibatkan serangkaian tahapan mulai dari data cleaning, transformasi, hingga aplikasi algoritma canggih untuk mengekstraksi pengetahuan yang relevan. Melalui data mining, kita dapat mengungkap wawasan penting yang mungkin tidak terlihat melalui metode analisis tradisional, sehingga mendukung klasifikasi wilayah prioritas yang lebih akurat dan terinformasi (Dhewayani et al., 2022).

2.1.4 Klasifikasi

Klasifikasi merupakan salah satu teknik supervised learning fundamental dalam data mining, di mana tujuannya adalah untuk membangun model yang dapat memprediksi kelas atau kategori dari sebuah instance data berdasarkan atribut-atribut yang dimilikinya (Tan et al., 2005). Proses ini melibatkan fase pelatihan, di mana algoritma belajar dari dataset yang telah diberi label (misalnya, wilayah yang sudah dikategorikan berdasarkan prioritas), dan fase pengujian, di mana model yang terlatih digunakan untuk memprediksi kelas pada data baru yang belum diberi label. Dalam penelitian ini, klasifikasi akan diterapkan untuk mengelompokkan wilayah-wilayah di Kota Bekasi ke dalam kategori prioritas penambahan tenaga medis, seperti "tinggi," "sedang," atau "rendah," dengan memanfaatkan berbagai fitur data seperti tingkat prevalensi penyakit, jumlah fasilitas kesehatan, dan rasio tenaga medis. Hasil klasifikasi ini akan menjadi dasar rekomendasi strategis.

2.1.5 Support Vector Machine (SVM) dalam Klasifikasi Kesehatan

Support Vector Machine (SVM) adalah algoritma supervised learning yang sangat kuat dan sering digunakan untuk tugas klasifikasi maupun regresi. Algoritma ini telah menjadi metode machine learning yang populer, terutama dalam skenario ketika data memiliki dimensi tinggi atau batas keputusan antar kelas kompleks dan tidak linear (Fitriani et al., 2023). Inti dari SVM adalah menemukan hyperplane optimal di ruang fitur yang secara maksimal memisahkan kelas-kelas data. Hyperplane ini merupakan batas keputusan yang memaksimalkan margin, yaitu jarak terpendek antara hyperplane dengan data point terdekat dari setiap kelas. Data point yang paling dekat tersebut disebut support vector; mereka adalah elemen kritis yang menentukan posisi serta orientasi hyperplane. Dengan memaksimalkan margin, SVM meningkatkan kemampuan generalisasi model, sehingga kinerjanya tetap andal bahkan pada data baru. Inilah alasan mengapa SVM sering menghasilkan model yang tangguh dan efektif dalam skenario dunia nyata.

Salah satu keunggulan penting dari SVM adalah kemampuannya dalam menangani masalah klasifikasi non-linear. Jika data tidak dapat dipisahkan secara linear di ruang aslinya, SVM menggunakan teknik kernel trick, yaitu metode pemetaan implisit ke ruang berdimensi lebih tinggi agar data menjadi lebih mudah dipisahkan secara linear. Berbagai fungsi kernel dapat digunakan, seperti linear, polynomial, Radial Basis Function (RBF) atau Gaussian, dan sigmoid. Pemilihan kernel yang tepat sangat bergantung pada karakteristik data serta kompleksitas hubungan antar kelas.

Selain itu, pengaturan parameter model seperti C (parameter penalti untuk kesalahan klasifikasi) dan gamma (pada kernel RBF) sangat krusial. Proses tuning parameter biasanya dilakukan dengan teknik validasi silang untuk memperoleh konfigurasi yang optimal.

Dalam bidang kesehatan, SVM telah terbukti mampu menangani analisis data medis yang kompleks. Misalnya, algoritma ini digunakan untuk diagnosis penyakit berdasarkan gejala maupun hasil laboratorium, prediksi risiko kesehatan dari riwayat medis dan gaya hidup, serta klasifikasi jenis sel dalam citra medis atau identifikasi pola pada data genomik dan proteomik (Nurqotimah et al., 2024). Fleksibilitas SVM dalam mengolah berbagai jenis data—baik numerik maupun kategorikal (setelah melalui proses encoding)—serta kemampuannya mencegah overfitting melalui prinsip maximal margin hyperplane, menjadikannya sangat relevan untuk penelitian ini. Oleh karena itu, SVM dipilih sebagai algoritma utama dalam mengklasifikasikan wilayah prioritas berdasarkan urgensi penyakit dan sebaran fasilitas kesehatan di Kota Bekasi.

Untuk mengukur efektivitas model SVM dalam melakukan klasifikasi, digunakan beberapa metrik evaluasi. Akurasi (Accuracy) mengukur proporsi keseluruhan prediksi yang tepat dengan rumus:

$$\text{Accuracy} = \frac{\text{Jumlah Total Prediksi}}{\text{Jumlah Prediksi Benar}} \quad (2.1)$$

Namun, akurasi bisa menyesatkan jika dataset tidak seimbang. Presisi (Precision) mengukur relevansi dari hasil positif yang diprediksi model, dirumuskan sebagai:

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}} \quad (2.2)$$

Presisi tinggi berarti alokasi sumber daya lebih efisien. Recall (Sensitivity / True Positive Rate) mengukur kelengkapan model dalam menemukan seluruh kasus positif, dengan rumus:

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatif (FN)}} \quad (2.3)$$

Recall tinggi memastikan tidak ada wilayah kritis yang terlewat. Selanjutnya, F1-Score merupakan rata-rata harmonik dari Precision dan Recall, yang memberikan metrik tunggal seimbang, terutama pada dataset tidak seimbang. Rumusnya adalah:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2.4)$$

Selain itu, Cohen's Kappa (κ) digunakan untuk mengukur tingkat kesepakatan antara prediksi model dengan label aktual, sekaligus mengoreksi kemungkinan kesepakatan yang terjadi secara acak. Rumusnya adalah:

$$K = \frac{P_o - P_e}{1 - P_e} \quad (2.5)$$

Terakhir, Matriks Konfusi (Confusion Matrix) digunakan sebagai alat visual yang merangkum jumlah True Positives (TP), True Negatives (TN), False Positives (FP),

dan False Negatives (FN) untuk setiap kelas. Matriks ini membantu mengidentifikasi jenis kesalahan yang paling sering dilakukan model dan memberikan gambaran menyeluruh mengenai performa klasifikasi.

2.1.6 CRISP-DM (Cross Industry Standard Process for Data Mining)

CRISP-DM (Cross-Industry Standard Process for Data Mining) merupakan metodologi standar yang menyediakan kerangka kerja sistematis dan berulang dalam proyek-proyek data mining dan analisis data. Metode ini membimbing proses ekstraksi wawasan dari data, mulai dari pemahaman masalah bisnis hingga implementasi solusi berbasis data. Dalam studi yang dilakukan oleh Sastya dan Nugraha (2023), pendekatan CRISP-DM diterapkan secara sistematis dalam menganalisis data penjualan MeatSolution untuk memahami perilaku pelanggan. Proses dimulai dari tahap business understanding, yaitu penetapan tujuan bisnis berupa analisis penjualan historis guna menyusun strategi pemasaran yang lebih baik. Tahap berikutnya, data understanding, dilakukan dengan eksplorasi data menggunakan Microsoft Excel dan Power BI. Selanjutnya, pada tahap data preparation, data dibersihkan dan disiapkan untuk analisis. Pada tahap modeling, digunakan teknik-teknik tertentu untuk mengidentifikasi pola perilaku konsumen. Hasil dari pemodelan kemudian dievaluasi (evaluation) untuk memastikan kesesuaiannya dengan kebutuhan bisnis. Terakhir, pada tahap deployment, wawasan seperti produk yang paling laris, bulan penjualan tertinggi, metode pembayaran yang paling sering digunakan, dan karakteristik pembeli disampaikan untuk mendukung pengambilan keputusan strategis perusahaan. Penerapan CRISP-

DM dalam konteks ini terbukti efektif dalam mengubah data menjadi informasi yang bernilai secara strategis (Sastya & Nugraha, 2023).

2.1.7 Pemodelan UML

Unified Modeling Language (UML) adalah bahasa standar dan notasi grafis yang digunakan secara luas untuk memodelkan, memvisualisasikan, merancang, dan mendokumentasikan sistem perangkat lunak. Dalam pengembangan sistem klasifikasi wilayah prioritas penambahan tenaga medis, UML merupakan alat fundamental untuk memetakan arsitektur sistem, memodelkan fungsionalitas, serta menggambarkan interaksi antar komponen dan pengguna secara komprehensif. Penggunaan UML akan meningkatkan kejelasan komunikasi antara tim pengembang dan pemangku kepentingan, sekaligus memastikan bahwa sistem yang dibangun selaras dengan persyaratan fungsional dan non-fungsional yang ditetapkan (Putri & Taufik, 2024).

2.1.7.1 Use Case Diagram

Use Case Diagram adalah salah satu jenis diagram perilaku dalam UML yang berfokus pada pemodelan fungsionalitas sistem dari sudut pandang pengguna (aktor). Diagram ini secara visual mengidentifikasi entitas eksternal yang berinteraksi dengan sistem (aktor) dan fungsi-fungsi utama atau layanan yang dapat mereka lakukan (use case) (Putri & Taufik, 2024). Tujuan utama dari Use Case Diagram adalah untuk menggambarkan apa yang dilakukan sistem tanpa merinci bagaimana hal itu dilakukan, sehingga mempermudah pemahaman tentang persyaratan fungsional dari berbagai perspektif pengguna. Dalam konteks sistem klasifikasi wilayah prioritas penambahan

tenaga medis, Use Case Diagram akan menggambarkan secara jelas bagaimana berbagai pengguna—seperti Administrator Sistem, Petugas Dinas Kesehatan, atau Analis Data—dapat berinteraksi dengan sistem. Contoh use case yang akan dimodelkan antara lain Mengelola Data Wilayah, Menjalankan Proses Klasifikasi SVM, dan Melihat Laporan Hasil Klasifikasi Prioritas Tenaga Medis. Pemodelan ini sangat membantu dalam mendefinisikan cakupan fungsional sistem secara keseluruhan dan memahami ekspektasi fungsional dari berbagai perspektif pengguna.

2.1.7.2 Diagram Activity

Activity Diagram adalah diagram UML lain yang digunakan untuk memodelkan alur kontrol atau alur kerja dari suatu proses bisnis atau operasional dalam sebuah sistem. Diagram ini menunjukkan urutan langkah-langkah yang harus dilakukan untuk menyelesaikan suatu aktivitas, termasuk titik awal, aktivitas individual, keputusan, percabangan, penggabungan, dan titik akhir (Nurqotimah et al., 2024). Untuk sistem klasifikasi wilayah prioritas penambahan tenaga medis, Activity Diagram akan memvisualisasikan alur kerja end-to-end. Ini akan merinci proses mulai dari input data kesehatan dan demografi, tahap preprocessing data yang melibatkan pembersihan dan transformasi, kemudian eksekusi algoritma SVM untuk klasifikasi, hingga penyimpanan hasil klasifikasi dan penyajiannya kepada pengguna melalui antarmuka sistem. Diagram ini sangat berguna untuk memahami urutan eksekusi, mengidentifikasi potensi paralelisme dalam proses, serta menganalisis bottleneck atau area yang perlu dioptimalkan dalam alur sistem.

2.1.8 Database dan DBMS

Sebuah database adalah kumpulan informasi atau data terstruktur yang terorganisir, biasanya disimpan secara elektronik dalam sistem komputer. Database Management System (DBMS) adalah perangkat lunak yang berinteraksi dengan pengguna akhir, aplikasi, dan database itu sendiri untuk menangkap dan menganalisis data. Sebuah DBMS menyediakan cara sistematis untuk membuat, mengambil, memperbarui, dan mengelola data.

Sofyan, Pratama, dan Dewi (2024) berfokus pada pengembangan DBMS menggunakan MySQL, menekankan peran penting sistem ini dalam mengelola dan mengatur data secara efisien. Kemampuan DBMS untuk menangani volume informasi yang besar dan menyediakan akses yang andal adalah fundamental untuk hampir semua aplikasi modern, dari situs web sederhana hingga sistem perusahaan yang kompleks. Lebih lanjut, Rizki, Nuraini, dan Wahyudi (2025) menggarisbawahi pentingnya keamanan database dengan melakukan tinjauan literatur sistematis tentang kerentanan deteksi SQL injection menggunakan machine learning. Ini menyoroti bahwa meskipun database dan DBMS sangat penting, keamanannya terhadap serangan berbahaya seperti SQL injection adalah kekhawatiran utama yang memerlukan tindakan perlindungan canggih.

2.1.9 MySQL

. MySQL adalah sistem manajemen database relasional (RDBMS) open-source berdasarkan Structured Query Language (SQL). Ini adalah pilihan populer untuk

aplikasi web dan banyak digunakan karena keandalan, kinerja, dan kemudahan penggunaannya.

Penelitian oleh Putra dan Sari (2024) tentang pelatihan siswa SMK dalam MySQL dan PHP, serta Sofyan, Pratama, dan Dewi (2024) tentang pengembangan DBMS menggunakan MySQL, dengan jelas menunjukkan prevalensi dan pentingnya MySQL dalam implementasi database praktis. Sifat open-source dan fitur-fiturnya yang kuat menjadikannya solusi utama bagi pengembang dan organisasi yang ingin menyimpan dan mengelola data mereka secara efektif. Fokus berkelanjutan pada keamanannya, seperti yang terlihat dalam karya Rizki, Nuraini, dan Wahyudi (2025) mengenai deteksi SQL injection, lebih lanjut menekankan peran pentingnya dalam lanskap digital kontemporer.

2.1.10 Software dan Pemrograman Terkait

Bagian ini akan menguraikan secara rinci perangkat lunak dan bahasa pemrograman utama yang menjadi fondasi dalam pengembangan dan implementasi sistem klasifikasi wilayah prioritas penambahan tenaga medis. Pemilihan alat-alat ini didasarkan pada kapabilitas, popularitas di kalangan data scientist, serta kesesuaiannya dengan tujuan proyek.

2.1.10.1 Editor Jupyter

Jupyter Editor, yang sering ditemukan dalam bentuk Jupyter Notebook, adalah lingkungan komputasi interaktif. Alat ini memungkinkan kita untuk menggabungkan

kode yang bisa dijalankan, persamaan, visualisasi, dan teks naratif dalam satu dokumen. Fleksibilitasnya menjadikan Jupyter pilihan populer dalam ilmu data, pembelajaran mesin (machine learning), dan riset akademik.

Menurut Handayani, Syahputra, dan Maulana (2022), Jupyter Notebook dapat diterapkan untuk menganalisis kriteria plagiarisme menggunakan teknik semantik. Ini menunjukkan kemampuan Jupyter tidak hanya untuk mengeksekusi kode, tetapi juga untuk menyajikan proses analitik yang kompleks dan hasilnya secara bertahap dan jelas. Hal ini sangat penting untuk tugas-tugas seperti deteksi plagiarisme yang melibatkan analisis teks mendetail dan penerapan algoritma.

2.1.10.2 Bahasa Pemrograman Python dan Library Terkait

Python adalah bahasa pemrograman tingkat tinggi yang serbaguna, dikenal karena kesederhanaan, keterbacaan, dan ekosistem library (perpustakaan) yang luas. Kualitas-kualitas ini menjadikannya pilihan utama untuk berbagai aplikasi, termasuk pengembangan web, analisis data, kecerdasan buatan, dan komputasi ilmiah. Library terkait secara signifikan memperluas fungsionalitas Python, menawarkan modul siap pakai untuk tugas-tugas khusus.

Lestari, Widodo, dan Santoso (2023) menunjukkan implementasi machine learning menggunakan library Python, khususnya Scikit-learn. Ini menggarisbawahi peran sentral Python dalam bidang machine learning, di mana library-nya yang kuat menyediakan alat yang diperlukan untuk tugas-tugas seperti manipulasi data, pelatihan

model, dan evaluasi. Kemudahan integrasi library ini memungkinkan peneliti dan pengembang untuk membangun dan menerapkan solusi machine learning yang kompleks secara efisien.

2.1.10.3 Library Scikit-learn

Scikit-learn adalah library machine learning perangkat lunak gratis untuk bahasa pemrograman Python. Ini menampilkan berbagai algoritma klasifikasi, regresi, dan pengelompokan, serta dirancang untuk beroperasi dengan library numerik dan ilmiah Python, yaitu NumPy dan SciPy.

Seperti yang disoroti oleh Lestari, Widodo, dan Santoso (2023), Scikit-learn adalah library fundamental untuk mengimplementasikan algoritma machine learning. Kumpulan alatnya yang komprehensif memungkinkan praktisi untuk menerapkan berbagai teknik machine learning, dari supervised learning hingga unsupervised learning, untuk memecahkan masalah dunia nyata. API yang ramah pengguna dan dokumentasi yang luas berkontribusi pada adopsi luasnya di lingkungan akademik dan industri untuk mengembangkan sistem cerdas.

2.1.10.4 Bahasa Pemrograman PHP

PHP (Hypertext Preprocessor) adalah bahasa scripting open-source populer yang sangat cocok untuk pengembangan web dan dapat disematkan ke dalam HTML. Umumnya digunakan untuk server-side scripting, di mana ia menghasilkan konten halaman web dinamis, memproses data formulir, dan berinteraksi dengan database.

Putra dan Sari (2024) membahas pelatihan siswa SMK dalam PHP bersama MySQL untuk interaksi database. Ini menunjukkan relevansi PHP yang berkelanjutan dalam pengembangan web, terutama untuk membuat aplikasi web dinamis yang bergantung pada konektivitas database. Kemudahan pembelajaran dan penggunaannya yang luas dalam sistem manajemen konten serta platform e-commerce menjadikannya keterampilan yang berharga bagi calon pengembang web.

2.2 Kajian Penelitian Terdahulu

Penelitian terkait dan yang relevan dengan topik "Klasifikasi Wilayah Prioritas Penambahan Tenaga Medis Berdasarkan Urgensi Penyakit dan Sebaran Fasilitas Kesehatan Menggunakan SVM di Kota Bekasi" difokuskan pada penggunaan algoritma Support Vector Machine (SVM) untuk klasifikasi dalam bidang kesehatan. Berikut beberapa paper penelitiannya:

2.2.1 Paper 1

Penelitian yang dilakukan oleh Nurqotimah, Khudori, & Pradini (2024) berjudul "Implementasi Algoritma Support Vector Machine (SVM) untuk Klasifikasi Penyakit Stroke" yang dipublikasikan pada *Journal of Applied Computer Science and Technology (JACOST)*, 5(1), pp. 1–8. Penelitian ini membahas penggunaan algoritma Support Vector Machine (SVM) untuk mengklasifikasikan atau mendiagnosis penyakit stroke berdasarkan data pasien. Latar belakang penelitian ini menyoroti stroke sebagai masalah kesehatan global yang memerlukan diagnosis cepat dan akurat. Model klasifikasi ini bertujuan untuk membantu tenaga medis dalam mendeteksi risiko stroke

lebih awal, berdasarkan faktor-faktor seperti tekanan darah, kadar gula darah, riwayat penyakit, dan gaya hidup.

Penelitian ini sangat relevan dengan skripsi ini karena memiliki kesamaan fundamental dalam hal penggunaan algoritma Support Vector Machine (SVM) sebagai metode utama yang diterapkan. Selain itu, tujuan utamanya adalah melakukan klasifikasi dalam domain kesehatan (klasifikasi penyakit). Kesamaan ini menjadi dasar kuat untuk mengadopsi dan menyesuaikan pendekatan SVM dalam mengklasifikasikan wilayah prioritas penambahan tenaga medis berdasarkan urgensi penyakit.

2.2.1.1 Tujuan Penelitian

Tujuan penelitian yang dilakukan oleh Nurqotimah, Khudori, & Pradini (2024) adalah untuk membangun dan mengimplementasikan model klasifikasi menggunakan SVM guna mendiagnosis atau mengidentifikasi kategori penyakit stroke pada pasien. Dengan demikian, penelitian ini diharapkan dapat memberikan alat bantu yang efektif bagi tenaga kesehatan dalam membuat keputusan diagnosis yang lebih cepat dan akurat, serta memfasilitasi penanganan pasien stroke secara lebih efisien. Untuk menguji model klasifikasi yang dihasilkan, penelitian ini kemungkinan menggunakan metrik evaluasi seperti akurasi, presisi, *recall*, atau F1-score.

2.2.1.2 Metodologi yang Digunakan

Penelitian ini menggunakan dataset yang berisi informasi pasien dengan berbagai fitur yang relevan untuk identifikasi stroke. Metode yang diterapkan adalah Support Vector Machine (SVM). Tahap pra-pemrosesan data kemungkinan meliputi

penanganan *missing values*, normalisasi fitur, dan pembagian dataset menjadi data *training* dan *testing*. Evaluasi kinerja model SVM dilakukan dengan metrik yang sesuai untuk masalah klasifikasi, seperti akurasi, presisi, *recall*, dan F1-score. Hasil evaluasi yang baik menunjukkan kemampuan model dalam mengklasifikasikan kasus stroke secara efektif.

2.2.1.3 Temuan Utama

Hasil dari penelitian ini menunjukkan bahwa model SVM yang dibangun memiliki kinerja yang baik dalam klasifikasi penyakit stroke, dengan tingkat akurasi yang tinggi. Hal ini menandakan bahwa SVM mampu mengidentifikasi pola-pola kompleks dalam data pasien yang membedakan antara pasien stroke dan non-stroke. Penelitian ini juga kemungkinan berhasil mengidentifikasi fitur-fitur dominan yang paling berpengaruh terhadap kejadian stroke, seperti tekanan darah tinggi atau riwayat diabetes, yang memberikan wawasan penting bagi pencegahan dan penanganan penyakit.

2.2.1.4 Kesimpulan Penelitian

Penelitian ini membuktikan bahwa algoritma Support Vector Machine (SVM) sangat efektif dalam masalah klasifikasi data di bidang kesehatan. Studi ini berhasil membangun model prediksi yang akurat untuk klasifikasi penyakit stroke berdasarkan data pasien. Keberhasilan ini menegaskan potensi SVM sebagai alat bantu diagnosis yang dapat meningkatkan kecepatan dan akurasi identifikasi kondisi medis. Dalam konteks skripsi ini, pendekatan serupa dapat diaplikasikan untuk mengklasifikasikan

wilayah berdasarkan tingkat urgensi penyakit dan sebaran fasilitas kesehatan, guna menentukan prioritas penambahan tenaga medis.

2.2.2 Paper 2

Penelitian yang dilakukan oleh Achmad, Adnan, & Rijal (2022) berjudul "Klasifikasi Penyakit Pernapasan Berbasis Visualisasi Suara Menggunakan Metode Support Vector Machine" yang dipublikasikan pada *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)*, 6(1), pp. 1–8. doi: 10.47970/siskom-kb.v6i1.330. Penelitian ini berfokus pada penggunaan Support Vector Machine (SVM) untuk mengklasifikasikan jenis-jenis penyakit pernapasan berdasarkan analisis visualisasi suara batuk atau pernapasan. Latar belakang penelitian menyoroti tantangan diagnosis dini penyakit pernapasan dan potensi *machine learning* dalam membantu proses tersebut melalui analisis suara yang non-invasif.

Penelitian ini juga sangat relevan dengan skripsi ini karena secara eksplisit menggunakan algoritma Support Vector Machine (SVM) untuk klasifikasi dalam konteks kesehatan. Meskipun datanya berbeda (suara dibandingkan data demografi dan penyakit), inti metodologi klasifikasi SVM pada data yang kompleks tetap sama. Relevansi ini mendukung argumen bahwa SVM adalah pilihan yang kuat untuk masalah klasifikasi kompleks di bidang kesehatan, termasuk klasifikasi wilayah prioritas.

2.2.2.1 Tujuan Penelitian

Tujuan utama penelitian Achmad, Adnan, & Rijal (2022) adalah mengembangkan model klasifikasi berbasis SVM yang dapat secara otomatis mengidentifikasi dan membedakan berbagai jenis penyakit pernapasan berdasarkan fitur-fitur yang diekstraksi dari visualisasi suara. Hal ini bertujuan untuk menciptakan alat diagnosis awal yang cepat dan objektif, terutama di daerah dengan keterbatasan akses terhadap *expert* medis.

2.2.2.2 Metodologi yang Digunakan

Penelitian ini mengumpulkan data suara pasien dengan berbagai kondisi penyakit pernapasan. Data suara ini kemudian diubah menjadi representasi visual (misalnya, spektrumgram) atau fitur-fitur numerik yang dapat diproses oleh algoritma. Metode Support Vector Machine (SVM) digunakan untuk melatih model klasifikasi. Proses ini melibatkan ekstraksi fitur dari visualisasi suara, normalisasi data, dan *tuning* parameter SVM untuk mendapatkan kinerja terbaik. Evaluasi model dilakukan dengan metrik klasifikasi standar seperti akurasi, presisi, *recall*, dan F1-score, untuk menilai kemampuan model dalam mengklasifikasikan penyakit pernapasan secara akurat.

2.2.2.3 Temuan Utama

Temuan kunci dari penelitian ini adalah bahwa SVM menunjukkan kinerja yang menjanjikan dalam mengklasifikasikan penyakit pernapasan berdasarkan visualisasi suara. Model yang dikembangkan mampu membedakan antara kondisi pernapasan yang berbeda dengan tingkat akurasi yang memuaskan. Hal ini membuka peluang baru

untuk diagnosis non-invasif dan pemantauan kondisi pernapasan menggunakan teknologi suara.

2.2.2.4 Kesimpulan Penelitian

Penelitian ini menegaskan efektivitas Support Vector Machine (SVM) dalam tugas klasifikasi data non-tradisional (visualisasi suara) untuk tujuan diagnostik kesehatan. Ini menunjukkan bahwa SVM memiliki kapabilitas untuk menangani kompleksitas data untuk mengidentifikasi pola relevan yang mengindikasikan kondisi kesehatan tertentu. Dalam skripsi ini, kemampuan SVM dalam mengklasifikasikan berdasarkan *fitur* data yang kompleks dapat diterapkan untuk mengelompokkan wilayah-wilayah di Kota Bekasi berdasarkan kombinasi urgensi penyakit dan sebaran fasilitas kesehatan.

2.2.3 Paper 3

Penelitian yang dilakukan oleh Nurochman & Ashiilah (2024) berjudul "Perbandingan SVM dan LSTM untuk Memprediksi Gangguan Kecemasan Berdasarkan Cuitan di Twitter" yang dipublikasikan pada *Jurnal Teknik Komputer (JTK)*, 10(2), pp. 45–52. Penelitian ini membandingkan kinerja Support Vector Machine (SVM) dengan *Long Short-Term Memory* (LSTM) dalam memprediksi gangguan kecemasan berdasarkan analisis sentimen cuitan di platform Twitter. Latar belakangnya adalah peningkatan masalah kesehatan mental dan potensi *social media* sebagai sumber data untuk deteksi dini.

Penelitian ini relevan karena secara langsung melibatkan algoritma Support Vector Machine (SVM) dan membandingkannya dengan algoritma *machine learning* lain untuk tugas klasifikasi dalam domain kesehatan/psikologi. Meskipun fokusnya adalah teks dan kesehatan mental, studi ini menggarisbawahi kekuatan SVM sebagai algoritma klasifikasi yang kompetitif, yang relevan untuk dipertimbangkan dalam masalah klasifikasi wilayah prioritas yang juga melibatkan data yang kompleks dan potensi perbandingan algoritma.

2.2.3.1 Tujuan Penelitian

Tujuan penelitian Nurochman & Ashiilah (2024) adalah untuk mengevaluasi dan membandingkan performa model klasifikasi menggunakan SVM dan LSTM dalam mendeteksi dan memprediksi gangguan kecemasan dari data teks (cuitan Twitter). Hal ini bertujuan untuk mengidentifikasi algoritma mana yang lebih efektif dalam menangkap pola-pola linguistik terkait kondisi kecemasan, sehingga dapat berkontribusi pada pengembangan sistem deteksi dini kesehatan mental.

2.2.3.2 Metodologi yang Digunakan

Penelitian ini mengumpulkan data cuitan dari Twitter yang kemudian melalui tahap pra-pemrosesan seperti *tokenization*, *stop-word removal*, dan *stemming*. Fitur-fitur teks kemudian diekstraksi. Dua model, yaitu Support Vector Machine (SVM) dan *Long Short-Term Memory* (LSTM), dilatih pada dataset ini. Metrik evaluasi seperti akurasi, presisi, *recall*, dan F1-score digunakan untuk membandingkan kinerja kedua

algoritma dalam mengklasifikasikan cuitan sebagai indikasi gangguan kecemasan atau tidak.

2.2.3.3 Temuan Utama

Hasil dari penelitian ini menunjukkan bahwa baik SVM maupun LSTM mampu melakukan klasifikasi gangguan kecemasan berdasarkan cuitan Twitter dengan cukup baik. Perbandingan detail menunjukkan bahwa mungkin ada skenario di mana satu algoritma sedikit mengungguli yang lain, tergantung pada kompleksitas pola dalam data teks. Namun, secara umum, SVM menunjukkan performa yang kompetitif, menegaskan kapabilitasnya dalam menangani data teks untuk tugas klasifikasi yang kompleks.

2.2.3.4 Kesimpulan Penelitian

Penelitian ini mengkonfirmasi bahwa Support Vector Machine (SVM) adalah algoritma klasifikasi yang kuat dan dapat bersaing dengan metode *deep learning* seperti LSTM, bahkan dalam konteks analisis teks dan kesehatan mental. Ini menunjukkan fleksibilitas SVM dalam menangani berbagai jenis data dan masalah klasifikasi. Temuan ini semakin memperkuat dasar pemikiran untuk menggunakan SVM dalam skripsi ini, di mana data spasial (sebaran fasilitas) dan data epidemiologis (urgensi penyakit) akan digunakan untuk mengklasifikasikan wilayah prioritas penambahan tenaga medis.

Kelima penelitian ini secara keseluruhan memberikan dasar teoritis dan metodologis yang kuat untuk mendukung penelitian skripsi Anda. Pemilihan algoritma

Support Vector Machine (SVM) terbukti sangat tepat karena kemampuannya dalam melakukan klasifikasi data yang kompleks dan beragam di bidang kesehatan. SVM mampu mengidentifikasi pola-pola rumit dari urgensi penyakit dan sebaran fasilitas kesehatan, yang krusial untuk menentukan wilayah prioritas penambahan tenaga medis di Kota Bekasi secara akurat dan berbasis data.

