

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Indonesia memiliki kontribusi 90% dari masalah gizi di antara 36 negara di dunia. Pada tahun 2010 hasil Riset Kesehatan Dasar (RISKESDAS) menunjukkan prevalensi stunting pada balita di Indonesia sebesar 35,6% yang terdiri atas kriteria pendek dan sangat pendek sekitar 17,1% dan 18,5%. Prevalensi persentase tersebut lebih tinggi dari persentase prevalensi kekurangan gizi sekitar 17,9%, gizi buruk 13,3% dan gizi lebih 14%.

Berdasarkan data dari Bappenas, selama 2018- 2019 Provinsi Jawa Timur merupakan salah satu provinsi yang menjadi wilayah prioritas penanganan permasalahan stunting. Pada tahun 2018, Jawa Timur memiliki 11 lokus untuk penanggulangan stunting (antara lain adalah Kabupaten Jember, Kabupaten Nganjuk dan Kab Lamongan), dan di tahun 2019 bertambah 1 kabupaten, yaitu Kabupaten Kediri. Meskipun data Riset Kesehatan Dasar tahun 2018 menunjukkan bahwa terjadi penurunan angka stunting dibandingkan pada tahun 2013, angka stunting di Jawa Timur masih berada pada angka lebih dari 30% 5 (Laksono & Megatsari,2017) .

Malnutrisi kronis selama pertumbuhan dan perkembangan awal digambarkan oleh stunting dari z-skor tinggi badan untuk usia (TB/U) kurang dari -2 standar deviasi (SD) berdasarkan standar pertumbuhan WHO. Dalam stunting umum mempengaruhi 1 dari 4 balita. Ada efek jangka pendek yang disebabkan oleh stunting yaitu penghambatan kemampuan untuk berbicara, perkembangan motorik dan kognitif, risiko kecacatan, penyakit infeksi, dan kematian. Ada juga dampak jangka panjang yang mengakibatkan risiko penyakit degeneratif, seperti penyakit jantung coroner, stroke, hipertensi dan diabetes mellitus. Selanjutnya bisa mengurangi produktivitas kerja di masa dewasa (zeniarja, widia, & Rahmad,2020).

Pada penelitian ini peneliti ingin membuktikan bahwa menentukan balita stunting menggunakan algoritma pengklasifikasian bisa digunakan, dimana

metode ini sering digunakan dalam ilmu data mining untuk menentukan suatu rekaman data baru ke salah satu dari beberapa jenis yang sudah di definisikan sebelumnya. Salah satu dari algoritma klasifikasi yaitu algoritma klasifikasi *K-Nearest Neighbor* (KNN). Algoritma klasifikasi *K-Nearest Neighbor* adalah metode klasifikasi yang menentukan kategori berdasarkan mayoritas kategori pada *k-Nearest Neighbor*, dilakukan dengan mencari kelompok  $k$  objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data testing (Junta, 2020). Nilai  $k$  yang terlalu kecil, maka hasil klasifikasi akan lebih terpengaruh oleh noise. Di sisi lain, jika nilai  $k$  terlalu tinggi akan mengurangi efek noise pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi lebih kabur. Nilai  $k$  yang bagus dapat dipilih dengan optimasi parameter, misalnya dengan menggunakan *cross-validation* (Banjarsari, Budiman & Farmadi, 2015).

Pada penelitian terdahulu yaitu pada penelitian Junta Zeniarja dengan judul Penerapan Algoritma *Naive Bayes* dan *Forward Selection* dalam Pengklasifikasian Status Gizi Stunting pada Puskesmas Pandanaran Semarang. Dan penelitian Hamsir Saleh dengan judul Analisa faktor penyebab stunting menggunakan algoritma *c4.5*. dengan adanya permasalahan tersebut maka penulis melakukan penelitian dengan judul “ Penerapan metode *K-Nearest Neighbor* (*K-NN*) untuk klasifikasi Stunting pada balita”.

## 1.2 Rumusan Masalah

Dalam penelitian ini penulis menitikberatkan permasalahan pada :

1. Berapa tingkat akurasi tertinggi pada hasil klasifikasi metode *K-Nearest Neighbor* terhadap data balita Stunting?
2. Berapa tingkat presisi tertinggi pada hasil klasifikasi metode *K-Nearest Neighbor* terhadap data balita Stunting?
3. Berapa tingkat *recall* tertinggi pada hasil klasifikasi metode *K-Nearest Neighbor* terhadap data balita Stunting?

### 1.3 Tujuan

Dari hasil analisis penelitian ini hal yang dituju antara lain;

1. Mengukur tingkat akurasi tertinggi dari hasil klasifikasi metode *K-Nearest Neighbor* terhadap data balita Stunting.
2. Mengukur tingkat presisi tertinggi pada algoritma *K-Nearest Neighbor* untuk klasifikasi data balita Stunting.
3. Mengukur tingkat *recall* tertinggi pada algoritma *K-Nearest Neighbor* untuk klasifikasi data balita Stunting

### 1.4 Manfaat

Dari hasil penelitian ini dapat dimanfaatkan terhadap:

1. Penelitian ini dapat digunakan sebagai referensi untuk mengembangkan penelitian lebih lanjut yang berkaitan dengan algoritma *K-Nearest Neighbor*.
2. Pengembangan Education Data Mining dengan atribut yang berbeda.

### 1.5 Batasan Masalah

1. Data penelitian yang digunakan yaitu balita Stunting sebanyak 200 record data di puskesmas Sumbermalang.
2. Perhitungan jarak menggunakan *Euclidean distance* dengan nilai  $k = 3, 5, 7$  dan  $9$ .
3. Skenario uji menggunakan *K Fold Cross Validation* dengan nilai  $k_f = 2, 4, 5$ , dan  $10$
4. Hasil dari penelitian ini berupa akurasi, presisi, dan *recall*.
5. Tools yang digunakan yaitu Rapidminer