

1. Pendahuluan

Latar Belakang

Stunting adalah salah satu masalah yang ada di Indonesia. *Stunting* atau yang disebut juga pendek adalah keadaan yang dialami anak berupa terhambatnya pertumbuhan anak. *Stunting* dapat disebabkan oleh masalah gizi kronis atau kurangnya asupan gizi dalam jangka waktu yang cukup panjang. Terhambatnya pertumbuhan anak yang terkena *stunting* berakibat pada rendahnya tinggi anak dibandingkan dengan tinggi anak-anak sekitarnya. Pada data *stunting* pada anak juga terdapat asosiasi negatif pada dengan kurangnya kemampuan motorik halus [1] dan kemampuan serta capaian kognitif pada anak [2]. *Stunting* merupakan indikator keseluruhan terbaik dari kesehatan anak dan refleksi akurat dari kesenjangan sosial. *Stunting* juga merupakan bentuk malnutrisi yang paling umum [3].

Berdasarkan hasil Survei Status Gizi Indonesia (SSGI), angka prevalensi *stunting* di Indonesia pada tahun 2022 berada di angka 21,6%. Angka prevalensi ini lebih kecil dibandingkan tahun 2021 yang berada di angka 24,4% [4]. Berdasarkan hal tersebut diketahui *stunting* masih diderita oleh 1 dari 5 anak di Indonesia. Angka tersebut masih jauh dari Rencana Pembangunan Jangka Menengah Nasional (RPJMN) yang berencana agar angka *stunting* di Indonesia berada di angka 14% pada tahun 2024 [5].

Mengontrol perkembangan *stunting* di Indonesia merupakan hal yang penting agar tujuan yang ingin dicapai pemerintah dapat dicapai. Salah satu hal yang penting adalah kemampuan dalam memprediksi arah perkembangan *stunting* di Indonesia berdasarkan data yang telah didapatkan. Kemampuan dalam memprediksi *stunting* dapat dijadikan acuan mengenai bagaimana efek dari usaha yang telah dilakukan pemerintah terhadap arah pertumbuhan *stunting* di Indonesia. Salah satu metode yang biasa digunakan untuk melakukan prediksi adalah pembelajaran mesin. Dengan pembelajaran mesin kita dapat melakukan prediksi *stunting* berdasarkan data yang ada.

Dalam studi ini dilakukan pengimplementasian dan komparasi efektivitas algoritma *artificial neural network* dan *gradient-boosted decision tree*. Algoritma *gradient-boosted decision tree* digunakan karena algoritma tersebut merupakan algoritma yang efisien dan terukur. Kontrol terhadap *overfitting* pada algoritma *gradient-boosted decision tree* juga lebih baik dibandingkan dengan algoritma lainnya. Tingkat sukses algoritma *gradient-boosted decision tree* tergolong tinggi pada kompetisi Kaggle. *Artificial neural network* adalah salah satu algoritma yang telah banyak diimplementasikan. *Artificial neural network* merupakan model yang terinspirasi dari jaringan neuron pada makhluk hidup[6]. *Artificial neural network* dipilih sebagai salah satu model prediksi pada penelitian ini karena kemampuan belajarnya yang cepat, stabil ketika dihadapkan dengan gangguan kecil, adaptif, efisien, dan dapat diimplementasikan secara paralel sehingga waktu pelatihan model dapat dipersingkat [7]. Pengimplementasian algoritma *artificial neural network* dan *gradient-boosted decision tree* sebagai model regresi yang dilakukan pada data yang didapat dari situs web kementerian dalam negeri Indonesia yang berisikan data *stunting* dari sebagian kabupaten/kota dari 34 provinsi di Indonesia pada tahun 2019 - 2023. Performa algoritma diukur dengan metrik *mean squared error* (MSE), *mean absolute percentage error* (MAPE), dan *R-squared*. Hasil prediksi dari implementasi *artificial neural network* dan *gradient-boosted decision tree* digunakan sebagai penentu performa dari setiap algoritma. Hasil perbandingan performa algoritma tersebut dapat dijadikan sebagai acuan dalam pemilihan algoritma yang lebih efektif dalam prediksi prevalensi *stunting* di Indonesia. Algoritma yang efektif dapat digunakan sebagai model prediksi prevalensi *stunting* yang optimal sehingga angka prevalensi *stunting* di Indonesia diketahui dan diprediksi secara lebih akurat.

Tujuan

Tujuan dari penelitian ini adalah:

1. Untuk mengetahui cara memproses data *stunting* di Indonesia.
2. Mengimplementasi *artificial neural network* dan *gradient-boosted decision tree* untuk memprediksi *stunting*.
3. Mengetahui performa *artificial neural network* dan *gradient-boosted decision tree* sebagai model prediksi *stunting* di Indonesia.

Batasan Masalah

Berikut adalah batasan yang terdapat pada penelitian ini:

- Data yang digunakan berupa data yang didapat dari situs web kementerian dalam negeri Indonesia yang berisikan data *stunting* dari sebagian kabupaten/kota dari 34 provinsi di Indonesia pada tahun 2019 - 2023.
- Metrik pengukuran yang digunakan sebagai pembanding kualitas model adalah metrik *mean squared error* (MSE), *mean absolute percentage error* (MAPE), dan *R-squared*.

Organisasi Tulisan

Pada bagian selanjutnya dijelaskan mengenai studi terkait yang menjelaskan mengenai referensi dan landasan

teori yang digunakan dalam penelitian, sistem yang dibangun yang menjelaskan mengenai metode yang digunakan pada penelitian, evaluasi yang berisi hasil pengujian dan analisis hasil pengujian, dan kesimpulan dari studi ini.

2. Studi Terkait

2.1. Tinjauan Pustaka

Sebelumnya telah ada beberapa penelitian mengenai *stunting* dengan algoritma pembelajaran mesin. S.M. Jubaidur Rahman dan koleganya pada tahun 2021, dilakukan penelitian mengenai faktor risiko *stunting*, *wasting*, dan kekurangan nutrisi pada anak yang dideteksi dengan pendekatan pembelajaran mesin. Algoritma yang digunakan pada penelitian tersebut adalah algoritma *logistic regression*, *random forest* (RF), dan SVM. Pada penelitian tersebut, prediksi terbaik didapat dari algoritma RF dengan akurasi 88,3% untuk prediksi *stunting*, 87,7% untuk prediksi *wasting*, dan 85,7% untuk prediksi kekurangan berat badan [8]. Fikrewold H Bitew beserta koleganya pada penelitiannya yang dipublikasikan pada tahun 2021 melakukan penelitian mengenai penggunaan algoritma pembelajaran mesin dalam prediksi kekurangan nutrisi pada balita di Ethiopia. Pada penelitian tersebut digunakan algoritma pembelajaran mesin *k-nearest neighbors* (k-NN), *random forest*, *neural network*, *extreme gradient boosting* dan *generalised linear models* dalam prediksi kekurangan nutrisi (*stunting*, *wasting*, kekurangan berat badan) di Ethiopia. Pada hasil penelitian tersebut, diketahui bahwa dengan implementasi *extreme gradient boosting tree*, didapatkan hasil prediksi kekurangan nutrisi di Ethiopia yang lebih unggul dibandingkan algoritma lainnya [9]. Penelitian *stunting* pada balita di Zambia dilakukan model pengklasifikasian dengan algoritma pembelajaran mesin *random forest* (RF), *SV clasification*, XGBoost, *linear regression*, dan *naive bayes* (NB). Pada penelitian tersebut didapatkan bahwa algoritma *random forest* adalah algoritma yang paling baik dengan tingkat akurasi 79% pada pengujian dan 61,6% pada pelatihan [10].

Di Indonesia terdapat penelitian mengenai linear model dalam memprediksi prevalensi *stunting* dengan pembelajaran mesin. Mambang beserta koleganya melakukan penelitian mengenai kemampuan model regresi dalam memprediksi *stunting* di Indonesia. Pada penelitian tersebut, digunakan berbagai macam model regresi yang terdapat pada *library* scikit-learn. Model yang digunakan adalah *linear regression* (LR), *ridge*, *lasso*, *orthogonal matching pursuit*, *tweedie regressor*, *polynomial regression with pipeline* dan *no pipeline*, SVR, *k-neighbors regressor*, dan MLPRegressor. Hasil yang didapatkan dengan pengujian menggunakan variabel maksimum didapatkan model terbaik *polynomial regression with pipeline* dengan MAPE 0,00, RMSE, 37,9, dan *R-squared* 1,00. Hasil pengujian dengan variabel rata-rata didapatkan model terbaik yaitu *polynomial regression* dengan MAPE 0,01, RMSE 3,32, dan *R-squared* 1,00. Dan untuk pengujian dengan variabel minimum model *polynomial regression with pipeline* mendapatkan hasil terbaik dengan MAPE 0.02, RMSE 3,32, dan *R-squared* 1,00 [11].

2.2. *Stunting* pada Anak

Menurut WHO, *stunting* (pendek) didefinisikan sebagai kondisi anak yang memiliki skor *height-for-age z* (HAZ) < -2 SD dan *severe stunting* (sangat pendek) didefinisikan sebagai kondisi anak yang memiliki skor *height-for-age z* (HAZ) < -3 SD [12], [13]. HAZ dihitung dengan umur dikurangi nilai median sesuai jenis kelamin dari populasi standar dan dibagi dengan SD dari populasi [14]. *Stunting* merupakan indikator keseluruhan terbaik dari kesehatan anak dan refleksi akurat dari kesenjangan sosial. *Stunting* juga merupakan bentuk malnutrisi yang paling umum [3].

Diestimasi, 161 juta anak di dunia terdampak *stunting* di dunia pada tahun 2013. Jutaan anak di dunia menderita beberapa tingkat pertumbuhan yang goyah (tidak stabil) [3]. Di Indonesia, berdasarkan hasil Survei Status Gizi Indonesia (SSGI), angka prevalensi *stunting* di Indonesia pada tahun 2022 berada di angka 21,6%. Angka prevalensi ini lebih kecil dibandingkan tahun 2021 yang berada di angka 24,4% [4].

2.3. *Artificial Neural Network*

Artificial neural network merupakan sebuah teknik pembelajaran mesin yang terinspirasi oleh model awal pemrosesan indra oleh otak. Pada *artificial neural network* dilakukan simulasi jaringan model neuron pada komputer. Model yang dapat dijadikan sebagai solusi dari berbagai masalah dapat dibuat dengan penerapan algoritma yang terinspirasi oleh neuron otak pada jaringan model neuron pada komputer.