

# Prediksi Tingkat Kelahiran Bayi di Kabupaten Bojonegoro dengan Menggunakan Algoritma Naive Bayes

Mariyatu Qibtiyah<sup>1</sup>, Nita Cahyani<sup>2</sup>

<sup>#</sup> *Statistika, Universitas Nadlatul Ulama Sunan Giri  
Jl. A. Yani No.10 Bojonegoro, 62115, Indonesia*

<sup>1</sup>[mariyatull15@gmail.com](mailto:mariyatull15@gmail.com), <sup>2</sup>[cahyani.nc@gmail.com](mailto:cahyani.nc@gmail.com)

*Abstract*— Perkembangan teknologi informasi pada zaman ini yang sangat pesat dimulai dari penemuan informasi baru pada big data dengan mencari suatu pola tertentu atau sering disebut dengan istilah Data Mining. Data mining merupakan proses dimana menggunakan teknik statistik, matematika, kecerdasan buatan, serta machine learning yang digunakan untuk mengekstraksi (proses pemisahan) dan mengidentifikasi informasi yang bermanfaat terkait dari berbagai database besar. Dengan berkembangnya teknologi informasi saat ini, terutama pada bidang data mining yang telah banyak digunakan dalam sistem informasi. Ada beberapa metode klasifikasi data mining yang sering digunakan untuk memprediksi atau meramalkan, seperti Algoritma Naive Bayes Classifier, Decision Tree, Neural Network, K-Nearest Neighbour, Artificial Neural Network, dan lain metode klasifikasi lain sebagainya. Metode yang digunakan pada penelitian ini menggunakan metode Algoritma Naive Bayes yang memanfaatkan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu ilmuwan yang memprediksi probabilitas di masa depan berdasarkan dengan pengalaman di masa sebelumnya atau masa lalu. Faktor-faktor yang mempengaruhi angka kelahiran sendiri yaitu faktor jenis kelamin laki-laki, jenis kelamin perempuan, dan pasangan usia subur. Angka kelahiran bayi sebagai indikator yang penting untuk mencerminkan keadaan derajat kesehatan di suatu masyarakat. Data yang digunakan dari tahun 1993 sampai dengan tahun 2022. Hasil analisis Algoritma Naive Bayes pada penelitian ini yaitu sebanyak 27 data diprediksi rendah dan 3 data diprediksi rendah pada tahun 2017, 2020, 2021 serta mendapatkan nilai akurasi (accuracy), Recall dan Precision sebesar 100%.

*Keywords*— Data Mining, Prediksi, Algoritma Naive Bayes

## I. PENDAHULUAN

Setiap tahunnya sebanyak 4,8 juta anak dilahirkan di Indonesia. Dengan total tingginya angka kelahiran ini menempatkan bangsa Indonesia pada urutan keempat setelah Cina, India, dan Amerika Serikat dengan populasi mencapai hampir 270 juta jiwa [1]. Sejak pelaksanaan program KB pada tahun 1970, Indonesia telah berhasil menurunkan angka kelahiran total dari 5,7 (1960) menjadi 2,45 anak per keluarga pada awal 2020. Salah satu faktor penyebab dari tingginya kelahiran di Indonesia rendahnya jumlah akseptor keluarga berencana di kalangan laki-laki, hanya 4,4% dibandingkan dengan cakupan perempuan usia subur yang telah menjadi akseptor KB mencapai 61,9%. Angka akseptor laki-laki di Indonesia ini jauh lebih rendah dibandingkan dengan negara tetangga seperti Filipina (24%), Malaysia (16,8%), dan Thailand (9%). Data ini mencerminkan

partisipasi laki-laki dalam program KB melalui penggunaan kontrasepsi belum dianggap umum di Indonesia [1].

Sedangkan rata-rata angka kelahiran bayi di Jawa Timur sendiri mencapai 500 ribu bayi per tahun. Dengan jumlah jiwa yang semakin banyak, maka diperlukan pangan, sandang, pangan, pendidikan dan lapangan pekerja, yang harus menjadi perhatian dari pemerintah. Selain angka kelahiran bayi yang tinggi, angka kematian ibu dan anak juga masih cukup tinggi. Setiap tahunnya sekitar 359 ibu meninggal per 100 ribu ibu melahirkan [2]. Salah satunya yaitu di kabupaten Bojonegoro, angka kelahiran bayi di kabupaten Bojonegoro untuk saat ini mengalami penurunan secara drastis. Faktor utama penyebab kelahiran bayi menurun di Kabupaten Bojonegoro adalah faktor kemiskinan. Hal tersebut relevan dengan data yang ada di Bojonegoro bahwa 1.169 penduduk sangat miskin, 4.516 penduduk miskin, dan 6.780 penduduk hampir miskin. Faktor yang paling mendasar penyebab angka kelahiran bayi menurun adalah angka kematian ibu

yang semakin tinggi. Selain itu, angka lain yang menjadi indikator mutu pelayanan kesehatan adalah angka kematian bayi [3].

Salah satu yang menjadi faktor penyebab terjadinya kenaikan jumlah penduduk di suatu negara adalah angka kelahiran ditambah dengan teknologi yang semakin canggih yang dapat membuat kelahiran seorang anak dapat dipercepat, hal tersebut apabila terjadi terus menerus maka yang akan terjadi adalah kepadatan penduduk [4]. Dalam mengatasi masalah tersebut perlu adanya analisis data untuk mengetahui prediksi pertambahan jumlah angka kelahiran bayi pada setiap tahun-nya. Menurut [5] anak yang lahir hidup adalah semua anak yang waktu lahir memperlihatkan tanda-tanda kehidupan, walaupun sesaat, seperti adanya detak jantung, bernafas, menangis, dan tanda-tanda kehidupan lainnya. Faktor-faktor yang mempengaruhi angka kelahiran sendiri menurut [6] ada 4 macam, jenis kelamin laki-laki, jenis kelamin perempuan, jumlah penduduk, dan pasangan usia subur. Angka kelahiran bayi sebagai indikator yang penting untuk mencerminkan keadaan derajat kesehatan di suatu masyarakat. Jumlah pertumbuhan penduduk yang semakin meningkat disuatu daerah membawa dampak dari sisi aspek kehidupan manusia [7].

Perkembangan teknologi informasi pada zaman ini yang sangat pesat dimulai dari penemuan informasi baru pada big data dengan mencari suatu pola tertentu atau sering disebut dengan istilah Data Mining [8]. Data mining merupakan proses dimana menggunakan teknik statistik, matematika, kecerdasan buatan, serta machine learning yang digunakan untuk mengekstraksi (proses pemisahan) dan mengidentifikasi informasi yang bermanfaat terkait dari berbagai database besar. Data mining terutama digunakan sebagai metode untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut Knowledge Discovery in Databases (KDD) [8]. Dalam pengaplikasiannya data mining memerlukan berbagai macam perangkat lunak analisis data untuk menemukan pola dan relasi (hubungan) data agar dapat digunakan untuk membuat prediksi atau peramalan [9]. Data Mining juga sering digunakan dalam penelitian untuk pola penjualan, penentuan jumlah produksi, keterkaitan pelanggan, dan lain sebagainya. Penelitian yang dilakukan untuk kepentingan pemerintah tentang kependudukan dalam program penekanan jumlah penduduk sangat jarang dilakukan. Pada hal ini penelitian tentang Data Mining dalam pemerintahan sangat diperlukan sebagai bahan informasi bagi pemerintahan [10].

Dengan berkembangnya teknologi informasi saat ini, terutama pada bidang data mining yang telah banyak digunakan dalam sistem informasi. Ada beberapa metode klasifikasi data mining yang sering digunakan untuk memprediksi atau meramalkan, seperti *Algoritma Naive Bayes Classifier*, *Decision Tree*, *Neural Network*, *K-Nearest Neighbour*, *Artificial Neural Network*, dan atau metode klasifikasi lain [11]. Definisi *Algoritma Naive Bayes* adalah sebuah metode atau analisis yang dilakukan dalam bentuk klasifikasi ataupun prediksi data baik berupa data training dan data testing [12]. *Algoritma Naive Bayes* ini memanfaatkan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu ilmuwan yang memprediksi probabilitas di masa depan

berdasarkan dengan pengalaman di masa sebelumnya atau masa lalu [13].

Salah satu penerapan *Algoritma Naive Bayes* adalah menentukan tingkat prediksi. Penentuan tingkat prediksi angka kelahiran bayi memungkinkan pemerintahan agar lebih mudah dalam menanggulangi angka kelahiran tinggi yang ada di kabupaten Bojonegoro [14]. Menurut [15] keuntungan menggunakan *Algoritma Naive Bayes* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (training data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. *Algoritma Naive Bayes* sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan. *Algoritma Naive Bayes* sendiri juga mudah di implementasikan dan mampu mendapatkan hasil baik dalam banyak kasus [16].

Tujuan penelitian ini untuk mengetahui statistik deskriptif angka kelahiran bayi beserta variabel yang diduga mempengaruhinya, untuk mengetahui hasil prediksi tingkat kelahiran bayi di Kabupaten Bojonegoro menggunakan *Algoritma Naive Bayes*. Manfaat dari penelitian ini diharapkan bisa membantu Pemerintah Kabupaten Bojonegoro dalam mengatasi masalah kependudukan di kabupaten Bojonegoro serta dapat memberikan kemudahan dalam memprediksi tingkat kelahiran pada tahun selanjutnya.

Pada penelitian ini, dilakukan analisis deskriptif statistik dan penambahan variabel pasangan usia subur sebagai pembaharuan dalam penelitian tingkat kelahiran di Kabupaten Bojonegoro. Dengan demikian perlu usulan penelitian yang berjudul “Prediksi Tingkat Kelahiran Bayi di Kabupaten Bojonegoro dengan Menggunakan *Algoritma Naive Bayes*” dengan harapan untuk memprediksi data angka kelahiran bayi selama kurun waktu 30 tahun sebagai bahan evaluasi dan analisa untuk mendukung strategi prediksi pada tahun berikutnya dengan tujuan untuk mengetahui dan menjelaskan faktor-faktor yang mempengaruhi tingkat kelahiran lebih mudah diperhitungkan kedepannya. mudah untuk mengikuti aturan format halaman paper ini adalah menggunakan format dalam dokumen ini. Simpanlah file ini dengan nama lainnya, lalu ketikkan isi makalah anda ke dalamnya.

## II. METODOLOGI PENELITIAN

### A. Sumber Data

Digunakan pendekatan kuantitatif dalam penelitian ini. Metode statistik yang diusulkan adalah statistik deskriptif dan *Algoritma Naive Bayes*. Dalam analisis data dilakukan dengan bantuan software Rapidminer. Pada penelitian ini, digunakan data sekunder berupa data hasil publikasi Kabupaten Bojonegoro data angka dari tahun 1993 sampai dengan tahun 2022

### B. Variabel Penelitian

Penelitian ini terdiri dari dua variabel, yaitu variabel terikat dan variabel bebas. Variabel terikat berupa angka kelahiran bayi dan variabel bebas berupa jenis kelamin laki-laki, jenis kelamin perempuan, dan pasangan usia subur. Berikut rincian variabel yang disajikan pada Tabel 1 berikut.

**Tabel 1. Variabel Penelitian**

Variabel	Nama Variabel	Skala Pengukuran
Y	Angka Kelahiran Bayi	Nominal
X <sub>1</sub>	Jumlah Jenis Kelamin Laki-Laki	Rasio
X <sub>2</sub>	Jumlah Jenis Kelamin Perempuan	Rasio
X <sub>3</sub>	Jumlah Pasangan Usia Subur (PUS)	Rasio

**C. Analisis Data**

Teknik analisa data adalah sebuah kegiatan berupa menganalisa suatu penelitian dengan mengontrol seluruh data yang berasal dari instrumen penelitian, seperti dokumen, catatan, rekaman dan lain-lain. Tujuannya adalah untuk memahami data dan memperoleh kesimpulan dengan lebih mudah. Teknik analisis data yang digunakan pada penelitian ini adalah sebagai berikut :

- 1) Mengumpulkan data yang akan digunakan untuk penelitian.
- 2) Menganalisis data secara deskriptif.
- 3) Mengubah skala variabel dependen (Y) yang awalnya berskala rasio menjadi skala nominal (tinggi dan rendah) dengan menggunakan rumus CBR :

$$CBR = B/P \times 1.000$$

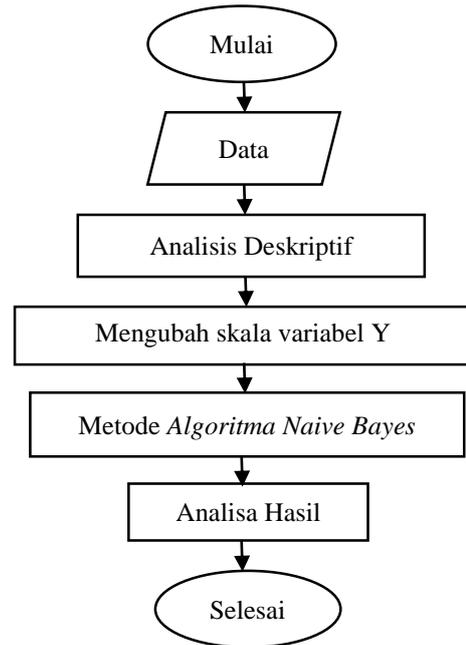
Dengan keterangan :

B = Banyaknya bayi yang lahir pada tahun tertentu  
 P = Jumlah penduduk pertengahan tahun  
 1.000 = Konstanta

- 4) Pembuatan model prediksi dengan menggunakan *Algoritma Naive Bayes*, dengan tahapan-tahapan sebagai berikut :
  - a) Menghitung probabilitas prior pada, yaitu probabilitas angka kelahiran tinggi dan rendah.
  - b) Menghitung probabilitas setiap variabel erhadap masing-masing kelas.
  - c) Menghitung perkalian probabilitas prior dengan atribut pada masing-masing kelas.
- 5) Mencari nilai probabilitas maksimal pada masing-masing kelas.
- 6) Menghitung ketepatan klasifikasi / prediksi dengan cara menghitung jumlah error (APER), akurasi, sensitivity dan specificity dalam proses pengklasifikasian.

**D. Diagram Alir**

Diagram alir pada penelitian ini ditunjukkan oleh Gambar 1 berikut ini.



Gambar 1. Diagram Alir Penelitian

**III. HASIL DAN PEMBAHASAN**

Hasil statistik deskriptif menunjukkan variabel tingkat kelahiran dan variabel yang diduga mempengaruhinya di Kabupaten Bojonegoro.

**A. Statistik Deskriptif**

Statistik Deskriptif merupakan proses menganalisa data atau dengan kata lain yaitu kita melaksanakan tindakan maupun perlakuan terhadap data untuk memperoleh tujuan tertentu baik berupa gambaran atas data maupun berupa kesimpulan terhadap kondisi atau kejadian di mana data tersebut diambil. Berikut hasil analisa statistik dekritif dari penelitian ini adalah :

**Tabel 2. Statistik Deskriptif Jenis Kelamin Laki-Laki**

Jenis Kelamin Laki-Laki	Nilai
Mean	658762.90
Standard Deviation	81781.467
Minimum	546500
Maximum	889600

Berdasarkan tabel diatas, nilai rata-rata jenis kelamin laki-laki di Kabupaten Bojonegoro adalah sebesar 658762.90 dengan standar deviasi 81781.467 yang menunjukkan bahwa jenis kelamin laki-laki di populasi yang diamati relatif homogen. Jenis kelamin laki-laki tertinggi di Kabupaten Bojonegoro pada tahun 2002 sebesar 889600. Di sisi lain, jenis kelamin laki-laki terendah di Kabupaten Bojonegoro yaitu pada tahun 2008 sebesar 546500.

Statistik Deskriptif dari Variabel Jenis Kelamin Perempuan di Kabupaten Bojonegoro dapat dilihat pada Tabel 3.

**Tabel 3.** Statistik Deskriptif Jenis Kelamin Perempuan

Jenis Kelamin Perempuan	Nilai
Mean	679135.23
Standard Deviation	94575.592
Minimum	557545
Maximum	976270

Berdasarkan tabel diatas, nilai rata-rata jenis kelamin perempuan di Kabupaten Bojonegoro adalah sebesar 679135.23 dengan standar deviasi 94575.592 yang menunjukkan bahwa jenis kelamin perempuan di populasi yang diamati relatif homogen. Jenis kelamin perempuan tertinggi di Kabupaten Bojonegoro pada tahun 2002 sebesar 976270. Di sisi lain, jenis kelamin perempuan terendah di Kabupaten Bojonegoro yaitu pada tahun 2004 sebesar 557545.

Statistik Deskriptif dari Variabel Pasangan Usia Subur (PUS) di Kabupaten Bojonegoro dapat dilihat pada Tabel 4.

**Tabel 4.** Statistik Deskriptif Pasangan Usia Subur (PUS)

Pasangan Usia Subur	Nilai
Mean	268361.27
Standard Deviation	23212.850
Minimum	206438
Maximum	296703

Berdasarkan tabel diatas, nilai rata-rata PUS di Kabupaten Bojonegoro adalah sebesar 268361.27 dengan standar deviasi 23212.850 yang menunjukkan bahwa PUS di populasi yang diamati relatif homogen. PUS tertinggi di Kabupaten Bojonegoro pada tahun 2017 sebesar 296703. Di sisi lain, PUS terendah di Kabupaten Bojonegoro yaitu pada tahun 2020 sebesar 206438.

Statistik Deskriptif dari Variabel Angka Kelahiran di Kabupaten Bojonegoro dapat dilihat pada Tabel 5 berikut ini.

**Tabel 5.** Statistik Deskriptif Angka Kelahiran

Angka Kelahiran	Nilai
Mean	8508.27
Standard Deviation	6199.002
Minimum	2665
Maximum	24686

Berdasarkan tabel diatas, nilai rata-rata Angka Kelahiran di Kabupaten Bojonegoro adalah sebesar 8508.27 dengan standar deviasi 6199.002 yang menunjukkan bahwa angka kelahiran di populasi yang diamati relatif homogen. Angka Kelahiran tertinggi di Kabupaten Bojonegoro pada tahun 2020 sebesar 24686. Di sisi lain, Angka Kelahiran terendah di Kabupaten Bojonegoro yaitu pada tahun 2004 sebesar 2665.

### B. Preprocessing Tingkat Kelahiran

Langkah pertama sebelum dilakukan prediksi dengan *Algoritma Naive Bayes* adalah preprocessing atau mengubah variabel prediktor menjadi bentuk kategorik dengan perhitungan tingkat angka kelahiran menggunakan rumus perhitungan CBR (*Crude Birth Rate*) yang digunakan untuk

menentukan kategori tingkat kelahiran tinggi dan rendahnya, rumus dari CBR adalah sebagai berikut :

$$CBR = \frac{B}{P} \times 1.000$$

**Tabel 6.** Perhitungan Rumus CBR

Tahun	Skala Data	Rumus CBR	Kategori
2000	Rasio 3780	$= \frac{3780}{728375} \times 1.000$ $= 0,0051 \times 1.000$ $= 5.1$	Hasil nilai CBR pada tahun 2000 adalah sebesar 5.1 < 20, maka tingkat kelahirannya “rendah”

dst

Setelah menghitung nilai CBR dari tahun 1993 sampai dengan tahun 2022. Langkah selanjutnya adalah menghitung nilai probabilitas prior untuk setiap kelas pada variabel data penelitian.

### C. Prediksi dan Tingkat Akurasi yang dihasilkan Algoritma Naive Bayes

#### 1. Menghitung Probabilitas Prior

##### 1.1. Menghitung Nilai P(Ci) dari Class (label) Keterangan P(Ci)

$$P(\text{Tingkat Kelahiran} = \text{tinggi}) = 3/30 = 0.1$$

$$P(\text{Tingkat Kelahiran} = \text{rendah}) = 27/30 = 0.9$$

Dari hasil diatas diperoleh hasil probabilitas hipotesis (P(Y)) yaitu untuk data kelas tingkat kelahiran tinggi didapat probabilitasnya sebesar 0.1 dan untuk kelas tingkat kelahiran rendah yaitu sebesar 0.9.

##### 1.2. Menghitung Nilai Mean

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

Variabel : Jenis Kelamin Laki-laki, Keterangan : Tinggi

$$\mu = \frac{620705 + 612155 + 673766}{3} = 635542$$

Variabel : Jenis Kelamin Laki-laki, Keterangan : Rendah

$$\mu = \frac{724060 + \dots + 677476}{27} = 661343$$

Variabel : Jenis Kelamin Perempuan, Keterangan : Tinggi

$$\mu = \frac{634378 + 636962 + 680760}{3} = 650700$$

Variabel : Jenis Kelamin Perempuan, Keterangan: Rendah

$$\mu = \frac{713001 + \dots + 673174}{27} = 682294.7037$$

Variabel : PUS, Keterangan : Tinggi

$$\mu = \frac{295669 + 206438 + 215892}{3} = 239333$$

Variabel : PUS, Keterangan : Rendah

$$\mu = \frac{287485 + \dots + 226252}{27} = 271586.6296$$

Variabel : Angka Kelahiran, Keterangan : Tinggi

$$\mu = \frac{23288 + 24686 + 23350}{3} = 23774$$

Variabel : Angka Kelahiran, Keterangan : Rendah

$$\mu = \frac{7366 + \dots + 12870}{27} = 6812$$

1.3. Menghitung Nilai Standar Deviasi

$$\sigma = \frac{\sqrt{\sum_i^n (x_i - \mu)^2}}{n-1}$$

Variabel : Jenis Kelamin Laki-laki, Keterangan : Tinggi

$$\sigma = \frac{\sqrt{(620705-635542)^2 + \dots + (612155-635542)^2}}{3-1} = 33377$$

Variabel : Jenis Kelamin Laki-laki, Keterangan : Rendah

$$\sigma = \frac{\sqrt{(592842-661343)^2 + \dots + (677476-661343)^2}}{27-1} = 85469$$

Variabel : Jenis Kelamin Perempuan, Keterangan : Tinggi

$$\sigma = \frac{\sqrt{(634378-650700)^2 + \dots + (673766-650700)^2}}{3-1} = 26064$$

Variabel : Jenis Kelamin Perempuan, Keterangan : Rendah

$$\sigma = \frac{\sqrt{(605523-682294)^2 + \dots + (673174-682294)^2}}{27-1} = 88096$$

Variabel : PUS, Keterangan : Tinggi

$$\sigma = \frac{\sqrt{(295669-239333)^2 + \dots + (215892-239333)^2}}{3-1} = 49016$$

Variabel : PUS, Keterangan : Rendah

$$\sigma = \frac{\sqrt{(258828-271586)^2 + \dots + (226252-271586)^2}}{27-1} = 17554$$

Variabel : Angka Kelahiran, Keterangan : Tinggi

$$\sigma = \frac{\sqrt{(23288-23774)^2 + \dots + (23350-23774)^2}}{3-1} = 789.85$$

Variabel : Angka Kelahiran, Keterangan : Rendah

$$\sigma = \frac{\sqrt{(3370-6812)^2 + \dots + (12870-6812)^2}}{27-1} = 3596$$

2. Menghitung Probabilitas Setiap Variabel Terhadap Masing-Masing Kelas.

Kemudian setelah mendapatkan nilai mean dan standar deviasi dari setiap variabel independen, maka dilanjutkan dengan menghitung probabilitas dari setiap variabel independen  $P(X|Y)$  di setiap variabel dengan menggunakan rumus sebagai berikut :

$$P(X_i = x_i | Y = y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} \exp \frac{-(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}$$

Dimana  $X_i$  adalah variabel independen,  $x_i$  adalah variabel independen berdasarkan data,  $Y$  adalah variabel dependen dengan kategori tinggi dan rendah ( $y_j$ ). Dari analisis ini akan diperoleh dari setiap variabel independen dan akan didapatkan hasil nilai probabilitas tingkat kelahiran tinggi dan rendah. Dibawah ini adalah contoh perhitungan dari tahun 2014.

- Tahun 2014

Variabel : Jenis Kelamin Laki-laki, Keterangan : Tinggi

$$= \frac{1}{\sqrt{2 \cdot 3,14 \cdot 33377}} \exp \frac{-(730879-635542)^2}{2 \cdot 33377^2} = 0.129$$

Variabel : Jenis Kelamin Laki-laki, Keterangan : Rendah

$$= \frac{1}{\sqrt{2 \cdot 3,14 \cdot 85469}} \exp \frac{-(730879-661343)^2}{2 \cdot 85469^2} = 0.019$$

Variabel : Jenis Kelamin Perempuan, Keterangan : Tinggi

$$= \frac{1}{\sqrt{2 \cdot 3,14 \cdot 26064}} \exp \frac{-(722164-650700)^2}{2 \cdot 26064^2} = 0.0622$$

Variabel : Jenis Kelamin Perempuan, Keterangan : Rendah

$$= \frac{1}{\sqrt{2 \cdot 3,14 \cdot 88096}} \exp \frac{-(722164-682294,7037)^2}{2 \cdot 88096^2} = 0.0015$$

Variabel : PUS, Keterangan : Tinggi

$$= \frac{1}{\sqrt{2 \cdot 3,14 \cdot 49016}} \exp \frac{-(291744-239333)^2}{2 \cdot 49016^2} = 0.0032$$

Variabel : PUS, Keterangan : Rendah

$$= \frac{1}{\sqrt{2 \cdot 3,14 \cdot 17554}} \exp \frac{-(291744-271586,629)^2}{2 \cdot 17554^2} = 0.006$$

Variabel : Angka Kelahiran, Keterangan : Tinggi

$$= \frac{1}{\sqrt{2 \cdot 3,14 \cdot 789,85}} \exp \frac{-(10202-23774)^2}{2 \cdot 789,85^2} = 0.0103$$

Variabel : Angka Kelahiran, Keterangan : Rendah

$$= \frac{1}{\sqrt{2 \cdot 3,14 \cdot 3596}} \exp \frac{-(10202-6812)^2}{2 \cdot 3596^2} = 1.846$$

Dst sampai dengan tahun 2022.

3. Menghitung Perkalian Probabilitas Prior Dengan Atribut Pada masing-Masing Kelas

-  $P(X|Ci)$  Tahun 2014

$P(X | \text{Keterangan} = \text{"Tinggi"})$

$$= 0.129 \times 0.0622 \times 0.0032 \times 0.0103 = 2.6446E-7$$

$P(X | \text{Keterangan} = \text{"Rendah"})$

$$= 0.019 \times 0.0015 \times 0.006 \times 1.846 = 3.156E-6$$

-  $P(X|Ci)$  Tahun 2015

$P(X | \text{Keterangan} = \text{"Tinggi"})$

$$= 0.0255 \times 0.106 \times 0.0024 \times 0.0116 = 7.5251E-8$$

$P(X | \text{Keterangan} = \text{"Rendah"})$

$$= 0.0019 \times 0.0015 \times 0.034 \times 6.43 = 6.230E-6$$

Dst sampai dengan tahun 2022.

4. Mencari Nilai Probabilitas Maksimal

-  $P(X|Ci) \cdot P(Ci)$  Tahun 2014

$P(X | \text{Keterangan} = \text{"Tinggi"}) \cdot P(\text{Keterangan} = \text{"Tinggi"})$

$$= 2.6446E-7 \times 0.1 = 2.6446E-8$$

-  $P(X | \text{Keterangan} = \text{"Rendah"}) \cdot P(\text{Keterangan} = \text{"Rendah"})$

$$= 3.156E-6 \times 0.9 = 2.8404E-6$$

Berdasarkan perhitungan akhir dengan mengalikan nilai probabilitas, dapat diketahui bahwa nilai  $P(X | \text{Keterangan} = \text{"Rendah"})$  lebih besar dari  $P(X | \text{Keterangan} = \text{"Tinggi"})$ . Maka kelas (label) pada data tahun 2014 adalah rendah.

-  $P(X|Ci) \cdot P(Ci)$  Tahun 2015

$P(X | \text{Keterangan} = \text{"Tinggi"}) \cdot P(\text{Keterangan} = \text{"Tinggi"})$

$$= 7.5251E-8 \times 0.1 = 7.5251E-9$$

-  $P(X | \text{Keterangan} = \text{"Rendah"}) \cdot P(\text{Keterangan} = \text{"Rendah"})$

$$= 6.230E-6 \times 0.9 = 5.607E-6$$

Berdasarkan perhitungan akhir dengan mengalikan nilai probabilitas, dapat diketahui bahwa nilai  $P(X | \text{Keterangan} = \text{"Rendah"})$  lebih besar dari  $P(X | \text{Keterangan} = \text{"Tinggi"})$ . Maka kelas (label) pada data tahun 2015 adalah rendah.

Berikut hasil perhitungan dengan algoritma naive bayes dari tahun 1993 sampai dengan tahun 2022 adalah seperti Tabel 7 sebagai berikut :

**Tabel 7.** Hasil Algoritma *Naive Bayes*

Tahun	Tingkat Kelahiran	Tahun	Tingkat Kelahiran
1993	Rendah	2008	Rendah
1994	Rendah	2009	Rendah
1995	Rendah	2010	Rendah
1996	Rendah	2011	Rendah
1997	Rendah	2012	Rendah
1998	Rendah	2013	Rendah
1999	Rendah	2014	Rendah
2000	Rendah	2015	Rendah
2001	Rendah	2016	Rendah
2002	Rendah	2017	Tinggi
2003	Rendah	2018	Rendah
2004	Rendah	2019	Rendah
2005	Rendah	2020	Tinggi
2006	Rendah	2021	Tinggi
2007	Rendah	2022	Rendah

Setelah diketahui hasil perhitungan menggunakan *Algoritma Naive Bayes*, maka didapatkan nilai *confusion matrix* seperti Tabel 8 dibawah ini :

**Tabel 8.** *Confusion Matrix*

	Rendah	Tinggi
Rendah	27	0
Tinggi	0	3

Bedasarkan pada hasil *confusion matrix* diatas dapat diketahui nilai TP (*true positive*) yaitu prediksi rendah yang diprediksi benar sebesar 27, nilai TN (*true negative*) yaitu prediksi tinggi diprediksi benar sebesar 3, nilai FN (*false negative*) yaitu prediksi tinggi yang diprediksi salah sebesar 0 dan FP (*false positive*) yaitu prediksi rendah yang diprediksi salah sebesar 0. Dari nilai *confusion matrix* diatas maka dapat digunakan untuk menghitung nilai *accuracy*, *recall* dan *precision* seperti dibawah ini.

Adapun hasil nilai *accuracy* (yaitu tingkat kedekatan nilai prediksi dengan nilai aktual atau sebenarnya) dihitung menggunakan rumus sebagai berikut :

$$Accuracy = \frac{(TP+TN)}{TP+FP+TN+FN} * 100\%$$

$$Accuracy = \frac{27+3}{27+3+0+0} * 100\%$$

$$Accuracy = \frac{30}{30} * 100\%$$

$$Accuracy = 100\%$$

Adapun hasil nilai *recall* (yaitu rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif) dihitung menggunakan rumus sebagai berikut :

$$Recall = \frac{TP}{TP+FN} * 100\%$$

$$Recall = \frac{27}{27+0} * 100\%$$

$$Recall = \frac{27}{27} * 100\%$$

$$Recall = 100\%$$

Adapun hasil nilai *precision* (yaitu rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif) dihitung menggunakan rumus sebagai berikut :

$$Precision = \frac{TP}{TP+FP} * 100\%$$

$$Precision = \frac{27}{27+0} * 100\%$$

$$Precision = \frac{27}{27} * 100\%$$

$$Precision = 100\%$$

#### IV. SIMPULAN

Dari data angka kelahiran di Kabupaten Bojonegoro dari tahun 1993 sampai dengan 2022 didapatkan nilai rata-rata jenis kelamin laki-laki di Kabupaten Bojonegoro adalah sebesar 658762.90 dengan standar deviasi 81781.467 yang menunjukkan bahwa jenis kelamin laki-laki di populasi yang diamati relatif homogen. Jenis kelamin laki-laki tertinggi di Kabupaten Bojonegoro pada tahun 2002 sebesar 889600. Nilai rata-rata jenis kelamin perempuan di Kabupaten Bojonegoro adalah sebesar 679135.23 dengan standar deviasi 94575.592 yang menunjukkan bahwa jenis kelamin perempuan di populasi yang diamati relatif homogen. Jenis kelamin perempuan tertinggi di Kabupaten Bojonegoro pada tahun 2002 sebesar 976270. Di sisi lain, jenis kelamin perempuan terendah di Kabupaten Bojonegoro yaitu pada tahun 2004 sebesar 557545. Nilai rata-rata PUS di Kabupaten Bojonegoro adalah sebesar 268361.27 dengan standar deviasi 23212.850 yang menunjukkan bahwa PUS di populasi yang diamati relatif homogen. PUS tertinggi di Kabupaten Bojonegoro pada tahun 2017 sebesar 296703. Di sisi lain, PUS terendah di Kabupaten Bojonegoro yaitu pada tahun 2020 sebesar 206438. Nilai rata-rata Angka Kelahiran di Kabupaten Bojonegoro adalah sebesar 8508.27 dengan standar deviasi 6199.002 yang menunjukkan bahwa angka kelahiran di populasi yang diamati relatif homogen. Angka Kelahiran tertinggi di Kabupaten Bojonegoro pada tahun 2020 sebesar 24686. Di sisi lain, Angka Kelahiran terendah di Kabupaten Bojonegoro yaitu pada tahun 2004 sebesar 2665. Hasil analisis *Algoritma Naive Bayes* untuk prediksi tingkat kelahiran bayi di Kabupaten Bojonegoro pada tahun 1993 sampai dengan tahun 2022 didapatkan hasil yaitu pada tahun 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2018, 2019, 2022 dan tingkat kelahiran bayi diprediksi tinggi pada tahun 2017, 2020, dan 2021. Saran untuk penelitian ini untuk penelitian berikutnya dapat digunakan data pengamatan yang berskala per kecamatan sehingga dapat memberikan informasi yang lebih akurat terkait prediksi kelahiran bayi yang ada di Kabupaten Bojonegoro, untuk penelitian berikutnya bisa menggunakan faktor pendukung yang lebih banyak dan lebih efisien untuk menentukan seberapa besar berpengaruh terhadap tingkat kelahiran bayi, untuk penelitian berikutnya, dapat menggunakan metode terbaru yang bisa digunakan pada analisis ini dan menambahkan data yang lebih banyak untuk mendapatkan hasil akurasi yang lebih tinggi.

DAFTAR PUSTAKA

- [1] William, “Angka Kelahiran di Indonesia Masih Tinggi,” <https://theconversation.com/amp/angka-kelahiran-di-indonesia-masih-tinggi-mengapa-mayoritas-laki-laki-ogah-ikut-kb-146577>, 2020. .
- [2] R. Jajeli, “Setiap Tahun 500 Ribu Dilahirkan di Jatim,” <https://news.detik.com/berita-jawa-timur/d-3493677/setiap-tahun-500-ribu-bayi-dilahirkan-di-jatim-program-kb-gagal>, 2018. .
- [3] Kominfo, “Melalui Program Sehati, Bupati Bojonegoro Pastikan Pemerataan Kesehatan Ibu dan Bayi,” <https://kominfo.jatimprov.go.id/berita/melalui-program-sehati-bupati-bojonegoro-pastikan-pemerataan-kesehatan-ibu-dan-bayi>, 2022. .
- [4] N. I. Parihah, “Prediksi Angka Kelahiran Bayi Pada Desa Tridaya Sakti Dengan Menggunakan Algoritma Naive Bayes,” vol. 1, no. 2, pp. 77–88, 2020.
- [5] BPS, “Penjelasan Kependudukan,” <https://www.bps.go.id/subject/12/kependudukan.htm?subjekViewTab2>, 2022. .
- [6] D. F. I. L. A. H. M. dan J. Roy, “Faktor-faktor yang Mempengaruhi Jumlah Kelahiran,” vol. 14, no. 1, pp. 8–19, 2018.
- [7] R. K. Yohanni Syahra, Ismawardi Santoso, “Implementasi Data Mining Untuk Prediksi Angka Kelahiran Bayi pada Desa Sibolangit Menggunakan Multi Regresi,” no. 1, pp. 687–690, 2019.
- [8] M. Idris, “Implementasi Data Mining dengan Algoritma Naive Bayes untuk Memprediksi Angka Kelahiran,” vol. 7, pp. 421–428, 2019.
- [9] H. A. D. Rani, “Sistem Prediksi Kondisi Kelahiran Bayi Menggunakan Klasifikasi Naive Bayes,” vol. 3, No.2, pp. 1–9, 2020.
- [10] D. P. dan W. P. M. Syifa Sintia Al Khautsar, “Algoritma Naive Bayes Untuk Memprediksi Kredit Macet Pada Koperasi Simpan Pinjam,” vol. 4, No.2, no. ISSN: 2460-4801, pp. 2447–6645, 2018.
- [11] J. T. W. Ginanjar Abdurrahman, “Analisis Klasifikasi Kelahiran Caesar Menggunakan Algoritma Naive Bayes,” vol. 3, no. 1, pp. 57–63, 2020.
- [12] F. Wijaya, “Implementasi Algoritma Naive Bayes dalam Klasifikasi Produk Ban Terlaris pada PT.Mitra Mekar Mandiri,” 2018.
- [13] R. Retnosari, “Analisis Kelayakan Kredit Usaha Mikro Berjalan pada Perbankan dengan Metode Naive Bayes,” vol. 8, no. p-ISSN : 2406-7733, 2021.
- [14] N. Cahyani, K. Fithriasari, and N. Iriawan, “On the Comparison of Deep Learning Neural Network and Binary Logistic Regression for Classifying the Acceptance Status of Bidikmisi Scholarship Applicants in East Java Methods,” pp. 83–90, 2018.
- [15] E. Manalu, “Penerapan Algoritma Naive Bayes untuk Memprediksi Jumlah Produksi Barang berdasarkan Data Persediaan dan Jumlah Pemesanan pada CV.Papa dan Mama Pastries,” vol. 1, No.2, no. ISSN 2088-3943, pp. 1–6, 2019.
- [16] A. Desiani, “Implementasi Algoritma Naive Bayes dan Support Vector Machine ( SVM ) Pada Klasifikasi Penyakit Kardiovaskular,” *Tek. Elektro dan Komputasi*, vol. 4, No.2, no. ISSN : 2685-1814, pp. 207–214, 2022.