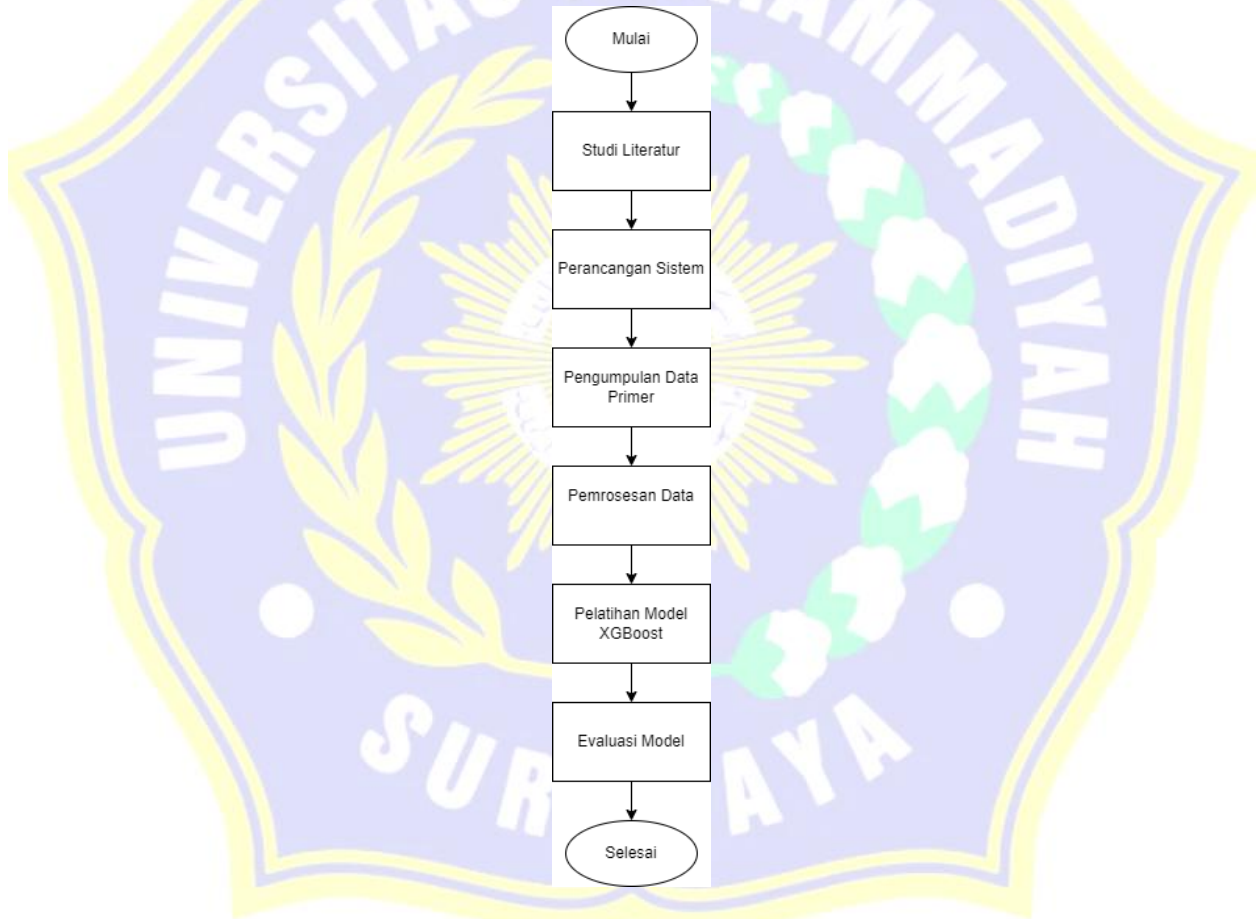


BAB III

METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan eksperimental dan komputasional. Pendekatan komputasional digunakan dalam proses pengolahan data dan pelatihan model machine learning menggunakan algoritma XGBoost. Alur penelitian yang dilakukan pada penelitian ini ditunjukkan pada Gambar 3.1. Flowchart penelitian menggambarkan tahapan penelitian secara keseluruhan mulai dari identifikasi permasalahan, pengumpulan data, perancangan sistem, implementasi, hingga proses pengujian dan evaluasi hasil sistem deteksi risiko diabetes.



Gambar 3. 1 Flowchart Penelitian

Berdasarkan Gambar 3.1, penelitian diawali dengan identifikasi permasalahan mengenai perlunya metode deteksi risiko diabetes yang bersifat noninvasif. Tahap

berikutnya adalah studi literatur dan pengumpulan data sebagai dasar dalam perancangan sistem. Selanjutnya dilakukan perancangan perangkat keras menggunakan ESP32, sensor GSR, dan sensor MAX30102 serta perancangan perangkat lunak berbasis website yang terintegrasi dengan algoritma XGBoost. Setelah sistem selesai diimplementasikan, dilakukan proses pengujian menggunakan data primer untuk memperoleh hasil klasifikasi risiko diabetes. Tahap terakhir adalah evaluasi performa sistem menggunakan confusion matrix yang menghasilkan nilai accuracy, precision, dan recall sebagai dasar dalam menilai kinerja sistem.

3.1 Pengumpulan Data

Pengumpulan data bertujuan untuk memperoleh informasi fisiologis pasien yang relevan dan faktual guna melatih serta menguji sistem deteksi risiko diabetes noninvasif. Untuk menjaga validitas penelitian dan memastikan model memiliki dasar pembelajaran yang nyata sebelum diimplementasikan, penelitian ini secara mutlak hanya menggunakan 100% data primer lapangan. Tidak ada penggunaan data sekunder dari internet, sehingga seluruh tahapan pelatihan maupun pengujian model algoritma Extreme Gradient Boosting (XGBoost) murni bertumpu pada data yang diakuisisi langsung dari pasien.

Data primer dipilih untuk penelitian ini karena mewakili kondisi fisiologis aktual para responden di lingkungan implementasi sistem. Data primer dikumpulkan secara langsung melalui pengukuran menggunakan sensor Galvanic Skin Response (GSR) dan sensor MAX30102, serta divalidasi menggunakan pembacaan glukometer sebagai acuan kebenaran. Hal ini dilakukan untuk mengurangi potensi bias yang mungkin timbul akibat penggunaan data sekunder yang bersumber dari sumber, populasi, atau prosedur pengukuran yang berbeda. Oleh karena itu, data yang digunakan saat melatih dan menguji model memiliki karakteristik yang sesuai dengan tujuan penelitian sehingga hasil klasifikasi menjadi lebih andal

Pengumpulan data dilakukan dalam beberapa tahap, yang meliputi persiapan perangkat keras, pengujian fungsi masing-masing sensor, dan pengumpulan data dari setiap peserta. Sebelum memulai pengukuran, fungsionalitas semua perangkat diperiksa, termasuk mikrokontroler ESP32, sensor GSR, sensor MAX30102, dan

koneksi jaringan yang digunakan untuk transfer data ke situs web. Langkah persiapan ini dilakukan untuk meminimalkan kesalahan pembacaan sensor yang dapat memengaruhi kualitas data yang diperoleh.

Para peserta diminta untuk meletakkan jari mereka pada sensor MAX30102 dan memegang elektroda sensor GSR pada posisi yang ditentukan selama pengumpulan data. Sistem merekam detak jantung, saturasi oksigen (SpO₂), dan respons kulit galvanik (GSR) para peserta ketika sinyal fisiologis telah stabil. Operator memasukkan usia dan Indeks Massa Tubuh (BMI) responden ke dalam sistem bersama dengan data yang dikumpulkan secara otomatis oleh sensor dari hasil pengukuran. Semua parameter kemudian dikirim ke server melalui jaringan Wi-Fi dan disimpan dalam basis data sebagai data penelitian.

Setiap hasil pengukuran dibandingkan dengan hasil dari glukometer, untuk menjaga kualitas data sebagai acuan (ground truth). Proses validasi dilakukan untuk memastikan bahwa label kelas yang digunakan dalam algoritma Extreme Gradient Boosting (XGBoost) sesuai dengan kondisi peserta. Data yang telah divalidasi kemudian diklasifikasikan ke dalam dua kelas, yaitu Normal (0) dan Berisiko Diabetes (1), yang dapat digunakan sebagai data pelatihan dan pengujian untuk model klasifikasi (Babajide Mustapha & Saeed, 2016).

Prosedur pengumpulan data dalam penelitian ini dilakukan secara sistematis, sehingga menghasilkan data fisiologis yang konsisten dan relevan secara langsung dengan implementasi sistem yang dikembangkan. Pengukuran dilakukan dengan cara yang sama untuk setiap subjek pada semua tahap proses pengumpulan data, sehingga variasi hasil pengukuran lebih mewakili kondisi fisiologis masing-masing individu dan bukan disebabkan oleh perbedaan perlakuan selama proses pengumpulan data. Oleh karena itu, data yang diperoleh diharapkan dapat menjadi landasan yang kuat untuk pelatihan, pengujian, dan evaluasi kinerja model XGBoost dalam sistem deteksi risiko diabetes noninvasif.

3.1.1 Variabel Penelitian

Populasi dalam penelitian ini meliputi pasien yang melakukan pemeriksaan kesehatan di fasilitas kesehatan tingkat pertama (klinik). Sampel yang digunakan

ditetapkan sebanyak 100 rekaman data primer. Pengambilan data ini terbagi secara independen menjadi dua kelompok, yaitu 100 sampel pasien digunakan secara khusus untuk pelatihan model komputasi, dan 30 sampel pasien berbeda digunakan untuk validasi atau pengujian sistem secara langsung di lapangan (Sedighi et al., 2026). Variabel penelitian dibatasi secara ketat menjadi dua jenis, yakni variabel independen dan variabel dependen. Variabel independen (fitur input) terdiri dari 5 parameter masukan sistem, yaitu Usia (tahun), Indeks Massa Tubuh/BMI, Heart Rate (BPM), Saturasi Oksigen (SpO₂), dan Resistansi Kulit (GSR). Penggunaan sensor wearable noninvasif ini sangat relevan untuk memantau variabilitas fisiologis penderita diabetes secara aman. Sementara itu, variabel dependen (target/label) merupakan hasil klasifikasi risiko diabetes yang bertumpu pada hasil pengukuran glukometer invasif sebagai kebenaran mutlak atau ground truth. Variabel target ini dikategorikan menjadi "Label 0" untuk merepresentasikan kondisi glukosa normal dan "Label 1" untuk merepresentasikan kondisi berisiko (Risma Ayu Fitriyani et al., 2023).

Dalam penelitian ini, variabel-variabel independen dipilih berdasarkan kemampuan masing-masing parameter dalam menggambarkan kondisi fisiologis yang berkaitan dengan risiko diabetes mellitus. Alasan pemilihan kelima variabel ini adalah karena variabel-variabel tersebut dapat diperoleh secara noninvasif atau melalui pengukuran sederhana tanpa pengambilan sampel darah. Selain itu, kombinasi usia, Indeks Massa Tubuh (IMT), denyut jantung, saturasi oksigen (SpO₂), dan Respons Kulit Galvanik (GSR) diharapkan dapat memberikan gambaran yang lebih lengkap mengenai kondisi fisiologis responden dibandingkan dengan penggunaan satu parameter saja. Pemilihan variabel-variabel ini juga didasarkan pada kemudahan implementasinya pada sistem berbasis Internet of Things (IoT) yang dikembangkan dalam penelitian ini (Risma Ayu Fitriyani et al., 2023).

Selama proses klasifikasi, variabel-variabel independen yang berbeda memiliki karakteristik yang berbeda pula. Variabel usia dan BMI diperoleh melalui input manual oleh operator berdasarkan data responden. Nilai detak jantung, SpO₂, dan GSR diperoleh secara otomatis melalui pembacaan dari sensor MAX30102 dan sensor respons kulit galvanik (GSR). Variabel-variabel ini kemudian digabungkan menjadi

satu vektor fitur sebagai masukan untuk algoritma Extreme Gradient Boosting (XGBoost). Model ini menggunakan beberapa parameter fisiologis secara bersamaan sehingga hubungan antarvariabel dapat dipelajari oleh model dan menghasilkan klasifikasi risiko diabetes yang lebih akurat dibandingkan dengan penggunaan satu variabel secara terpisah (Babajide Mustapha & Saeed, 2016).

Variabel dependen dalam penelitian ini adalah hasil klasifikasi risiko diabetes berdasarkan kebenaran dasar (ground truth) dari pembacaan glukometer. Penggunaan hasil glukometer sebagai acuan bertujuan untuk memastikan bahwa seluruh data yang digunakan dalam proses pelatihan dan pengujian dilabeli sesuai dengan kondisi aktual responden. Selain itu, label-label ini digunakan sebagai target pembelajaran dalam algoritma XGBoost sehingga model dapat mempelajari pola hubungan antara variabel independen dan kelas keluaran secara optimal. Hasil evaluasi sistem diharapkan menjadi lebih valid dengan proses pelabelan yang mengacu pada perangkat medis standar (Sedighi et al., 2026).

3.2 Pengolahan Data

Prosedur pengolahan data primer ini difokuskan untuk mengintegrasikan secara harmonis lima parameter masukan noninvasif yang meliputi karakteristik usia, indeks massa tubuh (BMI), frekuensi detak jantung (Heart Rate), saturasi oksigen darah (SpO2), serta nilai konduktivitas elektrik fisis kulit (Galvanic Skin Response/GSR). Tahapan komparatif dilakukan secara deskriptif komputasional dengan mengolah sinyal mentah yang dikirim oleh mikrokontroler ESP32 melalui proses pembersihan noise dan penyesuaian format fisis. Algoritma Extreme Gradient Boosting (XGBoost) berjalan secara stabil. Setiap paket dari kelima fitur input tersebut kemudian disandingkan dengan label hasil glukometer invasif (ground truth) yang diinput melalui website (Label 0 untuk Normal dan Label 1 untuk Berisiko). Proses penyandingan ini bertujuan untuk memastikan bahwa model prediktif yang dikembangkan memiliki landasan pembelajaran klinis yang kokoh serta mampu meminimalkan nilai kesalahan logaritmik (log loss) saat diuji menggunakan 30 data pasien validasi baru.

Semua data yang diperoleh dari sensor terlebih dahulu diproses secara awal untuk menjamin kualitas data yang akan digunakan dalam proses klasifikasi sebelum

dimanfaatkan dalam proses pelatihan model. Tahap ini bertujuan untuk mengurangi kemungkinan terjadinya kesalahan prediksi yang disebabkan oleh data yang tidak lengkap, tidak konsisten, atau berisik. Pemrosesan awal merupakan langkah penting dalam penelitian ini karena kualitas data masukan akan memengaruhi kemampuan algoritma XGBoost dalam membangun model klasifikasi yang optimal (Chen & Guestrin, 2016).

Langkah pertama dalam pemrosesan data adalah validasi data. Pada tahap ini, setiap parameter sensor, yaitu usia, Indeks Massa Tubuh (BMI), detak jantung, saturasi oksigen (SpO₂), dan respons kulit galvanik (GSR), diperiksa. Data yang tidak lengkap atau data yang dihasilkan dari kegagalan pembacaan sensor tidak digunakan dalam proses pelatihan model. Hal ini dilakukan untuk memastikan bahwa semua titik data yang digunakan dalam analisis mengandung kelima variabel masukan tersebut, sehingga model menerima informasi yang konsisten untuk semua sampel penelitian.

Langkah selanjutnya adalah proses pembersihan data untuk mengurangi pengaruh gangguan sinyal yang terjadi selama pengumpulan data. Gangguan ini dapat disebabkan oleh gerakan responden, perubahan posisi sensor, atau fluktuasi sesaat pada sinyal fisiologis. Oleh karena itu, pembacaan sensor disaring terlebih dahulu sebelum data dikirim ke server. Langkah ini bertujuan untuk menghasilkan data yang lebih stabil guna meningkatkan kualitas proses pelatihan algoritma XGBoost.

Format data kemudian dikonversi ke dalam struktur yang seragam untuk semua parameter sebelum disimpan ke dalam basis data. Data fisiologis yang dikumpulkan oleh ESP32 dikirimkan ke server situs web melalui protokol HTTP menggunakan JavaScript Object Notation (JSON). Saat server menerima data tersebut, data dikonversi ke dalam format tabel di mana setiap responden merupakan satu baris dan setiap variabel penelitian merupakan satu kolom. Struktur data tersebut kemudian dapat diproses lebih lanjut oleh algoritma pembelajaran mesin.

Terdapat langkah normalisasi untuk menyamakan rentang nilai di seluruh variabel, sehingga perbedaan skalanya tidak terlalu besar. Variabel seperti usia, BMI, detak jantung, SpO₂, dan GSR memiliki satuan pengukuran yang berbeda-beda dan oleh karena itu perlu distandardisasi sebelum digunakan dalam proses pelatihan model.

Rentang data yang lebih seragam dapat membuat proses pembangunan pohon keputusan algoritma XGBoost menjadi lebih stabil dan efisien, sehingga memungkinkan model untuk mempelajari hubungan antarvariabel dengan lebih baik (Chen & Guestrin, 2016).

Setelah semua langkah prapemrosesan selesai, dataset tersebut dibagi menjadi data pelatihan dan data pengujian sesuai dengan desain penelitian. Data pelatihan digunakan untuk membangun model klasifikasi menggunakan algoritma Extreme Gradient Boosting (XGBoost), sedangkan data pengujian digunakan untuk menguji kemampuan model dalam memprediksi data yang belum pernah digunakan dalam pelatihan. Prediksi model tersebut kemudian dibandingkan dengan hasil pengukuran glukosa yang diperoleh dari alat pengukur glukosa, yang digunakan sebagai acuan (ground truth). Perbandingan ini digunakan untuk menghitung akurasi, presisi, recall, dan matriks kebingungan. Indikator-indikator ini mencerminkan kinerja sistem.

3.2.1 Pengolahan Data Primer

Data fisis yang diperoleh secara langsung melalui sensor bio-optik MAX30102 dan sensor Galvanic Skin Response (GSR) yang terintegrasi dengan mikrokontroler ESP32 pada awalnya masih berupa data mentah (raw data). Oleh karena itu, data tersebut harus melalui beberapa langkah prapemrosesan sebelum dapat dieksekusi oleh algoritma, khususnya meliputi tahapan berikut:

1. Pembersihan Data (Data Cleaning): Meliputi pengurangan noise atau gangguan sinyal pada bacaan sensor fisik (MAX30102 dan GSR) saat pengambilan sampel di klinik, serta mengeliminasi rekaman data yang tidak akurat akibat pergerakan jari pasien.
2. Ekstraksi dan Penyelesaian Fitur: Menghitung nilai statistik yang mencerminkan sifat sinyal fisis (seperti rata-rata nilai konduktivitas kulit dari sensor GSR), serta menghitung angka Indeks Massa Tubuh (BMI) pasien berdasarkan variabel berat dan tinggi badan yang dikombinasikan dengan karakteristik usia.
3. Normalisasi Data: Proses menyeimbangkan skala numerik dari 5 fitur masukan (usia, BMI, Heart Rate, SpO2, dan nilai resistansi kulit) mengingat adanya

perbedaan format satuan fisis (BPM, persentase, dan Volt/Ohm), sehingga proses pembelajaran model oleh algoritma menjadi lebih stabil.

4. Pembagian Data (Data Splitting): Memisahkan total 130 kumpulan data primer klinis ke dalam dua kelompok independen tanpa pengacakan berulang, yakni 100 data dialokasikan murni sebagai data latih (training data) untuk membangun model algoritma, dan 30 data pasien baru dialokasikan sebagai data uji (testing data) untuk tahap implementasi inferensi. Pemisahan ini menjamin bahwa model diuji pada data responden yang belum pernah dipelajari sebelumnya.

3.2.2 Pemodelan Algoritma XGBoost

Algoritma Extreme Gradient Boosting (XGBoost) digunakan dalam penelitian ini sebagai metode klasifikasi untuk membangun model prediksi risiko diabetes. XGBoost merupakan algoritma machine learning berbasis ensemble learning dengan pendekatan boosting, di mana beberapa model sederhana (weak learners) digabungkan secara bertahap untuk membentuk model yang lebih kuat (strong learner). Dalam implementasinya, XGBoost menggunakan decision tree (Classification and Regression Trees/CART) sebagai base learner. Pendekatan boosting bekerja secara berurutan, di mana setiap model baru dibangun untuk memperbaiki kesalahan dari model sebelumnya. Proses kerja XGBoost dalam melatih kumpulan data primer pada penelitian ini dilakukan melalui tahapan sebagai berikut:

1. Inisialisasi Prediksi Awal : Model dimulai dengan menghasilkan prediksi awal terhadap data training yang berasal dari pecahan dataset primer. Pada klasifikasi biner (Normal dan Berisiko), prediksi awal berupa probabilitas dasar berdasarkan distribusi kelas di dalam dataset primer tersebut.
2. Perhitungan Fungsi Loss : Setelah prediksi awal diperoleh, dihitung nilai kesalahan menggunakan fungsi *loss* yang sesuai untuk klasifikasi biner, yaitu *log loss*. Nilai *loss* menunjukkan selisih antara prediksi model dan label aktual (*ground truth*) dari glukometer, dengan rumus:

$$L = - \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad [3]$$

Di mana :

y_i = label aktual (0 atau 1)

$\hat{y}_i$ = probabilitas prediksi

3. Perhitungan Gradient dan Hessian : XGBoost menghitung turunan pertama (gradient) dan turunan kedua (Hessian) dari fungsi loss terhadap prediksi. Informasi ini digunakan untuk menentukan arah dan besarnya perbaikan model pada iterasi berikutnya.

Gradient :

$$g_i = \hat{y}_i - y_i \quad [4]$$

Hessian :

$$h_i = \hat{y}_i(1 - \hat{y}_i) \quad [5]$$

Gradient digunakan untuk menentukan arah perbaikan, sedangkan hessian menentukan besar langkah perbaikan pada iterasi berikutnya.

4. Pembangunan Decision Tree Baru : Berdasarkan nilai gradient dan hessian, dibangun decision tree baru yang berfungsi untuk memperbaiki kesalahan model sebelumnya. Tree ini tidak berdiri sendiri, melainkan menjadi bagian dari model yang diperbarui.
5. Penambahan Model Secara Bertahap (Boosting) : Tree yang telah dibangun ditambahkan ke model sebelumnya dengan bobot tertentu (learning rate). Proses ini

dilakukan secara berulang hingga jumlah estimator tercapai atau hingga penurunan nilai loss tidak lagi signifikan.

$$F_t(x) = F_{t-1}(x) + \eta f_t(x) \quad [6]$$

$F_t(x)$ = model pada iterasi ke- t

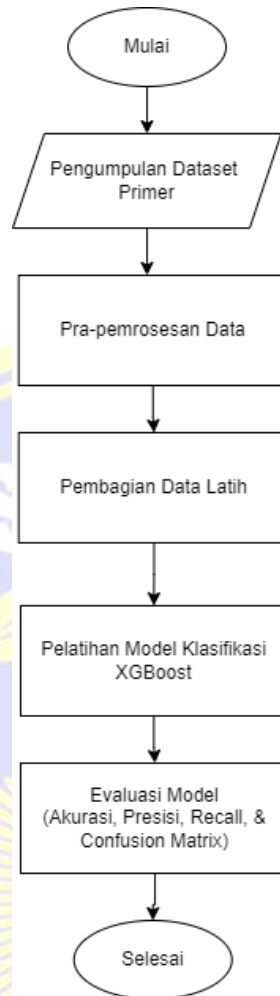
η = learning rate

$f_t(x)$ = tree baru

6. Regularisasi Model: Untuk mencegah overfitting, XGBoost menerapkan mekanisme regularisasi dengan mengontrol kompleksitas tree, seperti pembatasan kedalaman pohon (`max_depth`) dan parameter penalti lainnya. Regularisasi ini meningkatkan kemampuan generalisasi model terhadap data baru.

Setelah model selesai dilatih menggunakan data latih primer, model digunakan untuk melakukan proses inferensi pada data uji primer. Pada tahap implementasi sistem, data masukan berupa usia, indeks massa tubuh (BMI), Heart Rate (BPM), SpO2, serta nilai resistansi kulit (GSR) dimasukkan ke dalam model untuk menghasilkan probabilitas risiko diabetes yang kemudian diklasifikasikan ke dalam dua kelas, yaitu Berisiko (1) dan Normal (0) yang divalidasi dengan alat medis glukometer.

Alur kerja algoritma XGBoost yang digunakan dalam penelitian ini ditunjukkan pada Gambar 3.2. Flowchart tersebut menggambarkan tahapan pembentukan model mulai dari proses inisialisasi prediksi awal, perhitungan fungsi loss, pembentukan decision tree secara bertahap, hingga menghasilkan model klasifikasi yang digunakan untuk memprediksi risiko diabetes.

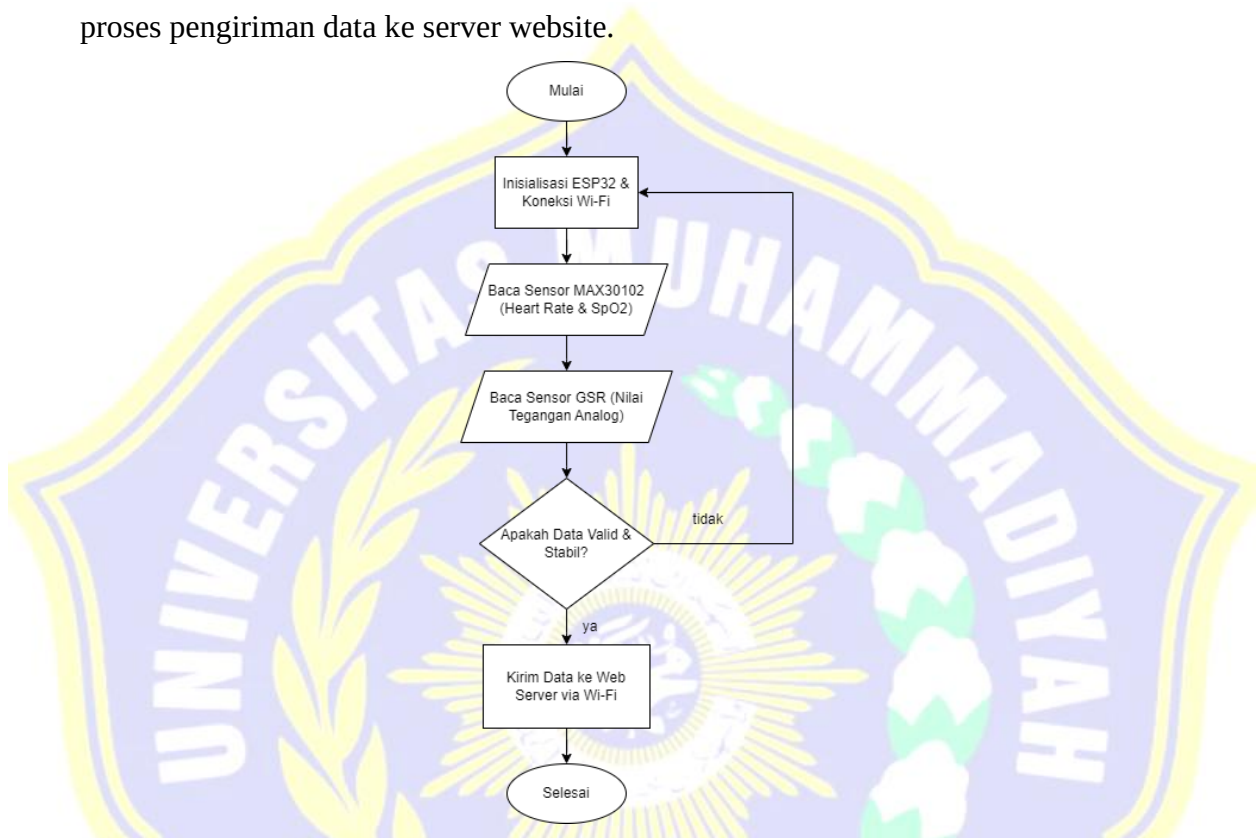


Gambar 3. 2 Flowchart Cara Kerja Algoritma

3.3 Rancangan dan Desain yang diusulkan

Perancangan sistem ini mencakup integrasi menyeluruh (end-to-end) antara perangkat keras pengakuisisi data fisiologis dan perangkat lunak berbasis platform website interaktif. Desain yang diusulkan bertujuan untuk mentransformasikan data pembacaan sensor noninvasif menjadi informasi diagnostik risiko diabetes secara waktu nyata (real-time). Alur data dirancang secara sistematis dimulai dari interaksi fisik jari pasien dengan sensor, pemrosesan sinyal mentah oleh mikrokontroler, transmisi data secara nirkabel, hingga eksekusi model algoritma pembelajaran mesin untuk menampilkan visualisasi tingkat risiko pada antarmuka pengguna di klinik.

Rancangan sistem secara keseluruhan diawali dengan proses akuisisi data menggunakan sensor GSR dan MAX30102 yang dikendalikan oleh mikrokontroler ESP32. Alur kerja perangkat keras tersebut ditunjukkan pada Gambar 3.3, yang menggambarkan tahapan pembacaan data sensor, pengolahan awal oleh ESP32, hingga proses pengiriman data ke server website.



Gambar 3. 3 Flowchart Perangkat Keras

Berdasarkan Gambar 3.3, proses diawali dengan inisialisasi mikrokontroler ESP32, sensor GSR, sensor MAX30102, dan koneksi Wi-Fi. Setelah seluruh komponen aktif, ESP32 membaca data fisiologis berupa nilai GSR, Heart Rate, dan SpO₂. Data tersebut kemudian diproses dan dikirimkan ke server menggunakan format JSON melalui protokol HTTP untuk selanjutnya digunakan pada proses klasifikasi menggunakan algoritma XGBoost.

Desain sistem yang diusulkan menerapkan konsep sistem end-to-end. Artinya, seluruh proses mulai dari pengumpulan data, transmisi, dan pemrosesan hingga penyajian hasil dilakukan dalam satu alur kerja tunggal yang sepenuhnya terintegrasi.

Keputusan ini diambil untuk membatasi transfer data secara manual yang dapat menyebabkan kesalahan pencatatan atau keterlambatan dalam penyampaian hasil inspeksi. Dengan koneksi ini, server dapat menerima pembacaan sensor untuk diproses menggunakan model pembelajaran mesin tanpa memerlukan intervensi tambahan dari operator.

Sensor MAX30102 mengukur detak jantung dan saturasi oksigen (SpO_2) selama fase pengumpulan data, sedangkan sensor Galvanic Skin Response (GSR) mengukur perubahan konduktivitas listrik kulit responden. Semua data fisiologis ini dibaca secara berkala oleh mikrokontroler ESP32, kemudian dilakukan pemeriksaan awal untuk memastikan data yang diperoleh berada dalam rentang pembacaan yang tepat. Setelah semua parameter berhasil diperoleh, data tersebut digabungkan menjadi satu paket data dan dikirim ke server situs web melalui jaringan Wi-Fi.

Perangkat keras tersebut berkomunikasi dengan server menggunakan Hypertext Transfer Protocol (HTTP) dan format pertukaran data JavaScript Object Notation (JSON). JSON dipilih karena strukturnya yang sederhana dan ringan sehingga dapat diproses dengan mudah oleh sistem backend. Saat server menerima data, sistem memvalidasi semua parameter sebelum menyimpannya ke dalam basis data. Data yang dikumpulkan kemudian dimasukkan ke dalam model Extreme Gradient Boosting (XGBoost) untuk memprediksi risiko diabetes berdasarkan pola yang dipelajari selama pelatihan.

Selanjutnya, model tersebut mengembalikan hasil klasifikasi ke antarmuka situs web, dan antarmuka situs web akan menampilkan hasil tersebut dengan cara yang ramah pengguna. Selain itu, situs web ini juga menampilkan nilai-nilai setiap parameter fisiologis yang diperoleh selama proses pengukuran, selain hasil klasifikasi, sehingga operator dapat memverifikasi data yang digunakan dalam proses prediksi. Dengan menggunakan desain sistem terintegrasi ini, proses skrining risiko diabetes dapat dilakukan secara otomatis, cepat, dan secara real-time. Diharapkan sistem ini dapat meningkatkan efisiensi layanan kesehatan pada tahap pemeriksaan awal.

Perancangan perangkat keras difokuskan pada interkoneksi fisis komponen biosensor noninvasif dengan unit pemroses utama mikrokontroler ESP32. Sensor bio-optik MAX30102 dihubungkan ke ESP32 menggunakan protokol komunikasi Inter-Integrated Circuit (I2C), di mana pin SDA (Serial Data) sensor dihubungkan ke pin GPIO 21 ESP32 dan pin SCL (Serial Clock) sensor dihubungkan ke pin GPIO 22 ESP32. Jalur komunikasi I2C ini digunakan untuk mentransmisikan data mentah kardiovaskular berupa frekuensi detak jantung (Heart Rate) dan kadar saturasi oksigen darah (SpO2) secara digital. Sementara itu, sensor analog GSR dihubungkan ke salah satu pin Analog-to-Digital Converter (ADC) pada ESP32, yaitu pin GPIO 34, untuk membaca fluktuasi tegangan listrik mikro pada epidermis kulit akibat aktivitas kelenjar keringat pasien. Untuk menjaga stabilitas tegangan referensi dan menghindari terjadinya galat pembacaan sensor fisis akibat gangguan arus listrik, seluruh rangkaian perangkat keras ini disuplai oleh catu daya eksternal 5 V yang stabil melalui regulator tegangan. Rancangan hubungan antar komponen perangkat keras pada sistem ditunjukkan pada Gambar 3.4. Wiring diagram ini memperlihatkan koneksi antara sensor GSR, sensor MAX30102, dan mikrokontroler ESP32 yang digunakan sebagai pusat pengendali sistem.



Berdasarkan Gambar 3.4, sensor MAX30102 dihubungkan ke ESP32 melalui komunikasi I2C menggunakan pin SDA dan SCL, sedangkan sensor GSR dihubungkan ke pin ADC ESP32 untuk membaca perubahan tegangan analog dari kulit pengguna. Konfigurasi tersebut memungkinkan seluruh data fisiologis diperoleh secara real-time sebelum diproses lebih lanjut oleh sistem.

Desain perangkat keras dalam penelitian ini dirancang dengan fokus pada kemudahan integrasi antarperangkat, stabilitas proses akuisisi data, dan komunikasi yang efisien antar komponen. Kami memilih ESP32 sebagai pusat kendali, karena perangkat ini mampu mengendalikan komunikasi digital melalui protokol I2C sekaligus membaca sinyal analog melalui Analog-to-Digital Converter (ADC). Kemampuan ini memungkinkan semua sensor dioperasikan secara bersamaan tanpa memerlukan mikrokontroler tambahan, sehingga rangkaian menjadi lebih sederhana dan mudah diimplementasikan.

Selama pengukuran, sensor MAX30102 digunakan untuk memperoleh parameter detak jantung dan saturasi oksigen (SpO_2) melalui metode Fotopletismografi (PPG). Sensor Galvanic Skin Response (GSR) digunakan untuk mengukur perubahan konduktivitas listrik kulit yang disebabkan oleh aktivitas kelenjar keringat. Penggunaan kedua sensor secara bersamaan memungkinkan perolehan beberapa parameter fisiologis dalam satu proses pengukuran. Data yang diterima kemudian diproses oleh ESP32, termasuk pembacaan data, penyaringan awal untuk menghilangkan noise dari sinyal, dan pemformatan data sebelum dikirim ke server.

Desain perangkat keras ini juga menekankan stabilitas proses pengukuran, selain memfasilitasi komunikasi antarperangkat. Bus komunikasi I2C dan bus ADC dirutekan secara terpisah untuk mencegah interferensi antara pembacaan data digital dari MAX30102 dan pembacaan sinyal analog dari sensor GSR. Untuk menstabilkan tegangan operasi setiap komponen, diperlukan catu daya yang stabil guna memastikan konsistensi pembacaan sensor selama proses akuisisi data. Dengan desain seperti ini, sistem diharapkan dapat menghasilkan data fisiologis yang lebih stabil dan mengurangi kemungkinan kesalahan pembacaan yang disebabkan oleh gangguan perangkat keras.

Semua elemen desain perangkat keras diintegrasikan menjadi satu kesatuan. Sensor berfungsi sebagai perangkat untuk pengumpulan data fisiologis, ESP32 berperan sebagai pusat kendali yang melakukan pemrosesan data awal, sedangkan situs web berfungsi sebagai platform untuk penyimpanan, visualisasi, dan klasifikasi data menggunakan algoritma Extreme Gradient Boosting (XGBoost). Integrasi ini memungkinkan pelaksanaan otomatis proses deteksi risiko diabetes – mulai dari pengumpulan dan transmisi informasi melalui Wi-Fi hingga penayangan hasil klasifikasi kepada pengguna secara real-time.

3.3.2 Rancangan Alur Kerja Mikrokontroler ESP32

Rancangan alur kerja perangkat lunak yang ditanamkan pada mikrokontroler ESP32 diprogram menggunakan lingkungan Arduino IDE dengan logika sekuensial yang konstan. Tahap awal alur kerja dimulai dengan inisialisasi jaringan Wi-Fi lokal klinik, protokol komunikasi I2C, serta pin ADC sensor GSR. Setelah koneksi nirkabel dan instrumen sensor terkonfirmasi aktif, mikrokontroler akan melakukan pemindaian data fisik secara berkala saat jari pasien diletakkan pada permukaan sensor. Untuk mengurangi dampak noise (gangguan sinyal fisis) akibat pergerakan kecil dari jari pasien, mikrokontroler menerapkan algoritma penyaringan digital berupa moving average filter (perataan nilai bergerak) secara konstan. Nilai rata-rata pembacaan fisis yang telah bersih kemudian dibungkus ke dalam format JavaScript Object Notation (JSON) yang memuat string data Heart Rate, SpO₂, dan nilai resistansi kulit (GSR). Protokol HTTP POST kemudian dieksekusi oleh ESP32 untuk menembakkan data JSON tersebut secara nirkabel menuju endpoint URL peladen website.

Mikrokontroler ESP32 memulai proses transmisi data setelah memastikan bahwa semua parameter sensor telah berhasil dibaca. Langkah ini memastikan nilai Detak Jantung, SpO₂, dan Respons Kulit Galvanik (GSR) telah berhasil diperoleh serta mencegah pengiriman data yang tidak lengkap ke server. Jika suatu sensor belum menghasilkan data yang valid, mikrokontroler akan terus mengulangi proses pembacaan hingga semua parameter tersedia. Mekanisme ini digunakan untuk menjaga konsistensi data dalam proses klasifikasi, sehingga setiap sampel selalu memiliki satu set lengkap lima variabel masukan.

Setelah semua parameter berhasil diperoleh, ESP32 mengemas data tersebut ke dalam satu paket data yang berisi nilai usia, Indeks Massa Tubuh (BMI), Detak Jantung, SpO₂, dan GSR. Data ini kemudian dikonversi ke format JavaScript Object Notation (JSON) agar dapat diproses oleh server situs web tanpa perlu mengonversi formatnya lagi. Kami memutuskan untuk menggunakan JSON karena ini merupakan format pertukaran data yang ringan, mudah dibaca dan ditulis baik oleh manusia maupun mesin, serta dapat diparsing oleh hampir semua bahasa pemrograman.

Pada fase komunikasi, ESP32 mentransfer data ke titik akhir API yang disediakan di server situs web menggunakan metode HTTP POST. Setelah server menerima data, sistem membalas bahwa transmisi berhasil. Kemudian, ESP32 membaca respons ini dan melanjutkan ke proses pembacaan data berikutnya. Tujuan dari mekanisme ini adalah untuk mengurangi risiko kehilangan data akibat gangguan komunikasi jaringan selama transmisi.

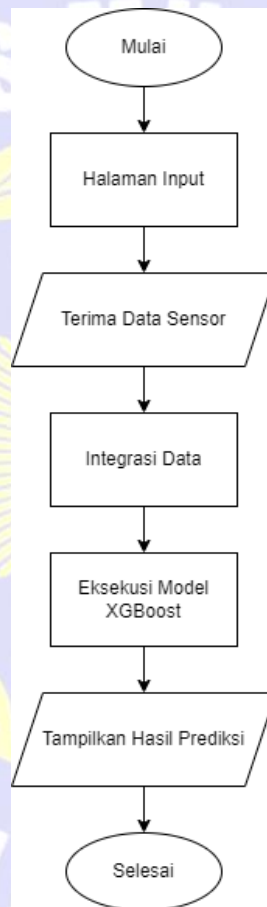
Pembacaan sensor, pemrosesan data awal, dan transfer data dilakukan secara berulang selama sistem aktif. Mekanisme ini memungkinkan pembaruan data fisiologis responden secara berkala, sehingga informasi yang ditampilkan di situs web selalu sesuai dengan hasil pengukuran terbaru. Desain alur kerja ini memungkinkan sistem deteksi risiko diabetes beroperasi secara otomatis dan real-time, mulai dari pengumpulan data hingga penyampaian hasil kepada pengguna, tanpa intervensi manual dalam proses pengukuran.

3.3.4 Rancangan Alur Kerja Website

Rancangan alur kerja platform website dirancang untuk menangani manajemen data, eksekusi model, dan penyajian antarmuka grafis. Sisi backend peladen yang dibangun menggunakan bahasa Python secara konstan mendengarkan (listening) pengiriman data JSON dari perangkat keras ESP32. Ketika paket data sensor diterima, sistem backend akan secara otomatis mencocokkannya dengan input karakteristik fisik pasien berupa usia dan BMI yang dimasukkan secara manual oleh tenaga medis melalui halaman form website. Alur kerja peladen kemudian mengumpankan kelima variabel tersebut ke dalam fungsi pemanggil model (inference engine) XGBoost untuk

memproses klasifikasi risiko secara instan. Hasil keluaran berupa status risiko (Normal atau Berisiko) dikirimkan kembali ke sisi frontend website menggunakan protokol komunikasi dua arah (WebSockets atau AJAX polling) agar visualisasi data pada halaman dashboard medis dapat diperbarui secara dinamis tanpa perlu memuat ulang seluruh halaman website.

Alur kerja website pada sistem deteksi risiko diabetes ditunjukkan pada Gambar 3.5. Flowchart ini menjelaskan proses pengolahan data mulai dari penerimaan data sensor hingga penyajian hasil klasifikasi kepada pengguna.



Gambar 3. 5 Flowchart Alur Kerja Website

Berdasarkan Gambar 3.5, website menerima data sensor yang dikirimkan oleh ESP32, kemudian menggabungkannya dengan data usia dan BMI yang diinput oleh pengguna. Selanjutnya, data tersebut diproses menggunakan model XGBoost untuk menghasilkan

prediksi risiko diabetes. Hasil klasifikasi kemudian ditampilkan pada halaman website secara real-time sehingga pengguna dapat melihat informasi risiko diabetes dengan mudah.

3.3.5 Pengembangan dan Evaluasi Model

Pengembangan model machine learning difokuskan pada optimalisasi kinerja algoritma XGBoost murni dengan memanfaatkan 100 rekaman dataset primer klinis sebagai data latih (training set) untuk menyusun struktur pohon keputusan pembentuk XGBoost. Proses pelatihan model bekerja secara iteratif menggunakan pendekatan boosting, di mana setiap decision tree baru dibangun untuk meminimalkan nilai kesalahan logaritmik (log loss) dari pohon keputusan sebelumnya berdasarkan nilai gradient dan Hessian. Selanjutnya, tahap evaluasi model dilakukan secara ketat dan independen dengan mengumpalkan 30 sampel data uji (testing set) dari pasien baru ke dalam sistem. Luaran prediksi risiko dari model terhadap 30 responden tersebut kemudian dibandingkan dengan label aktual hasil pengujian glukometer medis invasif (Label 0 untuk Normal dan Label 1 untuk Berisiko) sebagai parameter ground truth. Tingkat kesesuaian antara prediksi noninvasif sistem dan kondisi klinis aktual pasien dihitung secara kuantitatif memanfaatkan metrik evaluasi Confusion Matrix, yang meliputi perhitungan persentase akurasi (accuracy), presisi (precision), dan recall untuk membuktikan keandalan skrining dini risiko diabetes yang dikembangkan.

3.3.5.1 Confusion Matrix

Confusion matrix merupakan tabel yang digunakan untuk menggambarkan perbandingan antara hasil klasifikasi model dan kondisi aktual data. Confusion matrix terdiri dari empat komponen utama, yaitu:

- True Positive (TP): jumlah data diabetes yang diklasifikasikan sebagai diabetes
- True Negative (TN): jumlah data non-diabetes yang diklasifikasikan sebagai non-diabetes
- False Positive (FP): jumlah data non-diabetes yang salah diklasifikasikan sebagai diabetes

- False Negative (FN): jumlah data diabetes yang salah diklasifikasikan sebagai non-diabetes

3.3.5.2 Accuracy

Accuracy digunakan untuk mengukur tingkat ketepatan model secara keseluruhan, yaitu seberapa besar proporsi data yang berhasil diklasifikasikan dengan benar oleh model.

Rumus akurasi adalah:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad [7]$$

3.3.5.3 Precision dan Recall

Selain akurasi, evaluasi sistem juga dilakukan menggunakan nilai precision dan recall untuk memberikan gambaran yang lebih rinci mengenai kemampuan sistem dalam mendeteksi diabetes.

Precision digunakan untuk mengukur tingkat ketepatan sistem dalam memprediksi kondisi diabetes, dengan rumus:

$$Precision = \frac{TP}{TP + FP} \quad [[8]]$$

Sedangkan recall digunakan untuk mengukur kemampuan sistem dalam mendeteksi seluruh data diabetes yang sebenarnya, dengan rumus:

$$Recall = \frac{TP}{TP + FN} \quad [[9]]$$

3.4 Rencana Evaluasi Akhir

Tahap rencana evaluasi pada penelitian ini difokuskan pada pengujian performa sistem secara objektif dan terukur. Untuk memastikan sistem pendeteksi risiko diabetes ini berfungsi dengan baik, evaluasi dibagi ke dalam dua indikator pengujian utama, yaitu:

1. Evaluasi Model Pembelajaran Mesin (*Machine Learning*)

- Skenario Uji: Pengujian performa klasifikasi algoritma Extreme Gradient Boosting (XGBoost) dilakukan menggunakan 30 rekaman data uji (testing set) independen yang berasal dari pasien baru di klinik. Data validasi ini dipastikan belum pernah dilatihkan atau dikenali oleh model sebelumnya.
- Parameter Ground Truth: Luaran prediksi risiko dari sistem noninvasif akan dikomparasikan secara langsung dengan hasil validasi klinis glukometer medis invasif (Label 0 untuk Normal dan Label 1 untuk Berisiko).
- Metrik Evaluasi: Tingkat kesesuaian dihitung secara kuantitatif menggunakan matriks kebingungan (Confusion Matrix) yang meliputi nilai True Positive, True Negative, False Positive, dan False Negative. Dari nilai tersebut, akan ditarik persentase akurasi (accuracy), presisi (precision), dan sensitivitas (recall) guna menilai kemampuan keandalan deteksi sistem.

2. Evaluasi Stabilitas Perangkat Keras (*Hardware*)

- Keandalan Sensor: Menilai kestabilan pembacaan fisis parameter masukan dari sensor bio-optik MAX30102 (frekuensi detak jantung dan SpO2) serta sensor analog GSR saat dipasangkan secara langsung pada variasi kondisi fisik jari responden yang riil.
- Reduksi Noise: Menguji efektivitas unit pemroses mikrokontroler ESP32 dalam menerapkan algoritma perataan nilai (moving average filter) untuk mengurangi gangguan sinyal (noise) yang diakibatkan oleh pergerakan kecil dari jari pasien selama proses ekstraksi fitur berlangsung.