

<h3>Background</h3> <p>Stroke merupakan penyakit dengan tingkat kematian dan kecacatan yang tinggi. Masalah utama di dunia medis adalah keterlambatan deteksi dini, di mana pasien berisiko stroke sering diklasifikasikan sebagai sehat. Selain itu, model machine learning sering hanya berhenti di tahap eksperimen dan tidak siap digunakan dalam sistem nyata. Oleh karena itu, dibutuhkan sistem prediksi risiko stroke yang sensitif, konsisten, dan siap produksi melalui pendekatan MLOps.</p>	<h3>Solution</h3> <p>Solusi berupa aplikasi web prediksi risiko stroke berbasis Streamlit.</p> <p>Fitur utama:</p> <ul style="list-style-type: none">Input data pasien (profil, medis, gaya hidup)Pilihan model (Random Forest, SVM, Logistic Regression)Output probabilitas risiko stroke dengan visual peringatanRisk calibration berbasis pengetahuan medis <p>Out-of-scope: diagnosis klinis final (sistem hanya untuk skrining awal).</p>	<h3>Data</h3> <ol style="list-style-type: none">Sumber data: healthcare-dataset-stroke-data.csvData training dan inference berasal dari dataset yang samaLabel: stroke (0 = tidak stroke, 1 = stroke)Preprocessing:<ul style="list-style-type: none">Imputasi nilai BMIEncoding data kategorikalStandardisasi data numerikMasalah utama: class imbalance	<h3>Modeling</h3> <p>Tahap modeling dilakukan dengan pendekatan komparatif, di mana beberapa algoritma machine learning dilatih dan dievaluasi menggunakan pipeline preprocessing yang sama. Model yang digunakan meliputi Logistic Regression, Support Vector Machine (SVM), dan Random Forest. Evaluasi difokuskan pada metrik recall untuk memastikan kemampuan model dalam mendeteksi pasien berisiko stroke dan meminimalkan kesalahan false negative. Model dengan performa terbaik dipilih sebagai model utama, sementara model lain tetap disediakan sebagai pembanding.</p>	<h3>Feedback</h3> <p>Feedback diperoleh dari hasil prediksi sistem, perbandingan output antar model, serta analisis kasus dengan risiko tinggi. Feedback ini digunakan sebagai dasar untuk melakukan iterasi model, penyesuaian threshold risiko, dan peningkatan performa sistem di tahap selanjutnya.</p>
<h3>Value proposition</h3> <p>Sistem ini memberikan peringatan dini risiko stroke berbasis AI dengan sensitivitas tinggi. Nilai utama sistem:</p> <ol style="list-style-type: none">Mengurangi risiko false negative (pasien stroke dikatakan sehat)Mendukung keputusan medis sebagai second opinionModel dapat dijalankan secara konsisten di berbagai lingkungan (reproducible)Mudah digunakan melalui antarmuka web interaktif	<h3>Feasibility</h3> <p>Proyek feasible karena:</p> <ul style="list-style-type: none">Dataset publik tersediaMenggunakan library open-source (Scikit-learn, Streamlit)Infrastruktur sederhana (Docker, GitHub)Model dapat dijalankan di laptop lokal tanpa resource besar	<h3>Metrics</h3> <p>Metrik utama: Recall (Sensitivitas)</p> <p>Confusion Matrix digunakan sebagai alat evaluasi untuk menganalisis distribusi True Positive, False Negative, False Positive, dan True Negative. Recall diprioritaskan karena kesalahan false negative sangat berbahaya dalam konteks medis.</p>	<h3>Inferences</h3> <p>Inference dilakukan secara online (real-time) melalui aplikasi web berbasis Streamlit. Pengguna memasukkan data pasien, kemudian sistem melakukan preprocessing data, memanggil model terlatih, dan menampilkan probabilitas risiko stroke secara langsung tanpa perlu proses pelatihan ulang.k</p>	<h3>Project</h3> <p>Tim: Kelompok 9</p> <ol style="list-style-type: none">122450058 Adil Aulia Rahma Nurhidayah122450059 Nathanael Daniel Santoso122450061 Kharisa Harvany122450066 Cintya Bella122450102 Daris Samudra <p>Deliverables:</p> <ul style="list-style-type: none">Source code (GitHub)Model .pklAplikasi StreamlitDockerfileDokumentasi README <p>Timeline:</p> <ul style="list-style-type: none">Data & preprocessingModeling & evaluasiDeploymentFinal report & presentasi
<h3>Objectives</h3> <ol style="list-style-type: none">Membangun sistem prediksi risiko stroke berbasis machine learningMengimplementasikan workflow MLOps end-to-endMenangani permasalahan imbalanced data pada kasus medisMenyediakan online inference melalui aplikasi webMenjamin reproducibility dengan containerization (Docker)		<h3>Evaluation</h3> <p>Offline evaluation:</p> <ul style="list-style-type: none">Confusion MatrixRecall Score <p>Online evaluation:</p> <ul style="list-style-type: none">Observasi hasil prediksi di aplikasiAnalisis perbedaan output antar modelValidasi hasil dengan logika medis (risk calibration)		