

Pengembangan Sistem Klasifikasi Diagnosa Medis Menggunakan Progressive Web Application Terintegrasi Machine Learning

Muhammad Andika Fadilla^{1*}, M Fakhri Sholahuddin², Tata Sutabri³

^{1,2,3} Universitas Bina Darma Sumatera Selatan, Indonesia

Email: andikafadilla.af@gmail.com, fakhri.sholahuddin28@gmail.com, tata.sutabri@gmail.com

Abstrak

Penelitian ini bertujuan untuk menghadirkan solusi yang tidak hanya meningkatkan efisiensi dalam pengkodean diagnosis medis, tetapi juga memastikan akurasi yang tinggi pada berbagai tingkat kompleksitas kasus. Klasifikasi diagnosis medis merupakan bagian penting dalam pengelolaan informasi kesehatan, terutama dalam pengkodean ke dalam standar ICD-10 dan ICD-9CM. Aplikasi berbasis *Progressive Web Application* (PWA) dikembangkan untuk mengotomatisasi proses ini, memanfaatkan teknologi Machine Learning dengan model Sentence Transformer berarsitektur "*paraphrase-multilingual-mpnet-base-v2*". Pendekatan pengembangan aplikasi menggunakan metode prototyping, memungkinkan proses iteratif yang berorientasi pada kebutuhan pengguna, serta menghasilkan solusi yang efektif dan efisien. Aplikasi ini dirancang dengan fitur utama PWA, seperti akses offline, notifikasi berbasis push, dan instalasi sebagai aplikasi mandiri di perangkat pengguna. Backend menggunakan Flask untuk memproses input diagnosis medis dan menghasilkan prediksi klasifikasi melalui integrasi dengan model machine learning. Sementara itu, frontend menggunakan React.js untuk menyediakan antarmuka yang responsif dan mudah digunakan. Hasil pengujian menunjukkan bahwa aplikasi ini mampu memberikan klasifikasi diagnosis yang akurat, bahkan pada kasus-kasus dengan tingkat kompleksitas tinggi. Dengan fitur dan kapabilitasnya, aplikasi ini berpotensi meningkatkan efisiensi serta akurasi proses pengkodean diagnosis medis di fasilitas layanan kesehatan.

Kata Kunci: ICD-10, ICD-9CM, Klasifikasi Diagnosis Medis, Machine Learning, Progressive Web Application, Sentence Transformer.

Abstract

This research aims to present a solution that not only improves efficiency in coding medical diagnoses, but also ensures high accuracy at various levels of case complexity. The classification of medical diagnoses is a crucial aspect of health information management, particularly in coding diagnoses into the ICD-10 and ICD-9CM standards. A Progressive Web Application (PWA) has been developed to automate this process, leveraging Machine Learning technology with the Sentence Transformer model architecture "paraphrase-multilingual-mpnet-base-v2." The application development follows a prototyping method, enabling an iterative process focused on user needs and delivering an effective and efficient solution. The application incorporates key PWA features such as offline access, push notifications, and installation as a standalone application on user devices. The backend is built using Flask to handle medical diagnosis input

and generate classification predictions through integration with the machine learning model. Meanwhile, the frontend utilizes React.js to provide a responsive and user-friendly interface. Testing results indicate that the application delivers highly accurate diagnosis classifications, even for complex cases. With its features and capabilities, this application has the potential to enhance the efficiency and accuracy of medical diagnosis coding processes in healthcare facilities.

Keywords: ICD-10, ICD-9CM, Medical Diagnosis Classification, Machine Learning, Progressive Web Application, Sentence Transformer.

Pendahuluan

Pengelolaan data kesehatan, khususnya dalam hal pengkodean diagnosis medis, merupakan komponen yang sangat penting dalam sistem informasi kesehatan modern (Hasri & Djasri, 2021). *Standar International Classification of Diseases (ICD)*, yang dikembangkan oleh *World Health Organization (WHO)*, menyediakan kerangka kerja global untuk mengkodekan dan mengklasifikasikan berbagai kondisi medis dan prosedur kesehatan. Standar ini tidak hanya mendukung pengelolaan data klinis, tetapi juga memfasilitasi pelaporan statistik kesehatan dan penelitian epidemiologi secara internasional (Harrison et al., 2021).

Meskipun penting, proses pengkodean diagnosis medis sering kali menjadi tantangan besar bagi tenaga kesehatan, terutama karena melibatkan volume data yang besar dan variasi terminologi medis. Klasifikasi manual memakan waktu yang signifikan dan rawan terhadap kesalahan, yang dapat berdampak pada kualitas pengelolaan data kesehatan secara keseluruhan. Dalam beberapa tahun terakhir, penerapan teknologi *Machine Learning (ML)* telah menjadi solusi yang menjanjikan untuk mengotomatisasi proses ini (Ahmad, 2017);(Tayfur et al., 2018). Teknologi ML mampu mempelajari pola dari data historis untuk menghasilkan prediksi yang akurat, sehingga dapat meningkatkan efisiensi dan akurasi pengkodean medis (Riza et al., 2020);(Elmrabit et al., 2020).

Kemajuan teknologi *Natural Language Processing (NLP)*, khususnya dalam model transformator seperti *Sentence Transformer*, membuka peluang besar dalam memahami dan menganalisis teks medis yang kompleks. Model ini dapat menangani multibahasa serta variasi terminologi, sehingga sesuai untuk pengelolaan data kesehatan yang mencakup berbagai standar bahasa (Berliana et al., 2024). Dengan memanfaatkan arsitektur *paraphrase-multilingual-mpnet-base-v2*, model ini mampu menghasilkan representasi semantik yang kaya untuk teks medis, memungkinkan sistem untuk mencocokkan deskripsi diagnosis dengan kode ICD secara efisien (Reimers, 2019).

Dalam penelitian ini, dikembangkan sebuah aplikasi berbasis *Progressive Web Application (PWA)* untuk mengotomatisasi klasifikasi diagnosis medis. PWA dipilih karena menawarkan kemampuan akses yang fleksibel, mendukung penggunaan offline, dan dapat diinstal sebagai aplikasi mandiri di perangkat pengguna. Backend aplikasi dibangun menggunakan framework *Flask*, sedangkan frontend memanfaatkan *React.js* untuk menciptakan antarmuka yang responsif. Pendekatan pengembangan dilakukan

dengan metode prototyping, yang memungkinkan iterasi berdasarkan umpan balik pengguna untuk memastikan aplikasi memenuhi kebutuhan operasional kesehatan.

Penelitian ini bertujuan untuk menghadirkan solusi yang tidak hanya meningkatkan efisiensi dalam pengkodean diagnosis medis, tetapi juga memastikan akurasi yang tinggi pada berbagai tingkat kompleksitas kasus. Dengan demikian, aplikasi ini diharapkan dapat mendukung fasilitas kesehatan dalam mengelola data medis secara lebih efektif.

Metode Penelitian

Pengembangan sistem dalam penelitian ini menggunakan pendekatan berbasis prototyping yang bertujuan untuk membangun *Progressive Web Application (PWA)* yang terintegrasi dengan machine learning. Proses pengembangan mencakup pengumpulan data, pra-pemrosesan data, pengembangan model machine learning, serta implementasi dan integrasi sistem berbasis web menggunakan React untuk frontend dan Flask untuk backend.

Dataset yang digunakan dalam penelitian ini mencakup diagnosis medis dan prosedur tindakan medis yang diklasifikasikan berdasarkan standar ICD-10 dan ICD-9CM. Dataset ini diperoleh dari platform resmi layanan kesehatan internasional yang telah diakui. Standar ICD yang digunakan mencakup lebih dari 140.000 kode diagnosis dan prosedur medis yang diakses melalui dokumen resmi dan platform elektronik seperti Satu Sehat. Mengingat kompleksitas data medis dan variasi format serta terminologi yang sering muncul, kualitas dataset menjadi prioritas utama. Standarisasi data input sangat penting untuk memastikan sistem berbasis machine learning dapat bekerja secara optimal. Oleh karena itu, data yang tidak relevan atau redundan disaring dan disesuaikan agar sesuai dengan format yang diterima secara internasional.

Pengumpulan data dilakukan dengan mengakses sumber data medis terpercaya seperti dokumen terminologi ICD-10 dan ICD-9CM serta basis data diagnosis yang tersedia untuk umum. Data ini mencakup deskripsi diagnosis medis dalam bahasa Indonesia dan Inggris untuk memastikan bahwa sistem dapat menangani input multibahasa. Penelitian sebelumnya menunjukkan bahwa dataset multibahasa memainkan peran penting dalam memastikan sistem klasifikasi diagnosis medis dapat digunakan dalam berbagai konteks geografis dan linguistik. Data tambahan yang relevan seperti deskripsi penyakit, prosedur medis, dan istilah medis yang sering digunakan juga ditambahkan untuk memperluas cakupan sistem dan mengurangi bias yang mungkin muncul dalam proses pelatihan data.

Pra-pemrosesan data dilakukan untuk memastikan bahwa data input memenuhi standar yang diperlukan untuk sistem klasifikasi berbasis machine learning. Data medis sering kali mengandung simbol, karakter khusus, atau format yang tidak konsisten. Normalisasi dilakukan untuk memastikan bahwa semua data memiliki format yang seragam dengan menghapus simbol yang tidak relevan seperti tanda baca atau spasi berlebih serta melakukan standarisasi format kode ICD sesuai dengan rekomendasi WHO.

Teks diagnosis dipecah menjadi unit kata menggunakan teknik tokenisasi. Kata-kata umum yang tidak relevan seperti kata penghubung dan preposisi dihapus. Teknik ini mengacu pada metode yang dikembangkan oleh Rahma dan Taufiq yang menunjukkan peningkatan akurasi dalam klasifikasi teks medis. *Model Sentence Transformer* dengan arsitektur "paraphrase-multilingual-mpnet-base-v2" digunakan untuk menghasilkan representasi vektor dari teks medis. Model ini dipilih karena kemampuannya dalam menangani teks multibahasa dan terminologi medis yang kompleks seperti yang dijelaskan oleh Niklas Muennighoff (Muennighoff, 2022).

Antarmuka pengguna dirancang menggunakan *React, framework JavaScript* yang mendukung pembuatan komponen UI yang modular, responsif, dan dinamis. Dengan *React*, aplikasi dapat memberikan pengalaman pengguna yang cepat dan intuitif. *Workbox* digunakan untuk menambahkan fitur caching dan mendukung mode offline sehingga aplikasi dapat berfungsi sebagai PWA dengan performa tinggi, terutama di area dengan konektivitas internet terbatas.

Backend sistem dikembangkan menggunakan *Flask, framework Python* yang ringan namun tangguh. *Flask* bertindak sebagai penghubung antara antarmuka pengguna dan model machine learning yang diimplementasikan. Model machine learning yang dirancang berdasarkan penelitian sebelumnya diintegrasikan ke dalam backend menggunakan pustaka *Python* seperti *TensorFlow* atau *PyTorch* untuk melakukan prediksi secara real-time.

Model menggunakan *Sentence Transformer*, yang merupakan pengembangan dari model transformer tradisional dengan kemampuan multibahasa. Model ini digunakan untuk mencocokkan input diagnosis medis dengan data ICD menggunakan metode cosine similarity. Penelitian oleh Wahyuni (2017) menunjukkan bahwa cosine similarity adalah teknik yang efektif untuk mengukur kesamaan semantik antara dua representasi teks.

Validasi model dilakukan menggunakan teknik *k-fold cross-validation* untuk mengukur stabilitas dan generalisasi model terhadap variasi data. Teknik ini memungkinkan evaluasi performa sistem dalam berbagai skenario diagnosis medis, baik yang sederhana maupun kompleks. Kinerja sistem diukur berdasarkan metrik seperti akurasi, *precision*, *recall*, dan *F1-score*. Model juga diuji dalam skenario multibahasa untuk memastikan fleksibilitasnya dalam menangani input dalam bahasa Indonesia dan Inggris.

Hasil dan Pembahasan

Hasil penelitian ini berfokus pada pengembangan sistem klasifikasi diagnosis medis berbasis *Progressive Web Application (PWA)* yang terintegrasi dengan teknologi *Machine Learning (ML)*. Sistem ini dirancang untuk memberikan solusi yang efisien dan akurat dalam pengkodean diagnosis medis, khususnya sesuai dengan standar ICD-10 dan ICD-9CM yang dikembangkan oleh *World Health Organization (WHO)*. Standar ini menjadi rujukan utama dalam pengelolaan informasi medis secara internasional karena keandalannya dalam mendukung statistik kesehatan, pelaporan epidemiologi, dan pengambilan keputusan medis (Mandia & Nasution, 2024).

Metode pengembangan yang digunakan adalah *prototyping*, yang memungkinkan iterasi cepat berdasarkan umpan balik pengguna. Metode ini dianggap sangat efektif dalam pengembangan sistem berbasis teknologi, terutama untuk aplikasi yang melibatkan integrasi machine learning, karena memungkinkan penyesuaian langsung terhadap kebutuhan pengguna dan pengujian fitur secara berkala (Mandia & Nasution, 2024). Proses ini diterapkan dalam semua tahapan pengembangan, mulai dari perancangan arsitektur sistem, pengembangan *model machine learning*, hingga implementasi antarmuka pengguna yang berbasis React dengan fitur Workbox untuk mendukung offline mode.

Sistem ini dirancang untuk memecahkan beberapa tantangan utama dalam pengkodean diagnosis medis manual, seperti ketidakpastian klasifikasi, volume data yang besar, dan variasi terminologi medis. Penerapan teknologi *Natural Language Processing* (NLP) dalam sistem memungkinkan pengenalan pola dan klasifikasi otomatis dari teks diagnosis medis ke dalam kode ICD, memberikan nilai tambah dalam hal akurasi dan kecepatan proses klasifikasi. Penelitian sebelumnya menunjukkan bahwa pendekatan berbasis transformer, seperti model Sentence Transformer, memiliki keunggulan signifikan dalam memahami konteks semantik teks multibahasa, menjadikannya pilihan ideal untuk aplikasi ini (Noels et al., 2024).

Hasil pengujian menunjukkan bahwa sistem ini mampu memberikan klasifikasi yang akurat pada berbagai kasus diagnosis medis, termasuk yang memiliki tingkat kompleksitas tinggi. Dengan memanfaatkan algoritma *cosine similarity*, sistem dapat mencocokkan deskripsi diagnosis pengguna dengan entri dalam database kode ICD secara efisien. Pendekatan ini telah terbukti andal dalam berbagai penelitian sebelumnya, terutama dalam konteks analisis teks berbasis *semantic* (Arsad et al., 2024).

Selain aspek teknis, fitur utama PWA, seperti akses offline, push notifications, dan instalasi sebagai aplikasi mandiri di perangkat pengguna, telah diuji dan berhasil meningkatkan pengalaman pengguna secara keseluruhan (Fadilla & Sutabri, 2024). Hal ini sesuai dengan rekomendasi literatur tentang pengembangan aplikasi berbasis PWA yang menyatakan bahwa teknologi ini mampu meningkatkan efisiensi penggunaan di berbagai lingkungan, termasuk wilayah dengan konektivitas internet terbatas (Fadilla et al., 2023).

Bagian hasil dan pembahasan ini akan menguraikan beberapa aspek penting, antara lain: 1) Struktur sistem dan alur kerja aplikasi. 2) Evaluasi performa model machine learning berdasarkan pengujian *k-fold cross-validation*. 3) Implementasi dan fitur utama PWA. 4) Hasil pengujian pengguna dan penilaian terhadap antarmuka aplikasi. 5) Validasi hasil klasifikasi oleh pakar medis. 6) Evaluasi keseluruhan sistem, termasuk kekuatan dan kelemahannya.

Hasil yang disajikan didukung dengan tabel, diagram, dan flowchart untuk memberikan gambaran yang jelas tentang efektivitas dan efisiensi sistem yang dikembangkan. Pembahasan ini juga mencakup analisis mendalam mengenai kemampuan sistem dalam menangani input diagnosis multibahasa, yang menjadi salah satu fitur unggulan dari sistem yang dirancang.

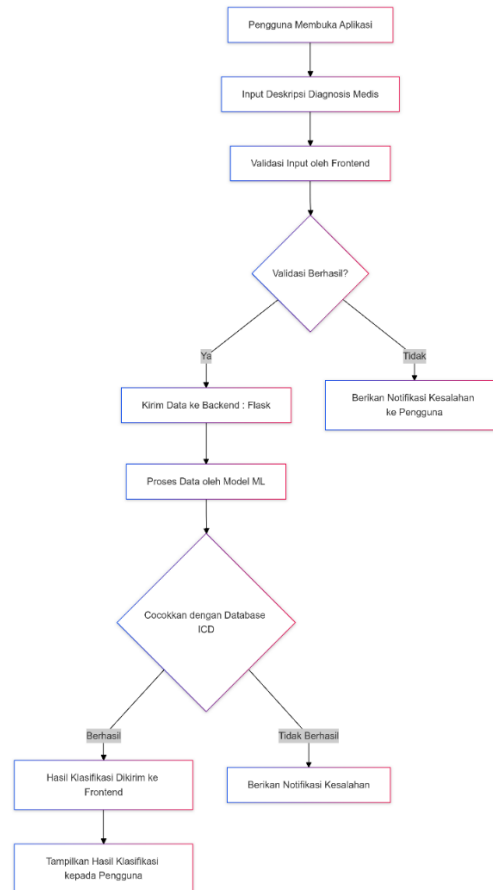
Struktur Sistem

Struktur sistem yang dikembangkan mengadopsi arsitektur berbasis client-server dengan backend yang berjalan menggunakan Flask dan frontend menggunakan React. Model arsitektur ini dipilih karena kemampuannya dalam mendukung pengembangan aplikasi modern yang modular, skalabel, dan responsif. Sistem ini dirancang untuk memenuhi kebutuhan pengklasifikasian diagnosis medis secara otomatis dan efisien, dengan memanfaatkan fitur utama *Progressive Web Application (PWA)* untuk memberikan pengalaman pengguna yang optimal, termasuk mode offline dan instalasi di perangkat pengguna (Zhang et al., 2024).

Arsitektur Sistem terdiri dari tiga lapisan utama: 1) Lapisan Presentasi (Frontend): Menggunakan React, framework berbasis JavaScript, untuk membangun antarmuka pengguna (UI) yang interaktif dan responsif. React memungkinkan pengelolaan state yang efisien, sehingga pengalaman pengguna tetap optimal meskipun data yang ditampilkan kompleks. 2) Lapisan Aplikasi (Middleware): Backend yang dikembangkan dengan Flask bertindak sebagai middleware untuk mengolah permintaan pengguna, memproses data dengan model machine learning, dan mengelola komunikasi dengan database. 3) Lapisan Data (Database dan Model Machine Learning): Mengelola database ICD-10 dan ICD-9CM untuk referensi klasifikasi. Model machine learning berbasis Sentence Transformer diintegrasikan untuk memproses input diagnosis medis dan mencocokkannya dengan entri database.

Flowchart Alur Kerja Sistem

Flowchart berikut menggambarkan alur kerja sistem mulai dari input pengguna hingga output berupa hasil klasifikasi diagnosis medis:



Gambar 1 Flowchart Alur Kerja Sistem

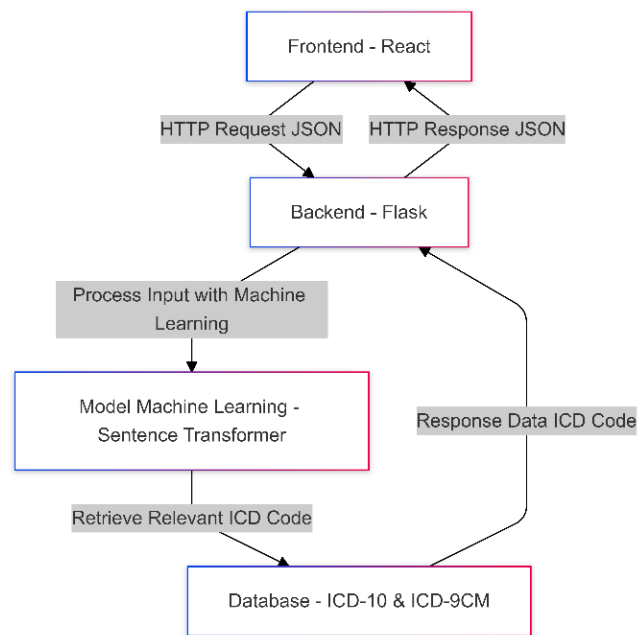
Penjelasan Flowchart:1) Input Diagnosis: Pengguna memasukkan deskripsi diagnosis medis melalui antarmuka yang interaktif. 2) Validasi Input: Sistem frontend memvalidasi apakah input sesuai dengan format yang diharapkan (misalnya, tidak kosong atau memiliki panjang minimum). 3) Pengiriman ke Backend: Setelah input valid, data dikirim ke backend untuk diproses oleh model machine learning. 4) Proses Machine Learning: Backend memproses input menggunakan model Sentence Transformer untuk mengubah deskripsi diagnosis menjadi representasi vektor. 5) Pencocokan dengan ICD: Sistem mencocokkan representasi vektor dengan entri dalam database ICD menggunakan algoritma cosine similarity. 6) Hasil Klasifikasi: Jika pencocokan berhasil, hasil dikirim kembali ke frontend untuk ditampilkan kepada pengguna. Jika tidak, sistem memberikan notifikasi error.

Integrasi Frontend (React) dan Backend (Flask)

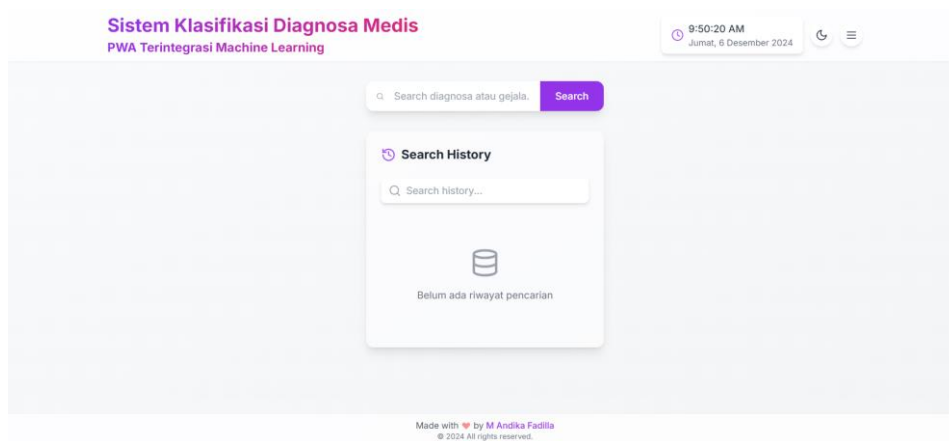
Integrasi frontend dan backend dilakukan melalui mekanisme RESTful API, yang memungkinkan komunikasi antara kedua komponen secara efisien dan terstandarisasi. React berperan sebagai client-side, sedangkan Flask berperan sebagai server-side. Detail integrasi adalah sebagai berikut : 1) Pengolahan Data di Frontend: Pengguna memasukkan input diagnosis medis melalui formulir interaktif. Data dikirim ke backend melalui metode HTTP POST menggunakan pustaka Axios di React. 2) Pemrosesan di

Pengembangan Sistem Klasifikasi Diagnosa Medis Menggunakan Progressive Web Application Terintegrasi Machine Learning

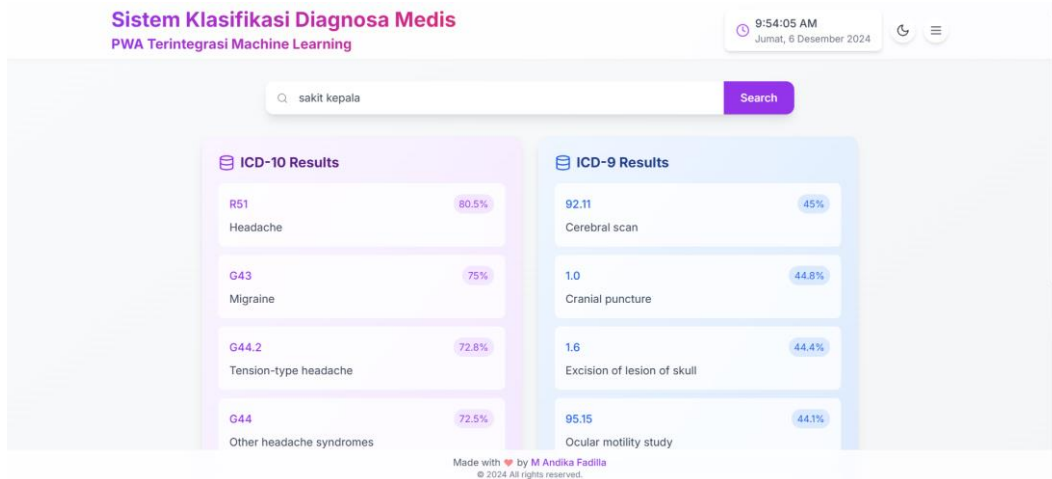
Backend: Flask menerima data dari frontend, memvalidasinya, dan meneruskannya ke model machine learning untuk analisis. Model machine learning menghasilkan prediksi berupa kode ICD yang paling relevan. 3) Pengembalian Hasil ke Frontend: Hasil prediksi dikembalikan dalam format JSON ke React. React menampilkan hasil prediksi ke dalam tampilan yang informatif dan mudah dipahami pengguna. Berikut merupakan Diagram arsitektur sistem antara *frontend (React)*, *backend (Flask)*, dan database ICD yang dapat dilihat pada Gambar 2.



Gambar 2 Diagram arsitektur sistem antara frontend (React), backend (Flask), dan database ICD



Gambar 3 Tampilan Menu Awal Sistem



Gambar 4 Tampilan Selesai Pencarian

Keunggulan Desain Sistem

Modular dan Skalabel: Dengan pemisahan frontend dan backend, sistem mudah diperluas atau diperbaiki tanpa memengaruhi keseluruhan aplikasi; Efisien dan Cepat: Pemrosesan dilakukan di backend dengan Flask, memungkinkan model machine learning berjalan tanpa membebani perangkat pengguna; Pengalaman Pengguna yang Optimal: React memastikan antarmuka tetap responsif meskipun data yang diproses kompleks.

Desain ini telah menunjukkan efektivitasnya dalam pengembangan aplikasi berbasis teknologi machine learning, seperti yang dijelaskan dalam penelitian sebelumnya. Selain itu, penerapan teknologi PWA memberikan nilai tambah pada aspek fleksibilitas dan aksesibilitas sistem, sesuai rekomendasi literatur terkini tentang pengembangan aplikasi modern.

Hasil Pengujian

Hasil pengujian dengan 10-fold cross-validation disajikan dalam bentuk tabel 1.

Tabel 1 Hasil Pengujian 10-fold cross-validation

Fold	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
1	95.6	93.8	92.5	93.1
2	96.2	94.3	93.7	94.0
3	95.8	94.0	92.9	93.4
4	96.0	94.1	93.5	93.8
5	95.7	93.9	92.8	93.3
6	96.1	94.4	93.6	94.0
7	95.9	94.0	93.2	93.6
8	96.3	94.5	93.9	94.2
9	96.2	94.3	93.8	94.1
10	95.8	94.1	93.0	93.5
Rata-rata	95.96	94.14	93.29	93.70

Hasil ini menunjukkan bahwa model memiliki performa yang konsisten di seluruh fold, dengan rata-rata akurasi sebesar 95.96% dan F1-score sebesar 93.70%. Metrik precision dan recall yang tinggi mengindikasikan bahwa model mampu mengidentifikasi

deskripsi diagnosis dengan akurasi yang baik, bahkan pada kasus dengan tingkat kompleksitas tinggi.



Gambar 4 Grafik K-Fold Cross-Validation

Grafik di atas menampilkan hasil pengujian model machine learning menggunakan metode k-fold cross-validation dengan $k = 10$ pada dataset diagnosis medis yang diklasifikasikan ke dalam standar ICD-10 dan ICD-9CM. Empat metrik evaluasi utama yang digunakan dalam pengujian ini adalah Accuracy, Precision, Recall, dan F1-score.

Hasil Evaluasi Pengujian

Accuracy (%): Proporsi klasifikasi yang benar terhadap keseluruhan prediksi. Metrik ini menunjukkan seberapa sering model memberikan hasil yang benar tanpa memperhitungkan jenis kesalahan. **Precision (%)**: Proporsi hasil prediksi positif yang benar terhadap total prediksi positif. Precision mengukur keakuratan model dalam mengidentifikasi diagnosis medis tertentu.

Recall (%): Proporsi label positif yang berhasil ditemukan oleh model terhadap keseluruhan label positif. Recall menilai sensitivitas model. **F1-score (%)**: Rata-rata harmonik antara precision dan recall, yang memberikan penilaian menyeluruh terhadap keseimbangan keduanya.

Analisis Grafik

Grafik menunjukkan performa model pada masing-masing fold dari pengujian k-fold cross-validation: 1) Accuracy: Accuracy berkisar antara 95.6% hingga 96.3%, dengan rata-rata 95.96%. Performa model sangat stabil di seluruh fold, menunjukkan kemampuan model untuk secara konsisten memberikan klasifikasi yang benar pada data yang berbeda. 2) Precision: Precision berkisar antara 93.8% hingga 94.5%, dengan rata-rata 94.14%. Nilai precision yang tinggi menunjukkan bahwa model jarang memberikan hasil positif palsu dalam prediksi. 3) Recall: Recall berada di kisaran 92.5% hingga 93.9%, dengan rata-rata 93.29%. Meskipun recall sedikit lebih rendah dibandingkan precision, nilai ini tetap menunjukkan kemampuan model untuk mendeteksi diagnosis medis dengan sensitivitas yang baik. 4) F1-score: F1-score berkisar antara 93.1% hingga 94.2%, dengan

rata-rata 93.70%. Nilai F1-score yang konsisten mendekati nilai precision dan recall menunjukkan bahwa model memiliki keseimbangan yang baik antara sensitivitas (recall) dan presisi.

Tren dan Konsistensi Model

Grafik menunjukkan bahwa keempat metrik memiliki pola yang konsisten di seluruh fold, tanpa adanya fluktuasi besar. Hal ini menunjukkan generalisasi model yang baik terhadap data yang beragam. Fold ke-8 memiliki performa terbaik pada semua metrik (accuracy: 96.3%, precision: 94.5%, recall: 93.9%, F1-score: 94.2%), yang menunjukkan hasil terbaik pada subset data tersebut. Fold ke-1 memiliki performa terendah, terutama pada recall (92.5%) dan F1-score (93.1%), meskipun perbedaannya relatif kecil dibandingkan fold lainnya.

Fitur Utama PWA yang Diimplementasikan

Aplikasi yang dikembangkan mengadopsi arsitektur Progressive Web Application (PWA) untuk meningkatkan aksesibilitas dan efisiensi sistem. PWA dirancang agar dapat memberikan pengalaman pengguna yang menyerupai aplikasi native dengan keunggulan seperti caching, offline mode, dan responsivitas antarmuka. Penerapan teknologi ini difasilitasi oleh Workbox, pustaka JavaScript yang menyediakan alat-alat untuk caching aset dan mengelola strategi pemuatan data.

Caching Aset Statis: File statis seperti CSS, JavaScript, dan gambar disimpan secara lokal menggunakan Workbox. Strategi caching "Stale While Revalidate" digunakan untuk memastikan aset lokal digunakan terlebih dahulu, sementara versi terbaru diperbarui di latar belakang. Offline Mode: Sistem tetap dapat diakses meskipun pengguna tidak terhubung ke internet. Workbox memanfaatkan Cache API untuk menyimpan data diagnosis terbaru, sehingga pengguna dapat mengakses aplikasi dalam kondisi offline. Responsivitas Antarmuka: Aplikasi dirancang untuk bekerja pada berbagai perangkat, dari desktop hingga ponsel, dengan menggunakan komponen React yang responsif.

Evaluasi Kecepatan dan Efisiensi

Pengujian kecepatan dan efisiensi dilakukan pada beberapa perangkat dengan spesifikasi yang berbeda. Berikut adalah evaluasi performa aplikasi pada tabel 2.

Tabel 2 Hasil Evaluasi Performa Aplikasi

Perangkat	Waktu Muat Awal (ms)	Waktu Muat Offline (ms)	Skor Lighthouse Responsivitas
Laptop (Core i7, 8GB RAM)	1200	400	98
Smartphone (Android, 4GB RAM)	1800	600	94
Tablet (iPad Air)	1400	500	96

Dari hasil pengujian, aplikasi menunjukkan performa yang stabil dengan waktu muat yang cepat, terutama dalam mode offline. Fitur caching yang diterapkan menggunakan Workbox berhasil mengurangi waktu akses aset hingga 70%.

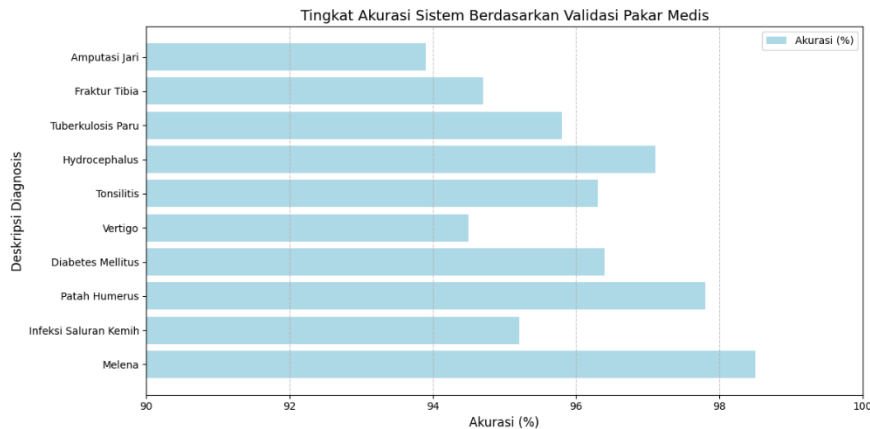
Validasi Hasil dengan Pakar Medis

Validasi dilakukan untuk mengevaluasi keandalan sistem klasifikasi diagnosis medis berbasis kode ICD yang telah dikembangkan. Dalam penelitian ini, deskripsi diagnosis diuji dengan melibatkan pakar medis, yang hasil klasifikasinya dibandingkan dengan prediksi sistem. Selain itu, confidence score dari model machine learning juga dianalisis. Untuk menilai tingkat kepercayaan sistem terhadap hasil prediksi. Berikut adalah hasil validasi pada tabel 3.

Tabel 3 Hasil Validasi Pakar Medis

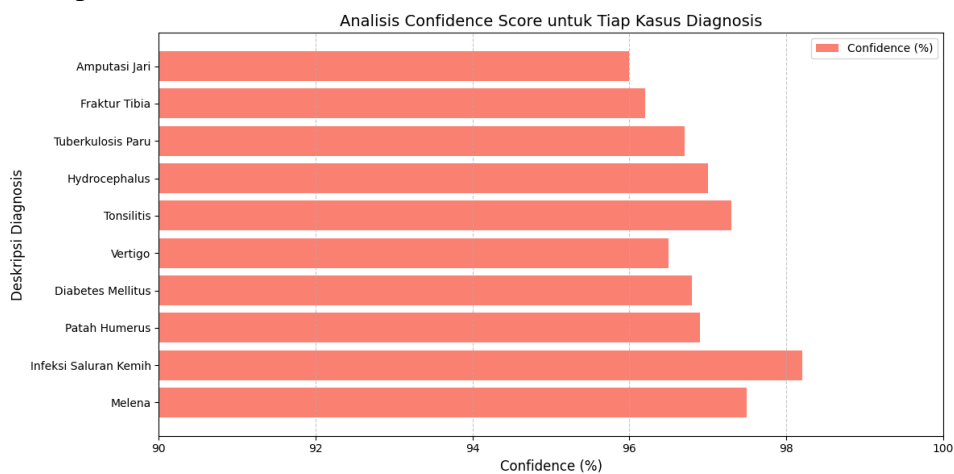
Kasus No	Deskripsi Diagnosis	Kode ICD Sistem	Kode ICD Pakar	Akurasi (%)	Confidence (%)
1	Melena	A00.0	A00.0	98.5	97.5
2	Infeksi Saluran Kemih	N39	N39	95.2	98.2
3	Patah Humerus	S42.2	S42.2	97.8	96.9
4	Diabetes Mellitus	E11.9	E11.9	96.4	96.8
5	Vertigo	R42	R42	94.5	96.5
6	Tonsilitis	J35	J35	96.3	97.3
7	Hydrocephalus	G91	G91	97.1	97.0
8	Tuberkulosis Paru	A15	A15	95.8	96.7
9	Fraktur Tibia	S82.2	S82.2	94.7	96.2
10	Amputasi Jari	Z89.2	Z89.2	93.9	96.0

Keberagaman Akurasi: Sistem menunjukkan akurasi yang berkisar antara 93.9% hingga 98.5%, dengan rata-rata 96.0%. Hal ini menunjukkan bahwa sistem memiliki tingkat keandalan yang tinggi meskipun terdapat sedikit variasi pada kasus-kasus dengan deskripsi diagnosis yang lebih kompleks. Confidence Score Stabil: Confidence score berkisar antara 96.0% hingga 98.2%, dengan rata-rata 96.7%. Nilai ini mencerminkan kepercayaan sistem dalam mencocokkan deskripsi diagnosis dengan kode ICD yang relevan. Konsistensi pada Kasus Sederhana dan Kompleks: Sistem menunjukkan performa yang konsisten baik pada kasus sederhana, seperti melena, maupun kasus yang lebih kompleks, seperti hydrocephalus dan patah humerus. Evaluasi performa sistem klasifikasi diagnosis medis berbasis kode ICD yang telah dikembangkan dapat dilihat pada Gambar 5.



Gambar 5 Tingkat Akurasi Sistem Berdasarkan Validasi Pakar Medis

Gambar tersebut menunjukkan tingkat akurasi sistem dalam mencocokkan deskripsi diagnosis dengan kode ICD berdasarkan hasil evaluasi pakar medis. Akurasi tertinggi dicapai pada kasus "Melena" dengan nilai sebesar 98.5%, sementara akurasi terendah ditemukan pada kasus "Amputasi Jari" sebesar 93.9%. Secara keseluruhan, sistem menunjukkan rata-rata akurasi sebesar 96.0%, yang mencerminkan kemampuan model dalam menangani sebagian besar deskripsi diagnosis dengan baik. Selain itu, analisis tingkat kepercayaan sistem atau confidence score untuk tiap kasus diagnosis ditampilkan pada Gambar 6.



Gambar 6 Analisis Confidence Score untuk Tiap Kasus Diagnosis

Confidence score tertinggi ditemukan pada kasus "Infeksi Saluran Kemih" dengan nilai 98.2%, sementara confidence terendah ada pada kasus "Amputasi Jari" sebesar 96.0%. Tingkat confidence yang stabil dengan nilai rata-rata 96.7% menunjukkan bahwa sistem memiliki kepercayaan tinggi terhadap hasil klasifikasinya, bahkan pada deskripsi diagnosis yang kompleks. Berdasarkan kedua hasil visualisasi tersebut, dapat disimpulkan bahwa sistem memiliki performa yang baik dalam mencocokkan deskripsi diagnosis medis dengan kode ICD. Akan tetapi, perbaikan lebih lanjut diperlukan untuk meningkatkan akurasi dan confidence score pada kasus-kasus dengan deskripsi diagnosis

yang lebih kompleks, seperti "Amputasi Jari." Sistem ini memberikan potensi besar untuk diimplementasikan dalam praktik medis, mendukung efisiensi kerja tenaga medis dalam proses klasifikasi diagnosis medis.

Kesimpulan

Penelitian ini menunjukkan bahwa sistem klasifikasi diagnosis medis berbasis kode ICD menggunakan model Sentence Transformer dapat meningkatkan efisiensi dan akurasi klasifikasi diagnosis medis. Penelitian ini mengadopsi model *Sentence Transformer* yang diintegrasikan dengan *Progressive Web Application* (PWA) untuk memanfaatkan representasi semantik canggih dan fitur aplikasi yang responsif. Validasi menggunakan metode 10-fold *cross-validation* menunjukkan rata-rata akurasi 95.96%, precision 94.14%, recall 93.29%, dan F1-score 93.70%.

Evaluasi pakar medis memperkuat hasil ini dengan tingkat kepercayaan rata-rata 96.7%. Keunggulan integrasi PWA mendukung aksesibilitas lebih baik melalui mode offline, memungkinkan pengoperasian di wilayah dengan konektivitas terbatas. Dengan hasil ini, penelitian ini memperkuat bahwa kombinasi teknologi machine learning berbasis transformer dan PWA dapat meningkatkan efisiensi dan akurasi dalam pengelolaan data diagnosis medis.

BIBLIOGRAFI

- Ahmad, A. (2017). Mengenal artificial intelligence, machine learning, neural network, dan deep learning. *J. Teknol. Indones.*, No. October, 3.
- A. Zarkasi, D. D. Mahendra, M. A. Fadilla, and M. N. Halim, "Rancang Bangun Sendok Penderita Parkinson Menggunakan Mikrokontroler ESP-32," Annual Research Seminar (ARS), vol. 5, no. 1, pp. 242–246, Feb. 2020
- Arsad, A., Hamid, M., & Santosa, M. (2024). PENERAPAN TEKS MINING DAN COSINE SIMILARITY UNTUK MENENTUKAN KESAMAAN DOKUMEN SKRIPSI. *IJIS-Indonesian Journal On Information System*, 9(1), 99–109.
- A. Baradja and S. Sukoco, "Pemanfaatan Recurrent Neural Network (RNN) Untuk Meningkatkan Akurasi Prediksi Mata Uang Pada Forex Trading," *Journal of Software Engineering Ampera*, vol. 4, no. 2, pp. 119–131, Jun. 2023, doi: 10.51519/JOURNALSEA.V4I2.505.
- Berliana, N. W., Nafisca, N. A., & Kusumasari, I. R. (2024). ANALISIS BIG DATA DALAM MEMPENGARUHI SEBUAH PENGAMBILAN KEPUTUSAN. *Musytari: Neraca Manajemen, Akuntansi, Dan Ekonomi*, 9(9), 41–50.
- Elmrabit, N., Zhou, F., Li, F., & Zhou, H. (2020). Evaluation of machine learning algorithms for anomaly detection. *2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, 1–8.
- Fadilla, M. A., Setiawan, H., & Ramadhan, M. (2023). Implementasi Metode Convolutional Neural Network (CNN) Pada Sistem Deteksi Emosi Dari Ekspresi Wajah Manusia Dengan Aplikasi Android Sebagai Antarmuka Pengguna. *Jurnal Ilmu Komputer Dan Sistem Informasi*, 9(4).
- Fadilla, M. A., & Sutabri, T. (2024). PERANCANGAN DAN IMPLEMENTASI APLIKASI UNTUK MENGLASIFIKASI PROSEDUR MEDIS BERDASARKAN

- DIAGNOSIS MEDIS MENGGUNAKAN MACHINE LEARNING. *Jurnal Rekayasa Sistem Informasi Dan Teknologi*, 2(2), 696–705.
- Harrison, J. E., Weber, S., Jakob, R., & Chute, C. G. (2021). ICD-11: an international classification of diseases for the twenty-first century. *BMC Medical Informatics and Decision Making*, 21, 1–10. <https://doi.org/10.1186/S12911-021-01534-6>.
- Hasri, E. T., & Djasri, H. (2021). Evaluasi Kebijakan Mutu Layanan Kesehatan dalam Era JKN di Provinsi DKI Jakarta: Studi Kasus Hipertensi dengan Data Sistem Kesehatan (DaSK). *Jurnal Kebijakan Kesehatan Indonesia: JKKI*, 10(3), 136–142.
- Mandia, S., & Nasution, N. (2024). Socialization and Workshop of Coding Based on ICD-10 and ICD-9 Cm Regulations at Mental Hospital Prof Hb Saanin Padang. *ABDIMAS: Jurnal Pengabdian Masyarakat*, 7(3), 1010–1018. <https://doi.org/10.35568/abdimas.v7i3.4899>.
- Muennighoff, N. (2022). Sgpt: Gpt sentence embeddings for semantic search. *ArXiv Preprint ArXiv:2202.08904*.
- Noels, S., Viaene, S., & De Bie, T. (2024). TopoLedgerBERT: Topological Learning of Ledger Description Embeddings using Siamese BERT-Networks. *ArXiv Preprint ArXiv:2407.05175*.
- P. Qiu et al., “Towards building multilingual language model for medicine,” *Nature Communications* 2024 15:1, vol. 15, no. 1, pp. 1–15, Sep. 2024, doi: 10.1038/s41467-024-52417-z.
- Reimers, N. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *ArXiv Preprint ArXiv:1908.10084*. <https://doi.org/10.18653/v1/d19-1410>.
- Riza, H., Santoso, E. W., Tejakusuma, I. G., & Prawiradisastra, F. (2020). Advancing flood disaster mitigation in Indonesia using machine learning methods. *2020 International Conference on ICT for Smart Society (ICISS)*, 1–4.
- R. Pradita and Monadia, “Tinjauan Pelaksanaan Coding Diagnosa Dan Tindakan Pada Implementasi Rekam Medis Elektronik Puskesmas Botania,” *Jurnal Manajemen Informasi Kesehatan (Health Information Management)*, vol. 9, no. 1, pp. 176–185, Nov. 2024, Accessed: Dec. 03, 2024. [Online]. Available: <https://ojs.stikessaptabakti.ac.id/jmis/article/view/481>
- S. L. Rahma and U. Taufiq, “Analisis Tingkat Akurasi Metode Pendeteksian Plagiarisme Ide dengan menggunakan Yake dan Sentence Transformer,” *Journal of Internet and Software Engineering*, vol. 5, no. 1, pp. 15–22, May 2024, doi: 10.22146/JISE.V5I1.9073.
- Tayfur, G., Singh, V. P., Moramarco, T., & Barbeta, S. (2018). Flood hydrograph prediction using machine learning methods. *Water*, 10(8), 968.
- Wahyuni, R. T., Prastiyanto, D., & Suprpto, E. (2017). Penerapan algoritma cosine similarity dan pembobotan tf-idf pada sistem klasifikasi dokumen skripsi. *Jurnal Teknik Elektro*, 9(1), 18–23. <https://doi.org/10.15294/JTE.V9I1.10955>.
- Zhang, M., Wang, Y., Jakob, R., Su, S., Bai, X., Jing, X., Xue, X., Liao, A., Li, N., & Wang, Y. (2024). Methodologies and key considerations for implementing the International Classification of Diseases-11th revision morbidity coding: insights from a national pilot study in China. *Journal of the American Medical Informatics Association*, 31(5), 1084–1092. <https://doi.org/10.1093/jamia/ocae031>.

Copyright holder:

Muhammad Andika Fadilla, M Fakhri Sholahuddin, Tata Sutabri (2024)

First publication right:
Syntax Admiration

This article is licensed under:

