

BAB I

PENDAHULUAN

Untuk mendukung proses klasifikasi status gizi pada balita, pemanfaatan teknologi kecerdasan buatan menjadi salah satu solusi yang semakin banyak dikembangkan, khususnya melalui pendekatan klasifikasi berbasis data tabular.

1.1 Latar Belakang

Gizi yang baik pada balita sangat berpengaruh terhadap pertumbuhan normal, perkembangan fisik, dan kecerdasan anak [1]. Pemenuhan gizi pada anak usia dibawah lima tahun perlu diperhatikan dalam menjaga kesehatan anak. Karena masa balita merupakan periode perkembangan yang rentan gizi. Balita merupakan kelompok masyarakat yang rentan gizi [2]. Pada kelompok tersebut mengalami siklus pertumbuhan dan perkembangan yang membutuhkan zat-zat gizi yang lebih besar dari kelompok umur lainnya. Permasalahan gizi ini sering terjadi pada anak dibawah lima tahun yang disebut masa emas, karena pertumbuhan fisik dan perkembangan psikomotor, mental, serta sosial yang sangat pesat. Namun, periode ini juga bersifat kritis jika balita tidak mendapatkan asupan gizi yang memadai sehingga tumbuh kembangnya bisa terhambat [3].

Di Indonesia, status gizi balita merupakan salah satu masalah kesehatan yang masih sering dijumpai. Kondisi gizi buruk dapat dipengaruhi oleh faktor langsung seperti asupan makanan, serta faktor tidak langsung seperti penyakit infeksi, sosial ekonomi, dan demografi [4]. Oleh karena itu, pemantauan status gizi secara rutin melalui fasilitas layanan dasar, seperti puskesmas dan posyandu, menjadi sangat penting. Penilaian status gizi didasarkan pada standar WHO yang tertuang dalam Keputusan Menteri Kesehatan Nomor 1995/Menkes/SK/XII/2020 melalui tiga indeks utama: BB/U, TB/U, dan IMT/U [5]. Selain itu, penilaian status gizi balita juga dapat dipengaruhi oleh berbagai variabel lain seperti pemberian ASI, riwayat sakit atau infeksi, kelengkapan imunisasi, pendidikan orang tua, frekuensi makan harian, sanitasi lingkungan, hingga tingkat pendapatan keluarga. Faktor-faktor tersebut berperan dalam menentukan kondisi kesehatan dan pertumbuhan balita secara menyeluruh [6]. Namun, dalam praktiknya, penilaian status gizi di lapangan khususnya di posyandu masih banyak dilakukan secara manual. Proses pengukuran

dan klasifikasi sering memerlukan perhitungan standar deviasi, pencatatan manual, serta interpretasi petugas berdasarkan grafik pertumbuhan. Kondisi ini berpotensi menimbulkan keterlambatan identifikasi, kesalahan pencatatan, dan ketidaktepatan klasifikasi [7]. Hal ini juga terjadi pada posyandu Desa Pesen, Kecamatan Kanor, Kabupaten Bojonegoro, yang masih menggunakan pembukuan manual untuk menentukan kategori gizi balita.

Permasalahan tersebut dapat diatasi menggunakan teknologi yang mampu melakukan klasifikasi status gizi secara cepat, akurat, dan konsisten. Pendekatan *machine learning* menjadi solusi inovatif karena mampu mengolah data dalam jumlah besar, mengenali pola, serta memberikan prediksi otomatis. Dua algoritma yang dipilih dalam penelitian ini adalah *Extreme Gradient Boosting* (XGBoost) dan *Light Gradient Boosting Machine* (LightGBM), yang termasuk dalam kelompok *Gradient Boosting Decision Tree* (GBDT). Pemilihan kedua algoritma ini bukan tanpa alasan. Dibandingkan metode lain seperti KNN, SVM, atau *Decision Tree* tunggal, XGBoost dan LightGBM mampu memberikan performa yang lebih tinggi pada data kompleks karena mengoptimalkan proses boosting secara lebih efisien [8]. Keduanya dikenal unggul dalam akurasi prediksi, mampu menangani data tidak seimbang, bekerja cepat pada dataset besar, serta memiliki mekanisme regularisasi yang kuat untuk mengurangi *overfitting*. Selain itu, kedua algoritma ini konsisten menjadi pilihan utama dalam berbagai penelitian kesehatan dan deteksi penyakit, sehingga secara empiris terbukti lebih stabil dan reliabel untuk permasalahan klasifikasi berbasis data tabular.

Teknik optimasi juga dilakukan untuk meningkatkan akurasi model secara maksimal, yaitu menggunakan metode GridSearchCV. Metode ini dipilih karena mampu melakukan pencarian kombinasi *hyperparameter* secara sistematis dan menyeluruh melalui mekanisme *Grid Search* yang digabungkan dengan *Cross-Validation*. Tidak seperti *Random Search* yang hanya mengevaluasi sebagian kombinasi secara acak, GridSearchCV memastikan setiap kombinasi parameter diuji sehingga peluang menemukan konfigurasi terbaik menjadi lebih tinggi. Proses validasi silang pada setiap kombinasi juga membuat model lebih tahan terhadap *overfitting* dan memiliki generalisasi yang lebih baik. Dengan keunggulan tersebut,

GridSearchCV menjadi metode yang lebih tepat ketika diperlukan optimasi yang akurat, terarah, dan dapat dipertanggungjawabkan secara metodologis [9].

Berbagai penelitian sebelumnya menunjukkan bahwa algoritma XGBoost dan LightGBM memiliki performa yang sangat baik dalam berbagai tugas klasifikasi. LightGBM terbukti unggul dalam efisiensi dan kecepatan, seperti pada studi Srinivasan (2024) yang mencatat akurasi 91,2% dan AUC 0,94 dalam analisis risiko transaksi blockchain, serta pada penelitian Noviandy et al. (2024) yang mencapai akurasi 98,13% dalam klasifikasi persetujuan pinjaman dengan dukungan *Random Search* dan interpretabilitas SHAP. Di sisi lain, XGBoost menunjukkan stabilitas dan akurasi lebih tinggi pada kasus-kasus tertentu, seperti deteksi serangan DDoS oleh Mahendra dan Putra (2024) dengan akurasi 97,62%, serta diagnosis penyakit kardiovaskular (Alemerein et al., 2024) dan klasifikasi kanker payudara (Wijayanto & Hartarto, 2024) yang masing-masing menunjukkan performa paling konsisten pada variasi dataset. Penerapan teknik optimasi seperti GridSearchCV pada berbagai studi juga terbukti mampu meningkatkan performa model secara signifikan, sebagaimana ditunjukkan pada penelitian Yulianti et al. (2022) dan Nemer et al. (2024) yang berhasil meningkatkan akurasi model hingga di atas 83 sampai 98%. Secara keseluruhan, kedua algoritma sama-sama unggul, dengan LightGBM lebih efisien untuk dataset besar, sedangkan XGBoost lebih stabil dan akurat pada berbagai kondisi data [10].

Berdasarkan referensi tersebut, penelitian ini bertujuan membandingkan performa XGBoost dan LightGBM dalam mengklasifikasi status gizi balita karena kedua algoritma ini terbukti unggul dalam akurasi, kecepatan, serta kemampuan menangani data kompleks dan tidak seimbang. Untuk memastikan performa terbaik, optimasi hyperparameter dilakukan menggunakan GridSearchCV yang memungkinkan pencarian parameter secara sistematis dan akurat melalui validasi silang, sehingga lebih dapat diandalkan dibanding metode optimasi lainnya. Evaluasi model menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score* guna memperoleh gambaran performa yang komprehensif. Hasil akhirnya diharapkan menghasilkan model yang tidak hanya akurat, tetapi juga efisien untuk diimplementasikan pada platform *website*, sehingga kader posyandu dapat menilai status gizi balita dengan cepat dan tepat sebagai dasar intervensi, seperti pemberian

makanan tambahan, pemantauan rutin, edukasi orang tua, hingga rujukan ke fasilitas kesehatan bila diperlukan.

1.2 Rumusan Masalah

Rumusan masalah penelitian yang dilakukan didapatkan berdasarkan uraian latar belakang adalah sebagai berikut :

1. Bagaimana cara penerapan GridSearchCV pada algoritma XGBoost dan LightGBM untuk klasifikasi status gizi pada balita ?
2. Bagaimana performa model LightGBM dan XGBoost sebelum dan sesudah di lakukan optimasi menggunakan GridSearchCV dalam klasifikasi status Gizi pada Balita ?
3. Bagaimana hasil perbandingan nilai parameter algoritma XGBoost dan LightGBM sebelum dan sesudah optimasi GridSearchCV dalam meningkatkan kinerja untuk klasifikasi status gizi pada balita?
4. Bagaimana implementasi model terbaik untuk klasifikasi status gizi dalam platform *website*?

1.3 Tujuan Penelitian

Pada rumusan masalah yang telah disusun, maka penelitian ini diarahkan untuk mencapai tujuan-tujuan sebagai berikut :

1. Mengetahui proses penerapan GridSearchCV pada algoritma XGBoost dan LightGBM untuk klasifikasi status gizi balita.
2. Menganalisis performa model LightGBM dan XGBoost sebelum dan sesudah dilakukan penyetelan *Hyperparameter* GridSearchCV
3. Mengetahui hasil analisis perbandingan performa algoritma XGBoost dan LightGBM sebelum dan sesudah dilakukan penyetelan *Hyperparameter* GridSearchCV.
4. Untuk mengembangkan dan mengintegrasikan model untuk klasifikasi gizi ke dalam platform *website* yang mudah diakses oleh kader posyandu.

1.4 Manfaat Penelitian

Pada penelitian ini diharapkan dapat membantu serta memberikan manfaat sebagai berikut :

1. Manfaat bagi Penulis

- Penulis dapat menerapkan hasil pengetahuan dan pembelajaran yang didapat selama kuliah pada penelitian
- Penulis dapat memperluas wawasan dan mengimplementasikan dibidang kesehatan
- Sebagai salah satu syarat kelulusan sarjana program studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur
- Dapat meningkatkan kemampuan dan keterampilan dalam menggunakan algoritma Gradien Boosting yang dipilih yaitu XGBoost dan LightGBM dengan tambahan optimasi *Hyperparameter* GridSearchCV untuk menyelesaikan masalah klasifikasi pada status gizi balita

2. Manfaat bagi Universitas

- Menambah Kumpulan skripsi dari salah satu mahasiswa Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” jawa Timur Mengenai Perbandingan Kinerja Algoritma GBM (XGBoost dan LightGBM) dengan Optimasi *Hyperparameter* GridSearchCV
- Menjadi sebuah tolak ukur bagi universitas dalam menentukan keberhasilan dan kemampuan penulis dalam pengaplikasian wawasan yang sudah didapatkan selama pendidikan jurusan Informatika.

3. Manfaat dibidang keilmuan

- Memperluas pengetahuan mengenai penerapan GridSearchCV dalam mengoptimalkan algoritma XGBoost dan LightGBM dalam klasifikasi status gizi pada balita.
- Menjadi referensi dalam pengembangan ilmu machine learning, khususnya dalam optimasi algoritma XGBoost dan lightGBM untuk klasifikasi status gizi pada balita. Hasil penelitian ini dapat menjadi acuan bagi akademis dalam memahami efektifitas *hyperparameter* tuning dalam meningkatkan performa model klasifikasi

4. Manfaat dibidang kesehatan

- Membantu dalam mengambil keputusan untuk mengidentifikasi status gizi pada balita secara lebih cepat dan akurat melalui teknologi *machine learning*.
- Meningkatkan efektivitas pencegahan gizi buruk sehingga dapat menerima penanganan yang lebih sesuai sejak dini.

1.5 Batasan Masalah

Adapun batasan masalah yang ditetapkan supaya pembahasan yang terdapat dalam penelitian ini tidak menyimpang, yaitu sebagai berikut :

1. Dataset yang digunakan berasal dari puskesmas kecamatan kanor kabupaten bojonegoro yang mencangkup variabel inti yang tersedia meliputi data antropometri balita seperti usia dalam bulan, berat badan, dan panjang badan, yang kemudian digunakan untuk menentukan variabel target utama yaitu status gizi balita yang diklasifikasikan menjadi Gizi Baik, Gizi Kurang, dan Gizi Buruk. Selain itu, dataset ini juga diperkaya dengan faktor-faktor pendukung lainnya yang memengaruhi status gizi, seperti indeks massa tubuh (*BMI*), riwayat pemberian ASI, riwayat sakit, kelengkapan imunisasi balita.
2. Pada penelitian ini hanya berfokus pada perbandingan kinerja algoritma XGBoost dan LightGBM sebelum dan sesudah optimasi menggunakan *Hyperparameter GridSearchCV*.
3. Evaluasi model dilakukan menggunakan metrik akurasi, presisi, recall, F1-score.
4. Output yang dihasilkan dalam penelitian ini yaitu tingkat keakurasaian hasil klasifikasi metode XGBoost dan LightGBM setelah dilakukan optimasi menggunakan *Hyperparameter GridSearchCV*.