BAB I
PENDAHULUAN

1.1 Latar Belakang

Kualitas sumber daya manusia merupakan aset fundamental bagi kemajuan sebuah negara. Indonesia masih menghadapi tantangan serius dalam pemenuhan gizi, khususnya masalah stunting yang berdampak pada kualitas generasi masa depan. Di tingkat regional, Provinsi Jawa Barat sebagai salah satu provinsi dengan populasi terbesar memiliki tantangan tersendiri dalam menekan angka prevalensi *stunting*. Data historis menunjukkan fluktuasi jumlah kasus diberbagai daerah yang dipengaruhi oleh faktor demografi dan intervensi kesehatan tahunan (Kemenkes RI, 2024). Ketidakpastian tren kenaikan atau penurunan kasus ini sering kali menyulitkan pemerintah daerah dalam merencanakan alokasi anggaran dan sumber daya intervensi secara akurat untuk tahun-tahun mendatang.

Untuk mengatasi ketidakpastian tersebut, diperlukan pendekatan berbasis data yang mampu memproyeksikan kondisi masa depan. Salah satu metode yang relevan adalah analisis deret waktu (*time series forecasting*) menggunakan *machine learning*. Berbeda dengan metode statistik konvensional, algoritma *machine learning* modern seperti *Random Forest* dan *XGBoost* memiliki kemampuan untuk menangkap pola nonlinear yang kompleks dari data historis tahunan, seperti tren jumlah balita stunting dan data demografi kelahiran di setiap wilayah.

Meskipun demikian, tinjauan terhadap literatur terkini menunjukkan bahwa mayoritas penelitian stunting di Indonesia masih didominasi oleh pendekatan klasifikasi pada data individu (Haris *et al*., 2023). Studi yang berfokus pada peramalan (*forecasting*) tren agregat tingkat Kabupaten dan Kota masih sangat terbatas. Padahal, informasi mengenai prediksi total kasus di masa depan jauh lebih strategis bagi pembuat kebijakan untuk menentukan prioritas wilayah penanganan. Selain keterbatasan topik, terdapat *research gap* dalam aspek metodologi optimasi. Penelitian prediksi yang ada sering kali hanya menerapkan model dengan parameter standar (*default*). Padahal, karakteristik data deret waktu (*time series*) yang memiliki dimensi temporal (tahun ke tahun) sangat sensitif terhadap konfigurasi *hyperparameter*. Model yang tidak dioptimasi berisiko menghasilkan error prediksi

(RMSE/MAE) yang tinggi, sehingga hasil peramalan menjadi kurang valid untuk dijadikan dasar kebijakan (Fajri dan Primajaya, 2023).

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk memprediksi jumlah kasus dan prevalensi stunting di Kabupaten dan Kota Provinsi Jawa Barat menggunakan pendekatan *time series regression*. Penelitian ini akan mengkomparasi performa algoritma *Random Forest* dan *XGBoost*, dengan kebaruan (*novelty*) pada penerapan optimasi *hyperparameter* menggunakan *Randomized search* untuk meminimalkan tingkat kesalahan peramalan. Hasil penelitian ini diharapkan dapat memberikan peta prediksi yang akurat bagi pemerintah Jawa Barat dalam menyusun strategi mitigasi stunting yang proaktif.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka rumusan masalah dalam penelitian ini sebagai berikut :

1. Bagaimana perbandingan performa prediksi antara algoritma *Random Forest Regression* dan *Extreme Gradient Boosting* (XGBoost) dalam memprediksi angka kejadian stunting di Provinsi Jawa Barat?
2. Seberapa efektif metode *Randomized search* dalam meningkatkan akurasi dari kedua model tersebut?
3. Bagaimana penerapan metode optimasi pada model *Random Forest* atau *XGBoost* dapat meningkatkan akurasi prediksi angka stunting?

1.3 Tujuan Penelitian

Bedasarkan latar belakang yang telah dijelaskan maka tujuan dari penelitian ini sebagai berikut :

1. Melakukan perbandingan kinerja base model algoritma *Random Forest* dan *XGBoost* untuk tugas regresi dalam prediksi stunting.
2. Menganalisis efektivitas penggunaan model metode *Randomized search* dalam upaya meningkatkan kinerja pada model *Random Forest* dan *XGBoost*.
3. Menetukan model manakah yang menghasilkan kinerja regresi paling optimal setelah melakukan perbandingan algoritma dan proses tuning menggunakan metode *Randomized Search*.

1.4 Manfaat

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut :

1. Bagi instansi Kesehatan hasil penelitian ini dapat menjadi solusi berbasis data memprediksi tren angka stunting dimasa depan, sehingga perencanaan program dapat dilakukan dengan baik bedasarkan bukti.
2. Bagi pengambil kebijakan dan perencana program, model prediksi ini dapat menjadi alat bantu pengambilan keputusan dalam menentukan kebijakan dan program di wilayah prioritas sehingga lebih efektif dan objektif untuk pencegahan stunting.
3. Bagi dunia akademik, penelitian ini dapat memperkaya kajian penerapan algoritma *machine learning Random Forest* dan *XGBoost* dalam menangani masalah regresi data serta menyajikan perbandingan antara base model dan tuned model menggunakan *Randomized Search.*