

## BAB II KAJIAN PUSTAKA

### 2.1 Penelitian Terdahulu

Referensi yang diambil dari peneliti terdahulu yakni Nani Purwati (2016) dengan judul penelitian “Klasifikasi Status Gizi Balita Berdasarkan Indeks Anthropometri Bb/U Dan Bb/Tb Menggunakan Jaringan Syaraf Tiruan”. Meneliti tentang klasifikasi status gizi menggunakan jaringan syaraf tiruan *backpropagation* berdasarkan indeks antropometri Bb/U yang akan menghasilkan status gizi kedalam gizi buruk, kurang, baik, dan lebih, serta mengklasifikasikan status gizi berdasarkan indeks antropometri Bb/Tb yang akan menghasilkan status sangat kurus, kurus, normal dan gemuk. Variabel yang digunakan dalam penelitian ini yaitu jenis kelamin, umur, berat badan, tinggi badan, dan status ekonomi dengan nilai akurasi sebesar 89,50% untuk indeks BB/U dan 88,50% untuk indeks BB/TB.

Untuk penelitian dari Yushintia Pramitarini, dkk (2013) berjudul “Analisa Rekam Medis Untuk Menentukan Status Gizi Anak Balita Menggunakan *Naïve Bayes Classifier*”. Variable yang digunakan dalam penelitian ini adalah yaitu jenis kelamin, umur, tinggi badan, berat badan, komorbid (penyakit yang sering diderita anak) diantaranya diare, edema, dan pneumonia. Pada penelitian ini dilakukan pengujian kelayakan penggunaan aplikasi dengan pembagian presentase data training dan data uji yang digunakan. Diperoleh hasil prosentase 70:30 maka nilai akurasinya 88%, 50:50 maka nilai akurasinya 55% dan 30:70 maka nilai akurasinya 28%.

Penelitian yang dilakukan Ervina Rizka Anandita (2014) melakukan penelitian yang berjudul “Klasifikasi Tebu Dengan Menggunakan Algoritma *Naïve Bayes Classification* Pada Dinas Kehutanan dan Perkebunan Pati”. Penelitian ini membahas tentang klasifikasi dan pemilihan tebu yang produktif atau tidak produktif. Metode yang digunakan yakni *Naïve Bayes Classifier* dan data yang digunakan menggunakan atribut tebu meliputi jenis tebu, umur panen, tinggi tanaman, diameter batang, daerah tanam, bobot batang,

rendeman, serta macam got dengan 28 data. Uji validitas menggunakan *Confusion matrix c* dan dapat dikatakan produktif atau tidak dengan presentase kinerja sebesar 73,3%.

Sedangkan penelitian kali ini berjudul “Identifikasi status gizi balita menggunakan metode *Naïve Bayes*”. Yang membedakan penelitian ini yakni menggunakan metode *Naïve Bayes* berdasarkan perhitungan kartu menuju sehat pada permasalahan identifikasi status gizi balita. Variable yang digunakan berupa umur, jenis kelamin, dan berat badan. Status gizi atau klasifikasinya terdiri dari gizi buruk, gizi kurang, gizi baik, dan gizi lebih. Untuk lebih jelasnya mengenai penelitian terdahulu dapat dilihat pada Tabel 2.1.

Tabel 2.1 Penelitian Terdahulu

No.	Judul	Objek	Kriteria	Output
1	Klasifikasi Status Gizi Balita Berdasarkan Indeks Anthropometri Bb/U Dan Bb/Tb Menggunakan Jaringan Syaraf Tiruan (Nani Purwati, 2016)	mengklasifikasikan status gizi menggunakan Jaringan Saraf Tiruan Backpropagation berdasarkan indeks antropometri BB/U dan BB/TB	Jenis kelamin, umur, berat badan, tinggi badan, dan status ekonomi	Nilai akurasi sebesar 89,50% untuk indeks BB/U dan 88,50% untuk indeks BB/TB
2	Analisa Rekam Medis Untuk	Pengujian pembagian presentase data	Jenis kelamin, umur, tinggi	Diperoleh hasil prosentase 70:30 maka

	Menentukan Status Gizi Anak Balita Menggunakan Naïve Bayes Classifier (Yushintia Pramitarini dkk, 2013)	training dan data uji pada kasus menentukan sttaus gizi balita	badan, berat badan, komorbid (penyakit yang sering diderita anak)	nilai akurasinya 88%, 50:50 maka nilai akurasinya 55% dan 30:70 maka nilai akurasinya 28%
3	Klasifikasi Tebu Dengan Menggunakan Algoritma Naïve Bayes Classification Pada Dinas Kehutanan dan Perkebunan Pati (Ervina Rizka Anandita, 2014)	Menerapkan metode naïve bayes untuk pengklasifikasian tebu yang dilakukan berdasarkan kelas produktif dan tidak produktif	Jenis tebu, umur panen, tinggi tanaman, diameter batang, daerah tanam, bobot batang, rendeman, serta macam got	Presentase kinerja sebesar 73,3% yang disebabkan oleh banyaknya data kontinu dibandingkan dengan data diskrit

## 2.2 Penilaian Status Gizi Balita

Status gizi adalah suatu ukuran mengenai kondisi tubuh seseorang yang dapat dilihat dari makanan yang dikonsumsi dan penggunaan zat-zat gizi di dalam tubuh. Status gizi dibagi menjadi tiga kategori, yaitu status gizi kurang, gizi normal, dan gizi lebih (Almatsier, 2005).

Status Gizi merupakan tanda-tanda penampilan seseorang akibat keseimbangan antara pemasukan dan pengeluaran zat gizi yang berasal dari

pangan yang dikonsumsi pada suatu saat berdasarkan kategori dan indikator yang digunakan. Seseorang diidentifikasi status gizi normal apabila terdapat keseimbangan antara jumlah energi yang masuk ke dalam tubuh dan energi yang dikeluarkan dari luar tubuh sesuai dengan kebutuhan individu. Energi yang masuk ke dalam tubuh dapat berasal dari karbohidrat, protein, lemak dan zat gizi lainnya (Nix, 2005).

Sedangkan status gizi kurang (*undernutrition*) merupakan keadaan gizi seseorang dimana jumlah energi yang masuk lebih sedikit dari energi yang dikeluarkan. Hal ini dapat terjadi karena jumlah energi yang masuk lebih sedikit dari anjuran kebutuhan individu. Sedangkan Status gizi lebih (*overnutrition*) merupakan keadaan gizi seseorang dimana jumlah energi yang masuk ke dalam tubuh lebih besar dari jumlah energi yang dikeluarkan (Nix, 2005).

Penilaian status gizi balita berguna untuk keperluan rujukan dari kelompok masyarakat atau puskesmas, selain itu dapat pula digunakan untuk pemantauan pertumbuhan anak yang sangat dibutuhkan orangtua bagi perkembangan anak. Ada berbagai cara untuk menilai status gizi salah satunya dengan cara pengukuran tubuh manusia yang dikenal dengan istilah "*Antropometri*".

Menentukan status gizi balita harus ada ukuran bakunya. Ukuran baku yang sekarang di gunakan di Indonesia adalah baku *World Health Organization National center for Health Statistics* (WHO-NCHS). Penilaian status gizi balita dipisahkan antara laki-laki dan perempuan. Ada beberapa macam *antropometri* yang umum digunakan yaitu Umur (U), Berat Badan (BB), Tinggi Badan (TB), Lingkar Perut (LP), Lingkar Lengan Atas (LLA), Lingkar Dada (LD), dan Lapisan Lemak Bawah Kulit (LLBK) dan Lingkar Kepala (LK).

Penilaian status gizi berdasarkan Indeks BB/U (Berat Badan menurut Umur), TB/U (Tinggi Badan menurut Umur), BB/TB (Berat Badan menurut Tinggi Badan) dengan standar baku *anthropometri* WHO-NCHS dapat digolongkan menjadi:

Tabel 2.2 Sebutan Status Gizi

No.	Indeks yang Dipakai	Sebutan Status Gizi
1	BB/U	Gizi Buruk, Gizi Kurang, Gizi Baik, dan Gizi Lebih
2	TB/U	Sangat Pendek, Pendek, Normal, dan Tinggi
3	BB/TB	Sangat Kurus, Kurus, Normal, dan Gemuk

### 2.3 Klasifikasi

Menurut Prasetyo (2012), klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Klasifikasi adalah salah satu pembelajaran yang paling umum di *data mining*. Klasifikasi didefinisikan sebagai bentuk analisis data untuk mengekstrak model yang akan digunakan untuk memprediksi label kelas. Kelas dalam klasifikasi merupakan atribut dalam satu set data yang paling unik yang merupakan variabel bebas dalam statistic.

Dalam klasifikasi terdapat dua proses yang dilakukan yaitu dengan membangun model untuk disimpan sebagai memori dan menggunakan model tersebut untuk melakukan pengenalan atau klasifikasi atau prediksi pada suatu data lain supaya diketahui di kelas mana objek data tersebut dimasukkan berdasarkan model yang telah disimpan dalam memori.

### 2.4 Klasifikasi Naïve Bayes

Naïve Bayes adalah salah satu algoritma pembelajaran induktif yang paling efektif dan efisien untuk *machine learning* dan data mining. Performa naïve bayes yang kompetitif dalam proses klasifikasi walaupun menggunakan asumsi keidependenan atribut atau tidak ada kaitan antar atribut (Tina R, 2013). Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan.

*Naïve bayes* termasuk ke dalam pembelajaran *supervised*, sehingga pada tahapan pembelajaran dibutuhkan data awal berupa data pelatihan untuk dapat mengambil keputusan. Pada tahapan pengklasifikasian akan dihitung nilai probabilitas dari masing-masing label kelas yang ada terhadap masukan yang diberikan. Label kelas yang memiliki nilai probabilitas paling besar yang akan dijadikan label kelas data masukan tersebut. *Naive bayes* merupakan perhitungan *teorema bayes* yang paling sederhana, karena mampu mengurangi kompleksitas komputasi menjadi multiplikasi sederhana dari probabilitas.

Keuntungan penggunaan Naive Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian (Saleh, 2015). Naïve Bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan.

Persamaan dari teorema *Bayes* adalah (Bustami, 2013):

$$P(H|X) = \frac{P(H|X) \cdot P(H)}{P(X)} \quad (2.1)$$

Dimana :

X = Data dengan *class* yang belum diketahui

H = Hipotesis data merupakan suatu *class* spesifik

$P(H|X)$  = Probabilitas hipotesis H berdasar kondisi X (posteriori probabilitas)

$P(H)$  = Probabilitas hipotesis H (prior probabilitas)

$P(X|H)$  = Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$  = Probabilitas X

Sedangkan untuk klasifikasi dengan data kontinyu digunakan rumus *Densitas*

*Gauss*:

$$P(X_i = x_i | Y = Y_j) = \frac{1}{\sqrt{2\pi} \sigma_{ij}} e^{-\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}} \quad (2.2)$$

Di mana:  $Y$  : Kelas yang dicari  
 $P$  : Peluang  $Y_i$  : Sub kelas  $Y$  yang dicari  
 $X_i$  : Atribut ke  $i$   $\mu$  : *mean*, rata – rata dari seluruh atribut  
 $X_i$  : Nilai atribut ke  $i$   $\sigma$  : Deviasi standar, varian dari seluruh atribut.

Alur metode naïve bayes dalam perhitungan penelitian ini yakni :

1. Baca *data training*
2. Hitung Jumlah dan probabilitas, namun apabila data numerik maka :
  - a. Cari nilai *mean* (rata-rata) dan standar deviasi (ukuran penyebaran data dalam sampel dan seberapa dekat titik data individu ke mean) dari masing-masing parameter yang merupakan data numerik. Adapun persamaan yang digunakan untuk menghitung nilai rata – rata hitung (*mean*) dapat dilihat sebagai berikut:

$$\mu = \frac{\sum_{i=1}^n X_i}{n} \quad (2.3)$$

atau

$$\mu = \frac{X_1 + X_2 + \dots + X_n}{n} \quad (2.4)$$

di mana :

$\mu$  : rata – rata hitung (*mean*)

$x_i$  : nilai sample ke  $-i$   $n$  : jumlah sampel



dan persamaan untuk menghitung nilai simpangan baku (standar deviasi) dapat dilihat sebagai berikut:

$$\sigma = \frac{\sum_{i=1}^n (X_i - \mu)^2}{n - 1} \quad (2.5)$$

di mana :

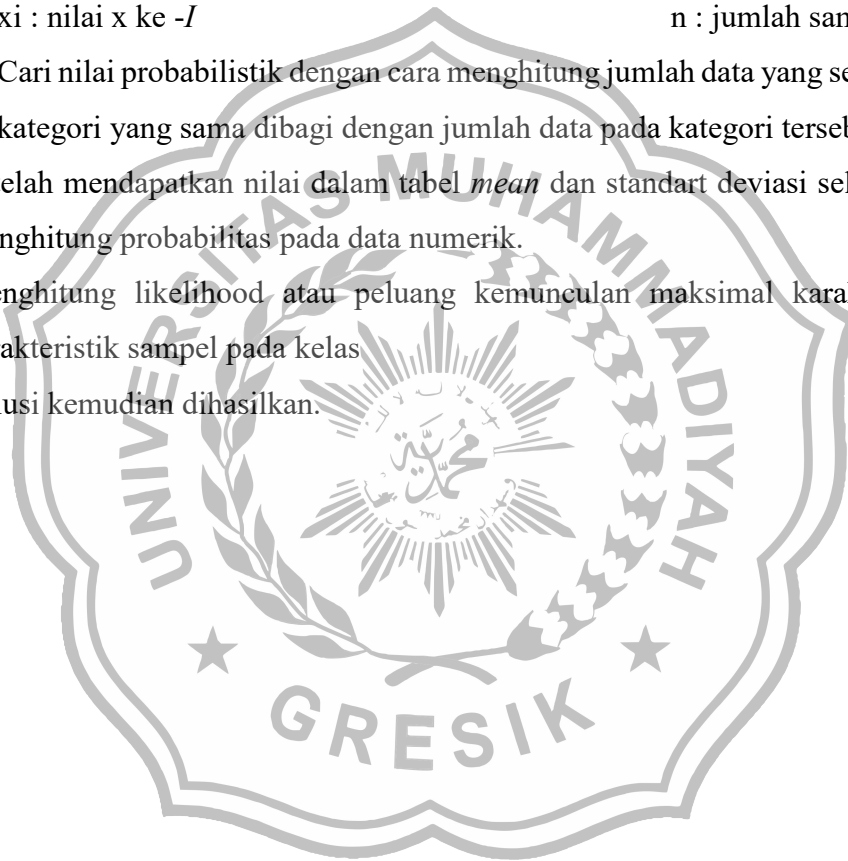
$\sigma$  : standar deviasi

$\mu$  : rata-rata hitung

$x_i$  : nilai  $x$  ke  $-I$

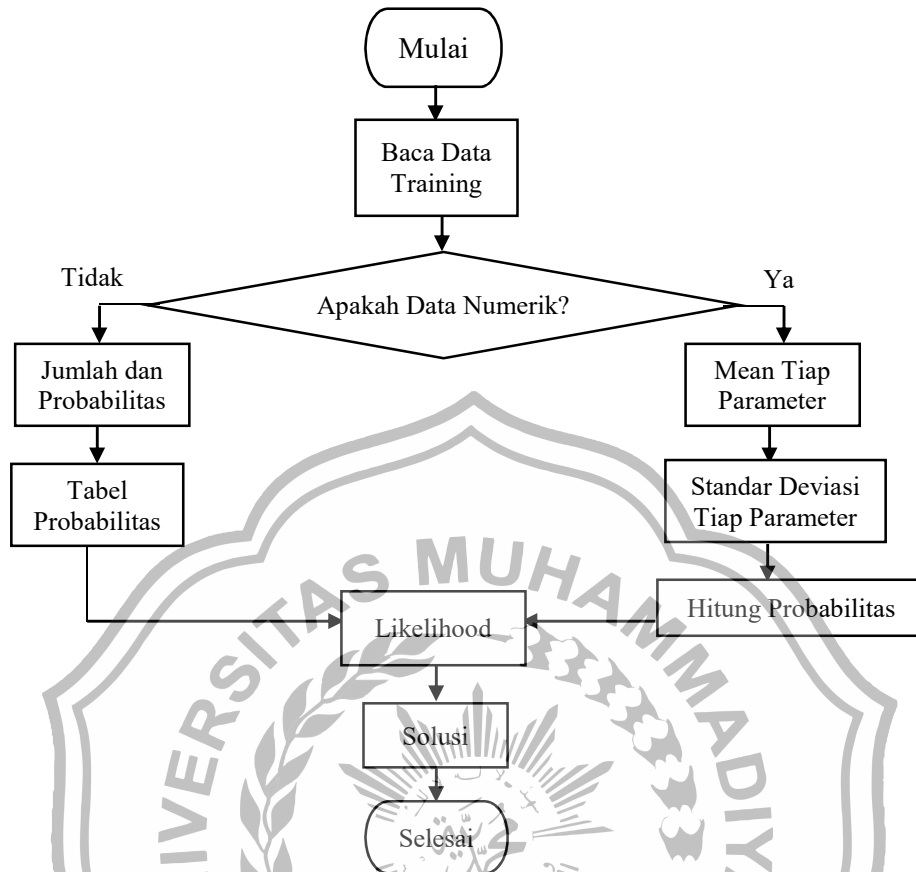
$n$  : jumlah sampel

- b. Cari nilai probabilitik dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.
3. Setelah mendapatkan nilai dalam tabel *mean* dan standart deviasi selanjutnya menghitung probabilitas pada data numerik.
4. Menghitung likelihood atau peluang kemunculan maksimal karakteristik-karakteristik sampel pada kelas
5. Solusi kemudian dihasilkan.





Penjelasan tersebut dapat dilihat dapat dilihat pada Gambar 2.1 sebagai berikut:



Gambar 2.1 Alur Metode Naïve Bayes

## 2.5 Perhitungan Evaluasi Klasifikasi

Untuk mengetahui apakah hasil klasifikasi yang didapatkan sudah akurat atau belum maka perlu dilakukan uji ketepatan hasil evaluasi. Untuk menghitung hasil evaluasi, dapat menggunakan perhitungan nilai akurasi. Perhitungan nilai akurasi dapat dilakukan dengan menggunakan persamaan 2.6 (Rodiysyah, 2013).

$$Akurasi = \frac{\text{jumlah klasifikasi benar}}{\text{jumlah data uji}} \times 100\% \quad (2.6)$$

Setiap data uji dilakukan prediksi menggunakan algoritma Naïve Bayes. Hasil perhitungan klasifikasi Naïve Bayes akan dibandingkan dengan hasil data *real* dan dihitung jumlah data yang tepat atau cocok dengan data *real*.

