



Klasifikasi Penyebaran Jaringan *Wifi Provider* Internet Menggunakan Algoritma XGBoost Berdasarkan Titik Koneksi Kabel Fiber Optik

Repaldo Manihuruk^{1*}, Tengku Mohd Diansyah²

^{1*,2}Program Studi Teknik Informatika, Fakultas Teknik Dan Komputer, Universitas Harapan Medan, Indonesia

^{1*}repaldo.manihuruk@gmail.com, ²dian.10.22@gmail.com

^{*}repaldo.manihuruk@gmail.com

Abstrak-Perkembangan teknologi internet berbasis fiber optik mendorong peningkatan kebutuhan akan distribusi jaringan WiFi yang stabil dan merata. Meskipun provider seperti XYZ telah memiliki infrastruktur fiber optik yang luas, permasalahan distribusi titik akses WiFi masih sering muncul, khususnya terkait ketidakseimbangan sebaran jaringan dan keterbatasan analisis berbasis data. Kondisi ini menimbulkan pertanyaan mengenai penentuan titik penyebaran jaringan WiFi yang optimal agar kualitas layanan tetap terjaga. Oleh karena itu, penelitian ini bertujuan untuk menganalisis pola penyebaran jaringan WiFi provider XYZ berdasarkan titik koneksi kabel fiber optik, menerapkan algoritma Extreme Gradient Boosting (XGBoost) dalam mengklasifikasikan kelayakan distribusi jaringan WiFi, serta mengevaluasi performa algoritma tersebut dalam meningkatkan efisiensi dan optimalisasi penyebaran jaringan. Penelitian ini mengajukan penerapan algoritma XGBoost sebagai metode klasifikasi dalam memprediksi kelayakan distribusi jaringan WiFi berdasarkan data pelanggan yang terkoneksi dengan kabel fiber optik. Fokus penelitian diarahkan pada proses pra-pemrosesan data pelanggan, pembangunan model klasifikasi menggunakan XGBoost, pengujian performa model dalam mengklasifikasikan lokasi yang layak dan tidak layak, serta penyeimbangan data untuk mengatasi masalah ketidakseimbangan kelas. Data penelitian terdiri dari 193 entri pelanggan XYZ yang dibagi menjadi data latih sebesar 80% dan data uji sebesar 20%. Hasil penelitian menunjukkan bahwa algoritma XGBoost mampu memberikan akurasi tinggi dalam klasifikasi distribusi jaringan WiFi. Dengan demikian, model ini dapat dijadikan dasar rekomendasi berbasis data dalam optimasi penyebaran jaringan sehingga provider mampu meningkatkan kualitas layanan internet yang lebih merata, stabil, dan efisien.

Kata kunci: XGBoost, klasifikasi, WiFi, fiber optik

Abstract-The rapid development of fiber-optic-based internet technology has led to an increasing demand for stable and evenly distributed WiFi networks. Although internet service providers such as XYZ have established extensive fiber-optic infrastructure, challenges in WiFi access point distribution remain common, particularly regarding uneven network coverage and limited data-driven analysis. These issues raise the question of how to determine optimal WiFi deployment locations to ensure consistent service quality. Therefore, this study aims to analyze the spatial distribution patterns of XYZ's WiFi network based on fiber-optic connection points, apply the Extreme Gradient Boosting (XGBoost) algorithm to classify the feasibility of WiFi distribution, and evaluate the performance of the proposed model in improving network distribution efficiency. This research employs XGBoost as a classification method to predict suitable and unsuitable WiFi deployment locations using customer data connected via fiber-optic cables. The study focuses on data preprocessing, model construction using XGBoost, performance evaluation in classifying feasible and non-feasible locations, and data balancing techniques to address class imbalance. The dataset consists of 193 XYZ customer records, divided into 80% training data and 20% testing data. The results demonstrate that the XGBoost algorithm achieves high classification accuracy in WiFi network distribution. Consequently, the proposed model can serve as a data-driven recommendation tool for optimizing WiFi deployment, enabling service providers to deliver more evenly distributed, stable, and efficient internet services.

Keywords: XGBoost, classification, WiFi, fiber optics

1. PENDAHULUAN

Jaringan WiFi merupakan teknologi komunikasi nirkabel yang memungkinkan perangkat terhubung ke internet melalui gelombang radio, sehingga memberikan fleksibilitas akses tanpa ketergantungan pada kabel fisik [1]. Perkembangan teknologi yang digunakan saat ini semakin berkembang khususnya dalam dunia telekomunikasi. Perubahan yang cepat pada jaringan telekomunikasi juga didorong oleh kebutuhan pengguna agar tetap terkoneksi kapanpun dan dimanapun. Berbagai fitur aplikasi baru seperti layanan multimedia, video conference, game online, dan layanan internet membutuhkan bandwidth yang lebih besar dengan kecepatan yang tinggi [2]. Fiber optik



memiliki keunggulan utama berupa bandwidth besar, latensi rendah, serta ketahanan terhadap gangguan elektromagnetik, sehingga mampu mendukung kebutuhan akses internet modern yang menuntut stabilitas dan kecepatan tinggi. Meskipun demikian, penerapan fiber optik juga memiliki keterbatasan, antara lain biaya instalasi yang relatif tinggi dan kompleksitas perencanaan jaringan, terutama dalam menentukan titik distribusi layanan yang optimal [3].

Dalam praktiknya, terdapat dua jenis kabel fiber optik yang umum digunakan, yaitu single-mode fiber untuk transmisi jarak jauh dan multi-mode fiber untuk jarak pendek [4]. Pemilihan jenis kabel ini berpengaruh langsung terhadap kualitas jaringan, jangkauan layanan, serta efisiensi distribusi bandwidth. Oleh karena itu, penyebaran jaringan WiFi yang terintegrasi dengan infrastruktur fiber optik memerlukan perencanaan yang matang agar kualitas layanan dapat terjaga secara merata di seluruh area cakupan [5].

Selain itu, pertumbuhan jumlah pelanggan yang terus meningkat dari waktu ke waktu menuntut XYZ untuk melakukan optimasi distribusi jaringan secara berkelanjutan. Tanpa perencanaan yang tepat, peningkatan jumlah pengguna dapat menyebabkan penurunan kualitas layanan, seperti penurunan kecepatan akses dan ketidakstabilan koneksi [6]. Oleh karena itu, diperlukan pendekatan yang mampu menganalisis berbagai faktor secara simultan, seperti lokasi kabel fiber optik, kepadatan pelanggan, serta karakteristik wilayah, guna menghasilkan keputusan yang lebih akurat dalam penentuan titik penyebaran WiFi [7].

Seiring berkembangnya teknologi, pendekatan berbasis machine learning mulai banyak diterapkan dalam pengelolaan dan optimasi jaringan. Salah satu algoritma yang menunjukkan performa unggul dalam berbagai permasalahan klasifikasi dan prediksi adalah Extreme Gradient Boosting (XGBoost) [8]. Algoritma ini merupakan pengembangan dari metode boosting berbasis pohon keputusan yang dirancang untuk meningkatkan akurasi model melalui proses pembelajaran bertahap serta regularisasi yang efektif [5]. Keunggulan XGBoost terletak pada kemampuannya menangani data berukuran besar, kompleks, serta memiliki hubungan nonlinier, sehingga banyak digunakan dalam bidang jaringan, keamanan siber, dan analisis data skala besar [9].

Sebuah penelitian telah mengkaji penerapan XGBoost dalam konteks jaringan, menunjukkan bahwa algoritma XGBoost mampu meningkatkan efisiensi serta akurasi dalam prediksi penyebaran jaringan WiFi. Namun, penelitian tersebut menekankan bahwa performa model sangat bergantung pada kualitas data yang digunakan, sehingga diperlukan tahap pra-pemrosesan yang optimal [10]. Sementara itu, penelitian lain membuktikan keunggulan XGBoost dalam klasifikasi serangan SYN dengan tingkat akurasi yang sangat tinggi pada dataset berskala besar. Meskipun demikian, penelitian tersebut juga mengidentifikasi permasalahan ketidakseimbangan kelas dan keberadaan nilai hilang sebagai faktor yang berpotensi menurunkan performa model apabila tidak ditangani secara tepat [11].

Berdasarkan kajian terhadap penelitian-penelitian sebelumnya, dapat diidentifikasi adanya celah penelitian (research gap). Sebagian besar penelitian terdahulu berfokus pada penerapan XGBoost dalam konteks keamanan jaringan atau prediksi umum tanpa secara spesifik mengkaji distribusi titik akses WiFi yang terintegrasi langsung dengan data pelanggan berbasis fiber optik. Selain itu, masih terbatas penelitian yang secara eksplisit mengombinasikan analisis kelayakan lokasi, penanganan ketidakseimbangan kelas, serta evaluasi performa model dalam konteks optimasi penyebaran jaringan WiFi pada ISP di Indonesia. Dengan demikian, diperlukan penelitian yang tidak hanya menerapkan algoritma XGBoost, tetapi juga menekankan pada kualitas data, proses pra-pemrosesan, dan evaluasi model secara menyeluruh untuk menghasilkan rekomendasi yang lebih aplikatif [12].

Tujuan penelitian ini dibuat untuk menganalisis pola penyebaran jaringan WiFi provider XYZ berdasarkan titik koneksi kabel fiber optik, menerapkan algoritma XGBoost dalam mengklasifikasikan kelayakan penyebaran jaringan WiFi, serta mengevaluasi performa algoritma tersebut dalam meningkatkan efisiensi distribusi jaringan. Penelitian ini juga menitikberatkan pada penanganan ketidakseimbangan data serta pengujian akurasi model dalam mengklasifikasikan lokasi yang layak dan tidak layak untuk penyebaran WiFi.

2. METODE PENELITIAN

2.1 Algoritma XGBoost

XGBoost (*Extreme Gradient Boosting*) adalah algoritma *Machine learning* yang termasuk dalam kategori ensemble learning, lebih spesifiknya dalam metode boosting. *XGBoost* dirancang untuk meningkatkan akurasi model prediksi dengan menggabungkan sejumlah model yang lebih sederhana, seperti pohon keputusan (decision trees), menjadi satu model yang lebih kuat. Proses analisis dengan klasifikasi algoritma XGBoost merupakan algoritma pengembangan dari Gradient Boosting untuk membangun model dengan prediksi yang lebih kuat sehingga memiliki performa baik dalam proses klasifikasi [13]. Algoritma ini digunakan untuk masalah regresi dan klasifikasi, dan sangat populer dalam berbagai kompetisi *Machine learning* karena kemampuannya dalam memberikan hasil yang sangat akurat dan efisien [14]. *XGBoost* menggabungkan konsep dari *gradient boosting*,

di mana model dibangun secara bertahap melalui beberapa iterasi. Setiap iterasi bertujuan untuk mengoreksi kesalahan yang dibuat oleh model sebelumnya. Dalam hal ini, model baru fokus pada data yang sulit diprediksi oleh model sebelumnya. Dengan cara ini, *XGBoost* dapat secara efektif mengurangi kesalahan dan meningkatkan akurasi prediksi dengan meminimalkan fungsi kerugian (loss function) [10].

Salah satu keunggulan *XGBoost* dibandingkan dengan algoritma boosting lainnya adalah regularisasi, yang membantu dalam mengurangi masalah *overfitting* [15]. Regularisasi adalah teknik yang digunakan untuk mengurangi kompleksitas model dan memastikan bahwa model tidak terlalu terfokus pada data pelatihan, yang bisa menyebabkan model menjadi kurang efektif pada data baru. *XGBoost* menggunakan L1 dan L2 *regularization*, yang membantu mengontrol ukuran koefisien model dan mengurangi *overfitting*. *XGBoost* mengoptimalkan fungsi kerugian (loss function) melalui teknik gradient boosting [12]. Fungsi kerugian untuk *XGBoost*, secara umum, dapat dinyatakan sebagai berikut:

$$L(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \Omega(f) \quad (1)$$

Persamaan $L(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \Omega(f)$ merupakan fungsi kerugian total yang digunakan dalam pelatihan model dan bertujuan untuk diminimalkan. Nilai y_i menyatakan nilai aktual data ke- i , sedangkan \hat{y}_i merupakan hasil prediksi model. Fungsi kerugian $l(y_i, \hat{y}_i)$ digunakan untuk mengukur tingkat kesalahan prediksi, yang bentuknya disesuaikan dengan jenis permasalahan, seperti *mean squared error* pada regresi. Selain itu, komponen regularisasi $\Omega(f)$ berperan dalam mengendalikan kompleksitas model dan mencegah *overfitting* dengan memberikan penalti terhadap model yang terlalu kompleks. Simbol f merepresentasikan model atau prediktor yang digunakan, umumnya berupa akar keputusan. Kombinasi antara fungsi kerugian dan regularisasi memungkinkan model mencapai keseimbangan antara akurasi dan kompleksitas.

Fungsi $\Omega(f)$ sendiri dihitung sebagai berikut:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (2)$$

Persamaan di atas merepresentasikan fungsi regularisasi yang digunakan untuk mengendalikan kompleksitas model pohon keputusan. Parameter γ berfungsi sebagai penalti terhadap jumlah daun pohon keputusan, sehingga mendorong terbentuknya model yang lebih sederhana. Variabel T menyatakan jumlah daun pada pohon keputusan, sedangkan w_j merupakan bobot pada daun ke- j . Selain itu, parameter λ merupakan koefisien regularisasi L2 yang berperan dalam membatasi besar bobot pohon, sehingga membantu mencegah terjadinya *overfitting* dan meningkatkan kemampuan generalisasi model.

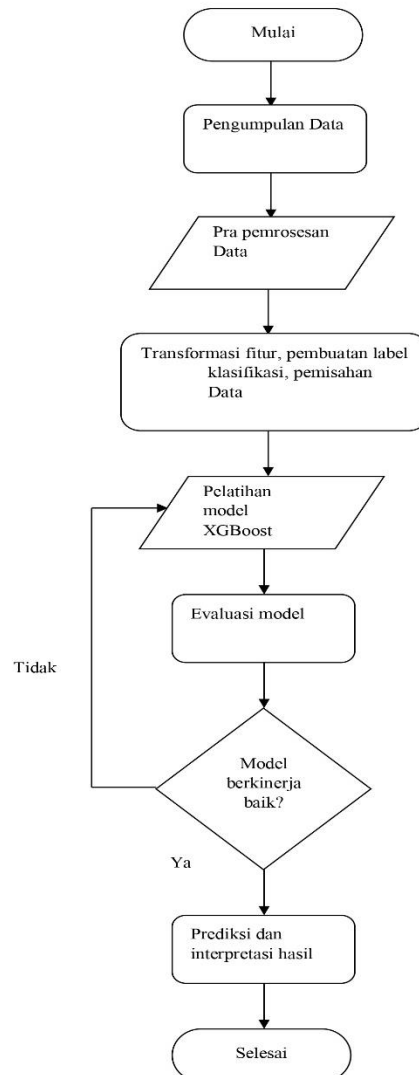
Loss Function (fungsi kerugian) adalah metrik yang digunakan dalam pembelajaran mesin dan deep learning untuk mengukur selisih (error) antara output prediksi model dengan nilai sebenarnya (ground truth). Tujuannya adalah meminimalkan nilai loss agar model semakin akurat dalam melakukan prediksi.

2.2 Alur Penelitian

Alur penelitian ini menjelaskan tahapan-tahapan yang dilakukan dalam proses klasifikasi penyebaran jaringan *WiFi* berdasarkan data pelanggan XYZ [16]. Seluruh proses dilakukan secara interaktif menggunakan *Google Colab*, yang memungkinkan eksekusi kode *Python* berbasis cloud tanpa memerlukan instalasi lokal [17]. Adapun tahapan dalam alur penelitian ini seperti pada gambar 1.

Adapun keterangan gambar 1 yaitu :

1. Pengumpulan Data
Tahap pertama dimulai dengan mengumpulkan dataset pelanggan dari penyedia layanan internet XYZ. Dataset berformat Excel ini berisi informasi seperti nama pelanggan, alamat, dan jenis paket layanan *WiFi* yang digunakan. File ini kemudian diunggah ke *Google Colab* untuk dianalisis lebih lanjut.
2. Pra-pemrosesan Data
Setelah data berhasil diunggah, dilakukan pembersihan data untuk menghapus header ganda, kolom yang tidak diperlukan, serta memperbaiki format kolom agar sesuai dengan struktur *DataFrame*. Langkah ini penting untuk memastikan bahwa data bersih dan dapat digunakan untuk pelatihan model.
3. Transformasi Fitur
Karena model *XGBoost* membutuhkan data numerik, maka dilakukan proses transformasi fitur, seperti mengubah nama paket menjadi nilai kategorikal atau numerik menggunakan teknik encoding. Selain itu, jika diperlukan, alamat dapat diolah untuk menghasilkan fitur wilayah atau kelompok berdasarkan nama jalan.

**Gambar 1.** Tahapan Alur Penelitian

4. Pengumpulan Data
Tahap pertama dimulai dengan mengumpulkan dataset pelanggan dari penyedia layanan internet XYZ. Dataset berformat Excel ini berisi informasi seperti nama pelanggan, alamat, dan jenis paket layanan *WiFi* yang digunakan. File ini kemudian diunggah ke *Google Colab* untuk dianalisis lebih lanjut.
5. Pra-pemrosesan Data
Setelah data berhasil diunggah, dilakukan pembersihan data untuk menghapus header ganda, kolom yang tidak diperlukan, serta memperbaiki format kolom agar sesuai dengan struktur *DataFrame*. Langkah ini penting untuk memastikan bahwa data bersih dan dapat digunakan untuk pelatihan model.
6. Transformasi Fitur
Karena model XGBoost membutuhkan data numerik, maka dilakukan proses transformasi fitur, seperti mengubah nama paket menjadi nilai kategorikal atau numerik menggunakan teknik encoding. Selain itu, jika diperlukan, alamat dapat diolah untuk menghasilkan fitur wilayah atau kelompok berdasarkan nama jalan.
7. Pembuatan Label Klasifikasi
Karena dataset tidak memiliki label kelayakan distribusi secara langsung, maka label dibuat berdasarkan logika atau aturan tertentu.
8. Pemisahan Data

Dataset kemudian dibagi menjadi dua bagian: data pelatihan dan data pengujian. Pembagian ini dilakukan agar model dapat dilatih menggunakan sebagian data, dan diuji akurasi menggunakan data yang tidak pernah dilihat sebelumnya.

9. Pelatihan Model XGBoost

Setelah data siap, model XGBoost dilatih menggunakan fitur-fitur yang telah diproses. Proses pelatihan ini dilakukan di *Google Colab* dengan bantuan pustaka XGBoost, yang sangat efisien untuk klasifikasi dengan performa tinggi.

10. Evaluasi Model

Model yang telah dilatih kemudian dievaluasi menggunakan metrik seperti akurasi, *precision*, *recall*, dan *confusion matrix*. Hal ini bertujuan untuk menilai seberapa baik model dalam mengklasifikasikan data dengan benar.

11. Prediksi dan Interpretasi Hasil

Setelah model diuji, sistem digunakan untuk memprediksi kelayakan penyebaran jaringan pada data baru atau data uji. Hasil prediksi ditampilkan dalam bentuk tabel serta dapat divisualisasikan dalam grafik atau diagram batang.

2.3 Data Penelitian

Penelitian ini menggunakan data pelanggan dari layanan internet XYZ yang terdiri atas 193 entri. Setiap entri mencakup informasi tentang nama pelanggan, alamat, serta jenis paket layanan internet yang digunakan. Data ini diperoleh dari dokumentasi internal dan digunakan untuk menganalisis serta mengklasifikasikan penyebaran jaringan *WiFi* berbasis titik koneksi kabel fiber optik menggunakan algoritma XGBoost. Sebelum digunakan, data telah dibersihkan dari duplikasi header dan entri kosong untuk memastikan validitas pengolahan data. Dalam proses pembangunan model klasifikasi, data ini dibagi menjadi dua bagian utama: data latih (*training*) sebanyak 80% dari total data, yang berjumlah 153 entri, dan data uji (*testing*) sebanyak 20% atau 40 entri. Pembagian ini bertujuan untuk memungkinkan model belajar dari sebagian besar data dan kemudian diuji kemampuannya dalam memprediksi atau mengklasifikasikan data baru secara akurat.

Tabel 1. Data Penelitian

Id	Nama	Alamat	Paket
1	Pelanggan 1	Jl Kenanga Raya	Home Ods
2	Pelanggan 2	Jl Bunga Kenanga	Home Ods
3	Pelanggan 3	Jl Bunga Kenanga	Home Ods
4	Pelanggan 4	Jl Jamin Ginting	Home Ods
5	Pelanggan 5	Jl Jamin Ginting	Home Ods
6	Pelanggan 6	Jl Durian	Home Ods
7	Pelanggan 7	Jl Durian	Home Ods
8	Pelanggan 8	Jl Setia Budi Pasar	Home Ods
9	Pelanggan 9	Jl Setia Budi Pasar	Home Ods
10	Pelanggan 10	Jl Setia Budi Pasar	Home Ods
11	Pelanggan 11	Jl Setia Budi Pasar	Home Ods
12	Pelanggan 12	Jl Setia Budi Pasar	Home Ods
13	Pelanggan 13	Jl Setia Budi Pasar	Home Ods
14	Pelanggan 14	Jl Setia Budi Pasar	Home Ods
15	Pelanggan 15	Jl Setia Budi Pasar	Home Ods
16	Pelanggan 16	Jl Setia Budi Pasar	Home Ods
17	Pelanggan 17	Jl Setia Budi Pasar	Home Ods
18	Pelanggan 18	Jl Setia Budi Pasar	Home Ods
19	Pelanggan 19	Jl Setia Budi Pasar	Home Ods
20	Pelanggan 20	Jl Setia Budi Pasar	Home Ods
...
193	Pelanggan 193	Jl Jend A H Nasution	Home Ods

2.4 Cara Kerja Algoritma XGBoost (Extreme Gradient Boosting)

Algoritma XGBoost (*Extreme Gradient Boosting*) adalah salah satu algoritma *machine learning* berbasis ensemble yang dikenal dengan kecepatan dan akurasi dalam menyelesaikan masalah klasifikasi maupun regresi [18]. Dalam penelitian ini, XGBoost digunakan untuk mengklasifikasikan sebaran jaringan *WiFi* berdasarkan data pelanggan dari layanan XYZ. Untuk mencapai hasil klasifikasi yang optimal, algoritma ini diterapkan melalui serangkaian tahapan sistematis sebagai berikut:

1. Pra-pemrosesan Data

Tahap pertama dimulai dengan pembersihan data dari header ganda, nilai kosong, serta informasi yang tidak relevan seperti nomor urut. Selanjutnya, data kategorikal seperti alamat dan jenis paket layanan diubah menjadi bentuk numerik melalui teknik encoding, agar dapat diproses oleh algoritma XGBoost.

2. Pembagian Dataset

Setelah data dibersihkan dan dikonversi, dataset dibagi menjadi dua bagian, yaitu data latih (80%) dan data uji (20%). Tujuan dari pembagian ini adalah agar model dapat belajar dari data latih dan dievaluasi menggunakan data uji yang belum pernah dilihat sebelumnya, sehingga hasilnya lebih objektif.

3. Pemilihan Fitur Penting

Pada tahap ini dilakukan pemilihan fitur yang dianggap relevan terhadap klasifikasi, seperti informasi alamat pelanggan dan jenis paket yang digunakan. Fitur-fitur ini menjadi variabel input utama dalam pelatihan model.

4. Pelatihan Model XGBoost

Model XGBoost dilatih menggunakan data latih yang telah diproses. Proses pelatihan ini melibatkan pengaturan parameter penting seperti *learning_rate*, *max_depth*, dan jumlah *n_estimators*. Pengaturan parameter ini bertujuan untuk menghindari overfitting dan memastikan model belajar dengan efektif.

5. Evaluasi Model

Setelah pelatihan selesai, model dievaluasi menggunakan data uji. Evaluasi dilakukan dengan menghitung metrik seperti akurasi, *precision*, *recall*, dan F1-score. Metrik-metrik ini memberikan gambaran seberapa baik model dalam mengklasifikasikan data secara benar.

6. Prediksi dan Penyempurnaan Model

Jika model menunjukkan performa yang baik, maka dapat digunakan untuk memprediksi lokasi baru yang potensial untuk penyebaran jaringan *WiFi*. Namun, jika akurasi belum optimal, maka dilakukan tuning parameter dan pelatihan ulang untuk mendapatkan hasil yang lebih baik.

2.4.1 Tahapan Alur Kerja Program

Gambar 1 menggambarkan tahapan sistematis dalam membangun model klasifikasi menggunakan algoritma XGBoost berbasis Python di platform Google Colab [19]. Setiap langkah dalam proses ini dirancang untuk memastikan bahwa data yang digunakan telah diproses dengan benar sebelum digunakan dalam pelatihan model, serta memastikan model dapat dilatih dan dievaluasi dengan baik.

Proses diawali pada tahap inisialisasi sistem yang menandakan dimulainya alur pelatihan model. Langkah pertama adalah unggah dataset berformat CSV yang berisi data pelanggan XYZ. Dataset ini merupakan sumber utama informasi, mencakup nama pelanggan, alamat, serta jenis paket layanan. File CSV diunggah ke Google Colab agar dapat diproses secara langsung dalam lingkungan cloud. Tahap berikutnya adalah import library, yaitu memanggil pustaka Python yang diperlukan dalam pemrosesan data dan pelatihan model. Pustaka yang digunakan meliputi *pandas* untuk manipulasi data, XGBoost untuk implementasi algoritma, *sklearn* untuk preprocessing dan evaluasi, serta *tabulate* dan *IPython.display* untuk menampilkan data dalam format tabel yang lebih rapi. Dengan adanya pustaka-pustaka ini, proses komputasi dapat berjalan lebih cepat dan terstruktur. Selanjutnya dilakukan pemuatan dataset (load dataset) menggunakan fungsi `pd.read_csv()`. Data yang dibaca disimpan dalam variabel `df_raw`. Pada tahap ini, dataset masih mentah sehingga memerlukan penyesuaian lebih lanjut. Oleh karena itu, dilakukan normalisasi nama kolom dengan mengubah seluruh nama kolom menjadi huruf kecil dan menghapus spasi menggunakan fungsi `str.strip()` dan `str.lower()`. Normalisasi ini bertujuan agar nama kolom konsisten dan tidak menimbulkan error saat pemanggilan dalam kode.

Setelah dataset siap, dilakukan validasi data awal dengan menampilkan beberapa baris data pertama menggunakan fungsi `tabulate`. Validasi ini penting untuk memastikan data berhasil dimuat dengan benar serta memeriksa kesesuaian format antar kolom. Kemudian, sistem melakukan pemeriksaan kolom paket. Kolom ini menjadi variabel utama karena merepresentasikan jenis layanan yang digunakan pelanggan. Apabila kolom "paket" tidak ditemukan, proses akan berhenti dan menampilkan error "paket tidak ditemukan". Hal ini menjadi kontrol awal agar proses tidak dilanjutkan dengan data yang tidak lengkap. Jika kolom paket tersedia, langkah berikutnya adalah encoding kolom paket menggunakan *LabelEncoder*. Proses ini mengubah data kategorikal, misalnya paket "Home", "Office", atau "XYZ", menjadi nilai numerik seperti 0, 1, atau 2. Encoding sangat penting karena algoritma XGBoost hanya dapat memproses data numerik. Setelah encoding selesai, dilakukan seleksi fitur dan

label. Pada tahap ini, fitur (X) berisi input data berupa paket_encoded atau alamat_encoded, sedangkan label (y) adalah target yang akan diprediksi, yaitu kategori hasil klasifikasi yang dibuat berdasarkan logika tertentu.

Tahap inti adalah inisialisasi model XGBoost. Pada tahap ini, ditentukan parameter utama seperti $learning_rate = 0.1$, $max_depth = 3$, dan $n_estimators = 100$. $Learning_rate$ mengontrol seberapa cepat model belajar dari kesalahan, max_depth membatasi kedalaman pohon agar tidak terlalu kompleks, sedangkan $n_estimators$ menentukan jumlah pohon yang akan dibangun. Secara matematis, model berusaha meminimalkan fungsi objektif:

$$Obj = \sum_{i=1}^n L(y_i, \hat{y}_i) + \Omega(f_t) \quad (3)$$

dengan L sebagai fungsi kerugian (loss function) dan Ω sebagai regularisasi untuk mengontrol kompleksitas pohon keputusan. Setelah parameter ditetapkan, model masuk ke tahap pelatihan (fit). Proses pelatihan bertujuan untuk meminimalkan selisih antara nilai aktual (y) dan hasil prediksi (\hat{y}_i). Model secara iteratif menambahkan pohon keputusan baru yang berfokus pada data yang masih salah diprediksi. Pada tahap ini, dihitung turunan pertama dari fungsi kerugian (gradient, g_i) dan turunan kedua (Hessian, h_i), yang digunakan untuk memperbaiki bobot setiap pohon. Dengan cara ini, model mampu meningkatkan akurasi secara bertahap pada setiap iterasi boosting. Tahap selanjutnya adalah prediksi output, di mana model yang telah dilatih digunakan untuk menghasilkan prediksi berdasarkan data baru atau data uji. Hasil prediksi dibandingkan dengan nilai aktual untuk mengukur akurasi model. Jika akurasi tinggi, maka model dapat dikatakan berhasil mempelajari pola distribusi jaringan dengan baik. Sebaliknya, jika akurasi rendah, maka dilakukan penyesuaian parameter dan pelatihan ulang. Akhirnya, proses ditutup dengan tahap selesai. Pada titik ini, model XGBoost telah berhasil dibangun, dilatih, dan siap digunakan sebagai alat bantu dalam menganalisis dan memprediksi sebaran jaringan WiFi XYZ. Dengan pendekatan ini, perusahaan dapat lebih mudah menentukan lokasi potensial untuk pengembangan jaringan berdasarkan pola data pelanggan yang telah dipelajari oleh model.

2.4.1.1 Perhitungan Manual

Algoritma XGBoost bekerja dengan meminimalkan fungsi objektif yang terdiri atas loss function dan regularisasi. Loss function (misalnya *log-loss* untuk klasifikasi biner) mengukur selisih antara prediksi dan label aktual, sementara regularisasi mengontrol kompleksitas pohon agar tidak terjadi overfitting [20]. Pada tiap iterasi, XGBoost membangun pohon baru dengan memanfaatkan pendekatan Taylor expansion orde-2, menghasilkan nilai gradien (g) dan Hessian (h). Bobot daun optimal dihitung dengan:

$$w_j^* = -\frac{G_j}{H_j + \lambda} \quad (4)$$

dan gain split dihitung dengan:

$$Gain = \frac{1}{2} \left(\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{G_T^2}{H_T + \lambda} \right) - \gamma \quad (5)$$

dengan G adalah jumlah gradien, H jumlah Hessian, λ parameter regularisasi L2, dan γ penalti jumlah daun.

Dari dokumen, 10 data pertama memuat alamat pelanggan. Untuk keperluan ilustrasi, kita beri label $y=1$ (layak) untuk 3 baris beralamat Jl. Setia Budi Pasar, dan $y=0$ (tidak layak) untuk 7 baris lainnya.

- Jumlah data: 10
 - Positif ($y=1$): 3
 - Negatif ($y=0$): 7
 - Prediksi awal ($f(0)$) diasumsikan $0 \rightarrow$ probabilitas awal $.p = 0,5$
1. Hitung Gradien & Hessian
Untuk log-loss:
 $g_i = p_i - y_i$ $h_i = p_i(1 - p_i)$
Karena semua $p_i = 0,5$:
 - Untuk $y=1$: $g = 0,5 - 1 = -0,5$ $h = 0,25$
 - Untuk $y=0$: $g = 0,5 - 0 = 0,5$ $h = 0,25$
 Akumulasi:
 - Kelas Setia Budi (3 data, $y=1$): $G_L = -1,5$ $H_L = 0,75$
 - Kelas selain itu (7 data, $y=0$): $G_R = +3,5$ $H_R = 1,75$
 - Total: $G_T = 2,0$ $H_T = 2,5$
 2. Hitung Bobot Daun Optimal
Dengan $\lambda=1$:

$$w_L = -\frac{-1.5}{0.75 + 1} = \frac{1.5}{1.75} = 0.8571$$

$$w_R = -\frac{3.5}{1.75 + 1} = -\frac{3.5}{2.75} = -1.2727$$

3. Hitung Gain Split

$$\begin{aligned} \text{Gain} &= \frac{1}{2} \left(\frac{(-1, 5^2)}{0,75 + 1} + \frac{(3, 5)^2}{1,75 + 1} - \frac{(2, 0)^2}{2,5 + 1} \right) \\ &= \frac{1}{2} (1,2857 + 4,4545 - 1,1429) = \frac{1}{2} (4,5973) = 2,298 \end{aligned}$$

Gain positif → split layak dipakai.

4. Update Prediksi

Dengan learning rate $\eta=0.3$:

- Untuk daun L:

$$\Delta f = 0,3 \times 0,8571 = +0,2571 \rightarrow p' = \sigma(0,2571 = 0,564$$

- Untuk daun R:

$$\Delta f = 0,3 \times -1,2727 = -0,3818 \rightarrow p' = \sigma(-0,3818 = 0,405$$

5. Kesimpulan Perhitungan Manual

- Gradien & Hessian dihitung dari label aktual vs prediksi awal.
- Bobot daun: $W_L = 0,8571$ $W_R = -1,2727$
- Gain split = 2.2987 → split efektif.
- Setelah 1 pohon, probabilitas naik untuk data “Setia Budi” (mendekati layak), turun untuk lainnya (mendekati tidak layak).

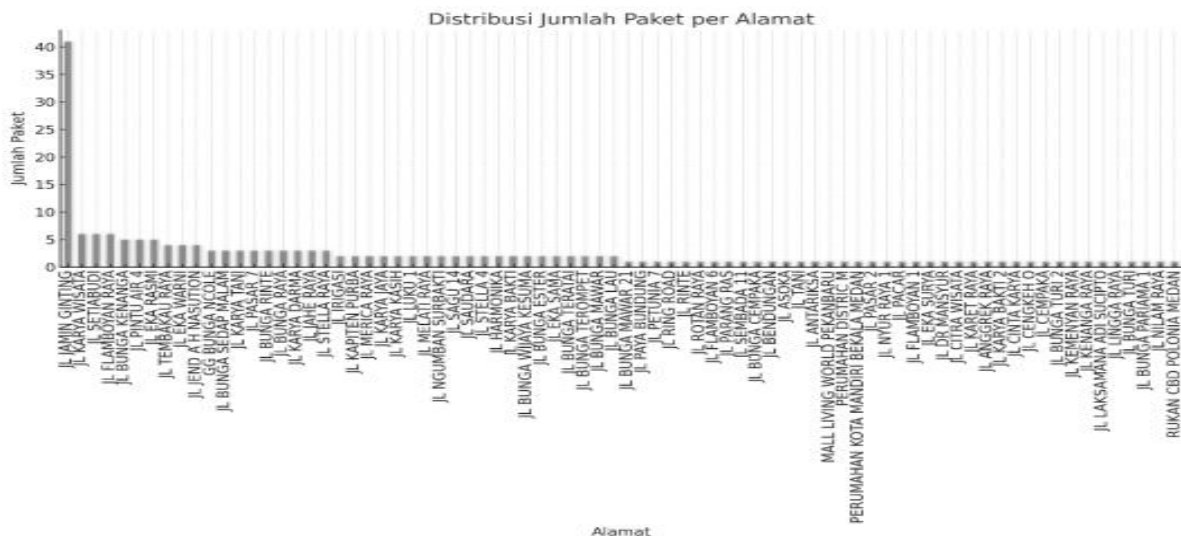
3. HASIL DAN PEMBAHASAN

3.1 Hasil

Pada bab ini akan dibahas secara menyeluruh hasil dari implementasi sistem klasifikasi penyebaran jaringan WiFi menggunakan algoritma XGBoost yang telah dirancang dan dianalisis pada bab sebelumnya. Pembahasan meliputi hasil dari setiap tahap yang telah dilalui mulai dari pra-pemrosesan data, pembuatan label klasifikasi, pelatihan model, hingga evaluasi kinerja model dalam memprediksi titik penyebaran jaringan yang optimal berdasarkan data pelanggan XYZ. Tujuannya adalah untuk menilai efektivitas sistem dalam memberikan solusi berbasis data yang mendukung pengambilan keputusan strategis dalam distribusi jaringan WiFi. Adapun tahapannya adalah sebagai berikut:

1. Deskripsi Dataset

Dataset terdiri dari 193 data pelanggan XYZ yang mencakup kolom nama, alamat, dan jenis paket layanan. Data ini mencerminkan titik-titik pemasangan layanan yang sudah ada, dan menjadi dasar dalam klasifikasi kelayakan distribusi jaringan WiFi.



Gambar 2. Distribusi Jumlah Dataset



2. Pra-Pemrosesan

Pra-pemrosesan data adalah tahap awal yang penting dalam proses analisis data dan pembuatan model machine learning. Tujuan utamanya adalah untuk menyiapkan data mentah agar layak digunakan dalam pelatihan algoritma. Dalam penelitian ini, dataset pelanggan XYZ yang memuat informasi nama, alamat, jenis paket layanan, dan label klasifikasi terlebih dahulu dibersihkan. Proses pertama adalah menghapus baris header ganda dan entri kosong yang berpotensi menyebabkan error dalam proses analisis. Selanjutnya, penyesuaian dilakukan pada nama kolom agar seragam, yaitu semua huruf kecil dan bebas spasi, agar mudah dipanggil dalam program Python. Kemudian, salah satu langkah krusial adalah melakukan transformasi data kategorikal menjadi bentuk numerik melalui teknik label encoding. Dalam hal ini, kolom paket yang semula berisi teks seperti "HOME 0DS" diubah menjadi angka 1, sehingga dapat diproses oleh algoritma XGBoost. Encoding ini tidak hanya menyederhanakan data, tapi juga memungkinkan model mengenali dan belajar dari pola-pola numerik secara optimal. Semua tahapan ini dilakukan menggunakan Python di platform Google Colab yang fleksibel dan efisien untuk komputasi berbasis cloud.

Tabel 2. Pra Pemrosesan Data

No	Nama	Alamat	Paket	Paket Encoded	Label
1	Pelanggan 1	Jl Kenanga Raya	HOME 0DS	1	1
2	Pelanggan 2	Jl Bunga Kenanga	HOME 0DS	1	0
3	Pelanggan 3	Jl Bunga Kenanga	HOME 0DS	1	1
4	Pelanggan 4	Jl Jamin Ginting	HOME 0DS	1	1
5	Pelanggan 5	Jl Bunga Wijaya	HOME 0DS	1	1
Kesuma					

Tabel 3. Hasil Pemrosesan

Paket	Paket_Encoded	Label
HOME 0DS	1	1
HOME 0DS	1	0
HOME 1DS	2	0
HOME 0DS	1	1
HOME 0DS	1	1
HOME 1DS	2	0
HOME 1DS	2	0
GAMERS 3DS	0	1
HOME 0DS	1	1
HOME 0DS	1	1

Kolom paket_encoded merupakan hasil dari proses label encoding terhadap kolom paket. Nilai 1 pada kolom ini menunjukkan bahwa paket layanan WiFi yang digunakan pelanggan adalah "HOME 0DS". Proses encoding ini mengubah nilai teks menjadi angka agar bisa diproses oleh algoritma machine learning. Setiap jenis paket akan memiliki nilai berbeda tergantung urutan kemunculannya, dan dalam data ini, "HOME 0DS" telah diberikan kode 1. Kolom label dalam dataset ini berfungsi sebagai indikator kelayakan distribusi jaringan WiFi pada titik pelanggan tertentu. Nilai pada label ini merupakan hasil dari klasifikasi berbasis logika atau aturan tertentu yang ditetapkan dalam penelitian, seperti jumlah pelanggan pada suatu alamat, jenis paket yang digunakan, atau tingkat permintaan di wilayah tersebut. Apabila sebuah titik memiliki label = 1, maka lokasi tersebut dianggap layak atau menjadi prioritas untuk dilakukan pengembangan atau distribusi jaringan WiFi lebih lanjut. Sebaliknya, jika nilai label = 0, maka titik tersebut dinilai kurang layak atau tidak menjadi prioritas utama untuk perluasan jaringan, karena memiliki jumlah pengguna rendah atau faktor lain yang membuatnya kurang optimal untuk investasi jaringan. Penentuan label ini sangat penting dalam membantu penyedia layanan seperti XYZ dalam menyusun strategi distribusi jaringan yang efisien dan tepat sasaran.

3. Split Data

Tahap split data merupakan proses penting dalam pembangunan model *machine learning* untuk memastikan bahwa model dapat belajar secara efektif dan dievaluasi secara objektif. Dalam penelitian ini, dataset pelanggan dibagi menjadi dua bagian, yaitu 80% data untuk pelatihan (training) sebanyak 153 data, dan 20% untuk pengujian

(testing) sebanyak 40 data. Pembagian ini dilakukan secara acak namun proporsional, artinya distribusi antara label yang berbeda (misalnya label 0 dan 1) tetap terjaga dalam kedua subset data. Tujuannya adalah agar model dapat belajar dari sebagian besar data yang tersedia dan kemudian diuji pada data yang belum pernah dilihat sebelumnya. Dengan pendekatan ini, performa model dapat dinilai secara lebih realistis dan menghindari bias, sehingga hasil klasifikasi yang dihasilkan lebih dapat diandalkan saat diterapkan pada data baru.

```
def preview_split(X_tr, X_te, y_tr, y_te, label_encoder):  
    # Rekonstruksi nama 'paket' dari 'paket_encoded'  
    X_tr_view = X_tr.copy()  
    X_te_view = X_te.copy()  
    X_tr_view['paket'] = label_encoder.inverse_transform(X_tr_view['paket_encoded']).astype(int).values  
    X_te_view['paket'] = label_encoder.inverse_transform(X_te_view['paket_encoded']).astype(int).values
```

Gambar 3. Kodingan Split Data

Ringkasan Split:

- X_train: (144, 1), y_train: (144,)
- X_test : (48, 1), y_test : (48,)

Gambar 4. Split Data

4. Pelatihan Model

Tahap pelatihan model dengan XGBoost merupakan inti dari proses klasifikasi dalam penelitian ini. Model dibangun menggunakan library XGBoost pada bahasa pemrograman Python dan dijalankan melalui platform Google Colab yang menyediakan lingkungan komputasi berbasis cloud. Selama pelatihan, beberapa parameter penting disesuaikan untuk mengoptimalkan performa model, di antaranya: `learning_rate` yang mengatur kecepatan pembelajaran model, `max_depth` yang menentukan kedalaman maksimal pohon keputusan, `n_estimators` yaitu jumlah pohon yang digunakan, serta `objective` yang menentukan jenis tugas (dalam kasus ini adalah klasifikasi). Dengan mempelajari pola-pola yang terdapat dalam data pelanggan, seperti jenis paket dan alamat, model dilatih untuk mengenali karakteristik titik yang layak untuk distribusi jaringan WiFi secara optimal. Proses pelatihan ini bertujuan untuk menghasilkan model prediktif yang mampu memberikan rekomendasi penyebaran jaringan yang lebih akurat dan berbasis data.

```
X_train = df_train[['paket_encoded']]  
y_train = df_train['label']  
  
model = XGBClassifier(  
    learning_rate=0.1,  
    max_depth=3,  
    n_estimators=100,  
    objective='binary:logistic',  
    use_label_encoder=False,  
    eval_metric='logloss'  
)  
  
model.fit(X_train, y_train)  
  
# =====  
# 🚀 OUTPUT AKHIR  
# =====  
print("\n🚀✅ Model XGBoost berhasil dilatih menggunakan file Data_Training.csv.")  
print(f"📊 Jumlah data latih: {len(df_train)}")
```

Gambar 5. Proses Kodingan Menggunakan Data Latih

Potongan kode pada gambar 5 merupakan tahapan inti dalam proses pelatihan model machine learning menggunakan algoritma XGBoost. Proses dimulai dengan pemisahan antara fitur dan target dari data latih (`df_train`). Dalam hal ini, kolom `paket_encoded` digunakan sebagai fitur input (`X_train`), sedangkan kolom label dijadikan target output (`y_train`) yang ingin diprediksi oleh model. Selanjutnya, model XGBoost dikonfigurasi dengan parameter tertentu. Parameter `learning_rate` diatur ke 0.1 untuk mengontrol kecepatan pembelajaran, `max_depth` ditetapkan ke 3 untuk membatasi kedalaman pohon keputusan agar menghindari overfitting, dan `n_estimators` diisi dengan 100 sebagai jumlah pohon yang akan dibentuk. Parameter `objective` disetel ke 'binary:logistic' karena tugas yang dihadapi adalah klasifikasi biner, dan `use_label_encoder` dinonaktifkan sesuai rekomendasi versi terbaru library XGBoost. Model kemudian dilatih menggunakan metode `.fit()` dengan memasukkan data fitur dan label. Setelah proses pelatihan selesai, sistem mencetak pesan yang menyatakan bahwa model telah berhasil dilatih menggunakan file `Data_Training.csv`, serta menampilkan jumlah baris data latih yang digunakan. Output ini menjadi indikator awal bahwa model siap digunakan untuk tahap selanjutnya seperti pengujian atau prediksi.

```
Model XGBoost berhasil dilatih menggunakan file Data_Training.csv.  
Jumlah data latih: 153  
/usr/local/lib/python3.11/dist-packages/xgboost/core.py:158: UserWarning: [18:26:19] WARNING: /workspace/src/learner.cc:740:  
Parameters: { "use_label_encoder" } are not used.
```

Gambar 6. Hasil Output

Gambar 6 menunjukkan bahwa proses pelatihan model XGBoost telah berhasil dijalankan menggunakan data dari file `Data_Training.csv`. Pesan yang ditampilkan menyatakan bahwa model berhasil dibangun dengan jumlah data latih sebanyak 153 baris, sesuai dengan proporsi 80% dari keseluruhan data yang dibagi sebelumnya. Hal ini mengindikasikan bahwa model telah berhasil menerima dan mempelajari pola dari fitur yang telah disiapkan, yaitu kolom `paket_encoded` sebagai masukan dan label sebagai target. Di bawahnya terdapat peringatan (*warning*) dari library XGBoost yang menyatakan bahwa parameter `use_label_encoder=False` tidak lagi digunakan atau diabaikan, karena sejak versi terbaru XGBoost, label encoder bawaan sudah tidak diperlukan. Peringatan ini bukanlah error, melainkan hanya informasi dari sistem, dan tidak memengaruhi jalannya pelatihan model. Dengan demikian, output ini menandakan bahwa model siap digunakan untuk langkah selanjutnya, seperti pengujian atau evaluasi performa menggunakan data testing.

5. Evaluasi Model

Tahap evaluasi model merupakan langkah penting setelah proses pelatihan selesai dilakukan. Evaluasi bertujuan untuk mengukur seberapa baik model yang telah dibangun mampu melakukan prediksi terhadap data baru yang belum pernah dilihat sebelumnya. Dalam penelitian ini, evaluasi dilakukan menggunakan beberapa metrik kinerja utama, yaitu akurasi, *precision*, *recall*, dan *F1-score*. Akurasi mengukur proporsi prediksi yang benar dari seluruh data uji, sedangkan *precision* menilai tingkat ketepatan model dalam mengklasifikasikan suatu kelas tertentu. *Recall* digunakan untuk mengetahui sejauh mana model mampu menangkap seluruh data dari kelas positif, dan *F1-score* memberikan nilai rata-rata harmonis antara *precision* dan *recall*, yang sangat berguna ketika terjadi ketidakseimbangan kelas. Selain evaluasi berbasis angka, hasil prediksi juga dianalisis menggunakan *confusion matrix*, yaitu tabel yang memperlihatkan perbandingan antara prediksi dan label sebenarnya dalam bentuk matriks. Visualisasi ini membantu mengidentifikasi berapa banyak kesalahan klasifikasi yang terjadi, serta seberapa baik model dalam membedakan kelas target. Untuk melengkapi interpretasi, hasil prediksi model dapat divisualisasikan dalam bentuk grafik atau plot, sehingga memudahkan pemahaman terhadap kinerja model secara menyeluruh dan intuitif.

Hasil Prediksi (10 data teratas):		
Paket	Label Sebenarnya	Prediksi Model
HOME 0DS	1	0
HOME 0DS	1	0
HOME 2DS	0	0
HOME 1DS	1	0
HOME 1DS	0	0
HOME 1DS	1	0
HOME 0DS	0	0
HOME 1DS	1	0
HOME 1DS	1	0
HOME 1DS	0	0

Gambar 7. Hasil Prediksi

Gambar 7 menampilkan tabel hasil prediksi model terhadap 10 data pertama dari dataset uji (*Data_Testing.csv*). Kolom-kolom dalam tabel meliputi jenis paket layanan pelanggan (Paket), label sebenarnya (Label Sebenarnya), dan prediksi dari model (Prediksi Model). Dalam konteks ini, label 1 menunjukkan bahwa lokasi pelanggan dianggap layak atau prioritas untuk distribusi jaringan WiFi, sedangkan label 0 berarti tidak layak atau bukan prioritas. Dari tabel tersebut, terlihat bahwa sebagian besar prediksi model adalah 0, bahkan untuk data yang sebenarnya berlabel 1. Sebagai contoh, untuk paket HOME 0DS dan HOME 1DS, model memprediksi 0 meskipun label sebenarnya adalah 1. Hal ini menunjukkan bahwa model mengalami underfitting, yaitu cenderung memprediksi ke satu kelas saja dan tidak cukup sensitif membedakan lokasi layak dan tidak layak. Masalah ini dapat disebabkan oleh kurangnya variasi fitur, data yang tidak seimbang, atau model yang terlalu sederhana. Perlu dilakukan analisis lanjutan untuk mengetahui penyebabnya dan meningkatkan performa model.

4. KESIMPULAN

Model klasifikasi berbasis algoritma XGBoost yang dibangun menggunakan data pelanggan XYZ telah berhasil melalui seluruh tahapan pemodelan, mulai dari pembersihan data, transformasi fitur, hingga proses pelatihan dan evaluasi. Pembersihan data, khususnya penghapusan header ganda, serta penerapan encoding pada variabel kategorikal dan pemisahan data latih serta data uji, terbukti mampu menyederhanakan struktur data dan memastikan proses pelatihan berjalan secara sistematis. Namun demikian, hasil evaluasi menunjukkan bahwa performa model masih tergolong rendah dengan tingkat akurasi sebesar 38%. Model menunjukkan ketidakseimbangan prediksi yang signifikan, di mana seluruh prediksi cenderung mengarah pada satu kelas dominan (label 0), sementara kelas minoritas (label 1) sama sekali tidak terdeteksi. Hasil penelitian ini menunjukkan perlunya penanganan lebih lanjut terhadap ketidakseimbangan data, seperti penerapan teknik resampling atau penyesuaian parameter model, agar kinerja klasifikasi dapat ditingkatkan secara signifikan pada penelitian selanjutnya. Dengan demikian hasilnya memberikan rekomendasi berbasis data untuk mengoptimalkan penyebaran jaringan WiFi sehingga dapat meningkatkan kualitas dan keandalan layanan internet, mendukung tercapainya koneksi WiFi yang lebih stabil dan merata di berbagai lokasi sesuai dengan kebutuhan pengguna, serta menjadi referensi bagi penelitian selanjutnya khususnya dalam penerapan algoritma XGBoost untuk analisis dan optimasi jaringan telekomunikasi.

REFERENSI

- [1] A. P. Sinaga, I. Syahputra, Melati, and Nurbaiti, "Optimalisasi Jaringan Wifi (Wireless Fidelity) sebagai Fasilitas Pendukung Akademik Mahasiswa (Studi Kasus di UINSU)," *Cognoscere J. Komun. dan Media Pendidik.*, vol. 2, no. 4, pp. 18–25, 2024, doi: 10.61292/cognoscere.244.
- [2] W. M. Prayoga and A. Sani, "Perancangan Jaringan Fiber To the Home (Ftth) Menggunakan Teknologi Gigabit Passive Optical Network (Gpon)," *EBIDEkonomi Bisnis Digit.*, vol. 1, no. 2, pp. 179–188, 2023,



- doi: 10.37365/ebid.v1i2.220.
- [3] Y. A. Hegia Theodosius Sitepu, Miftahul Ilmi, "Jurnal Teknologi Digital dan Sistem Informasi," *J. Teknol. Digit. dan Sist. Inf.*, vol. 1, no. 2, pp. 107–115, 2024.
- [4] W. A. Fadila, Q. Aini, and F. A. Wahyudi, "Perkembangan Teknologi Pemanfaatan Fiber Optik Dalam Industri Telekomunikasi Untuk Koneksi Jaringan," *Opt. J. Pendidik. Fis.*, vol. 8, no. 2, pp. 309–320, 2024, doi: 10.37478/optika.v8i2.4204.
- [5] H. Nugroho and S. A. Siagian, "Analisis Bandwidth Jaringan Wifi," *ICT Penelit. dan Penerapan Teknol.*, vol. 4, no. 6, pp. 35–43, 2013.
- [6] B. W. Aulia, M. Rizki, P. Prindiyana, and S. Surgana, "Peran Krusial Jaringan Komputer dan Basis Data dalam Era Digital," *JUSTINFO | J. Sist. Inf. dan Teknol. Inf.*, vol. 1, no. 1, pp. 9–20, 2023, doi: 10.33197/justinfo.vol1.iss1.2023.1253.
- [7] K. Insani and F. A. Alijoyo, "Bisnis Jaringan Internet Fiber Optic di Era Digitalisasi untuk Pemerataan Infrastruktur Telekomunikasi," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 6, no. 4, pp. 799–805, 2024, doi: 10.47233/jteksis.v6i4.1572.
- [8] J. M. A. S. Dachi and P. Sitompul, "Comparison Analysis of the XGBoost Algorithm and Random Forest Ensemble Learning Algorithm in Credit Decision Classification," *J. Ris. Rumpun Mat. Dan Ilmu Pengetah. Alam*, vol. 2, no. 2, pp. 87–103, 2023.
- [9] S. E. Herni Yulianti, Oni Soesanto, and Yuana Sukmawaty, "Gani, A. G., Firdaus, I., Amien, J. A., Informatika, T., Komputer, F. I., Riau, U. M., & Matching, S. (2020). Jurnal Computer Science and Information Technology (CoSciTech). Jurnal Sistem Informasi Universitas Suryadarma, 3(2), 1–19.," *J. Math. Theory Appl.*, vol. 4, no. 1, pp. 21–26, 2022.
- [10] A. Muhammad Augie Rudianto, E. Sakti Pramukantoro, and D. Kurnianingtyas, "Implementasi Sistem Deteksi Anomali pada Jaringan Komputer dengan Pendekatan XGBoost dan Data SNMP," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 2, pp. 2548–964, 2025, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [11] M. D. Noverta Effendi*1, Witrihan Ramadhani2, Fitri Farida3, "Jurnal Computer Science and Information Technology (CoSciTech) things," *J. Comput. Sci. Inf. Technol.*, vol. 5, no. 2, pp. 358–366, 2024.
- [12] A. Fauzihan, B. Sajiwo, B. Rahmat, and A. Junaidi, "Udaran (Ispu) Menggunakan Xgboost Dengan Teknik Imbalanced Data," vol. 12, no. 3, 2024.
- [13] J. Iskandar, V. C. Mawardi, and J. Hendryli, "Analisis Media Sosial Penyedia Layanan Internet Menggunakan Algoritma XGBOOST," *Semin. Nas. Corisindo*, pp. 78–83, 2022.
- [14] M. Asnawi and B. Hadi, "7644-Article Text-18743-3-10-20241015," *J. Econ. Manag. Account. Technol.*, vol. 7, no. 2, pp. 379–389, 2024.
- [15] Nazwa Fadhil, "Perbandingan Akurasi Algoritma Xgboost Dan Svr Dalam Prediksi Harga Cryptocurrency," *J. Ilmu Komput. dan Sist. Inf.*, vol. 13, no. 1, 2025, doi: 10.24912/jiksi.v13i1.32865.
- [16] N. M. Surbakti *et al.*, "Penggunaan Bahasa Pemrograman Python dalam Pembelajaran Kalkulus Fungsi Dua Variabel," *Algoritma. J. Mat. Ilmu Pengetah. Alam, Kebumihan dan Angkasa*, vol. 2, no. 3, pp. 98–107, 2024, doi: 10.62383/algoritma.v2i3.67.
- [17] R. Y. Hayuningtyas, Wina Yusnaeni, and Ida Darwati, "Penerapan Algoritma XGBoost Dalam Menganalisa Keberlanjutan Pelanggan Tour dan Travel," *Evolusi J. Sains dan Manaj.*, vol. 13, no. 2, pp. 25–32, 2025, doi: 10.31294/evolusi.v13i2.9748.
- [18] M. Anggito Herlambang Hadisuwarno *et al.*, "Komparasi performa model machine learning algoritma XGBoost dan Random Forest pada studi kasus mendeteksi stunting," *AITI J. Teknol. Inf.*, vol. 22, no. 2, pp. 266–278, 2025.
- [19] R. Gelar Guntara, "Pemanfaatan Google Colab Untuk Aplikasi Pendeteksian Masker Wajah Menggunakan Algoritma Deep Learning YOLOv7," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 5, no. 1, pp. 55–60, 2023, doi: 10.47233/jteksis.v5i1.750.
- [20] U. N. Mandiri, "OPTIMALISASI ALGORITMA XGBOOST MENGGUNAKAN HYPERPARAMETER TUNING DAN MULTIPLE PREPROCESSING UNTUK PREDIKSI HARGA MOBIL BEKAS TESIS Diajukan sebagai salah satu syarat untuk memperoleh gelar Magister Ilmu Komputer (M. Kom)," 2021.