

PREDICTION OF BABY BIRTH RATE USING NAÏVE BAYES CLASSIFICATION ALGORITHM IN RANDAU VILLAGE

Magda Kitty Hartono^{*1}, Hendry²

^{1,2}Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana, Indonesia

Email: ¹672018195@student.uksw.edu, ²hendry@uksw.edu

(Naskah masuk: 27 April 2022, Revisi : 15 Mei 2022, diterbitkan: 20 Agustus 2022)

Abstract

The birth rate is one of the factors increasing the rate of population growth. Birth or fertility can affect the population, getting more lower of birth rate in an area, the higher the life expectancy in that area. The number of births in Randau Jekak Village is increasing every year. The Naïve Bayes algorithm can be used to predict the future births rate because this algorithm is a simple algorithm and uses a lot of data as information in collecting data groups, and with data mining techniques it can see the predictive pattern of each variable and test. The testing data and the training data are expected to help the Village Institution or Office in Randau Jekak to suppressing the rate of population growth which increases every year. The results of this study can be concluded that the Naïve Bayes Algorithm is suitable for predicting the birth rate of babies in Randau Jekak Village with the classification technique, the birth rate in Randau Jekak Village in 2021 is High. The results of this data will be very useful for the Randau Jekak Village office in suppressing the rate of population growth in the coming year.

Keywords: Birth Rate, Data Mining, Naïve Bayes, Prediction.

PREDIKSI ANGKA KELAHIRAN BAYI DENGAN ALGORITMA KLASIFIKASI NAÏVE BAYES DI DESA RANDAU

Abstrak

Angka kelahiran merupakan suatu faktor dari meningkatnya angka pertumbuhan penduduk. Kelahiran atau fertilitas dapat mempengaruhi jumlah penduduk. Tingkat kelahiran di Desa Randau Jekak setiap tahun semakin tinggi, semakin tinggi angka kelahiran di suatu daerah maka semakin rendah pula usia harapan hidup di daerah tersebut. Algoritma Naïve Bayes merupakan suatu algoritma yang memungkinkan untuk mengklasifikasikan variabel tertentu dengan menggunakan metode statistik dan probabilitas dengan menggunakan data yang sangat banyak sebagai informasi dalam pengumpulan kelompok data serta dengan teknik data mining yang dapat melihat pola prediksi setiap variabel serta pengujian kepada data testing serta data training dapat digunakan untuk memprediksi angka kelahiran di masa yang akan datang karena algoritma ini termasuk ke dalam algoritma yang sederhana yang nantinya dapat membantu Lembaga atau Kantor Desa di Randau Jekak dalam menekan laju pertumbuhan penduduk yang meningkat setiap tahun. Atribut yang digunakan dalam penelitian ini adalah Tahun, Jenis Kelamin, Jumlah Penduduk, Angka Pernikahan, Angka Kelahiran, Jumlah Wanita Usia Subur serta Keterangan Angka Kelahiran. Dari hubungan antar atribut tersebut nantinya akan mendapatkan hasil klasifikasi tingkat angka kelahiran pada tahun tertentu dengan hasil tinggi atau rendah. Dari hasil penelitian ini dapat disimpulkan bahwa Algoritma Naïve Bayes cocok digunakan untuk memprediksi angka kelahiran bayi di Desa Randau Jekak dengan teknik klasifikasi maka tingkat angka kelahiran di Desa Randau Jekak pada tahun 2021 adalah “Tinggi”. Hasil dari data ini nantinya sangat berguna untuk pihak kantor Desa Randau Jekak dalam menekan laju pertumbuhan penduduk di tahun mendatang.

Kata kunci: Data Mining, Kelahiran, Naïve Bayes, Prediksi.

1. PENDAHULUAN

Manfaat dari perkembangan sekarang yang sangat cepat ini yaitu mencari pola dari informasi baru yang di dapat dari data yang dianggap sangat besar adalah istilah dari data mining [1]. Data mining

adalah suatu ilmu kecerdasan buatan untuk menggali pola-pola data yang dibutuhkan untuk diubah menjadi informasi baru yang berguna, data mining adalah suatu rangkaian proses yang melakukan teknik pembelajaran komputer untuk

menganalisis data secara otomatis [2]. Penelitian untuk menekan laju pertumbuhan penduduk masih jarang dilakukan, padahal penelitian seperti pengelolaan data penduduk atau angka pertumbuhan penduduk sangat penting untuk bahan informasi bagi pemerintah [3]. Dengan Teknik data mining penulis mengharapkan penelitian ini dapat membantu dalam memprediksi angka kelahiran di Desa Randau masa yang akan datang.

Kelahiran merupakan salah satu faktor naiknya jumlah penduduk [4]. Angka kelahiran yang tinggi dapat menyebabkan ledakan pertumbuhan penduduk, angka kelahiran adalah salah satu indikator penting dalam pelaksanaan pemerintah dalam pengendalian jumlah penduduk [5]. Daerah dengan jumlah penduduk yang tinggi tidak menjamin bahwa kemakmuran penduduk di daerah tersebut tinggi pula [6]. Pertumbuhan penduduk merupakan keseimbangan dinamis antara jumlah penduduk yang terus menerus meningkat, penduduk akan bertambah oleh jumlah kelahiran bayi (fertilitas) [7].

Selain kelahiran, faktor meningkat dan menurunnya pertumbuhan penduduk adalah dari kematian dan migrasi [8]. Tingginya angka kelahiran di Indonesia memerlukan perhatian khusus dari pemerintah seperti melaksanakan keluarga berencana dan pembangunan secara komprehensif [9]. Meningkatnya angka pertumbuhan penduduk di suatu daerah dapat berdampak pada masyarakat di daerah tersebut, dan jika laju pertumbuhan penduduk meningkat dari sisi kuantitas tetapi tidak meningkat secara kualitas maka hal ini dapat memicu atau menimbulkan masalah baru bagi pembangunan suatu daerah atau wilayah [10].

Algoritma Naïve Bayes adalah suatu teknik pengklasifikasian data dengan menggunakan salah satu cabang ilmu matematika yaitu statistik dan probabilitas yang dikemukakan oleh Thomas Bayes yaitu seorang ilmuwan Inggris, Thomas Bayes juga mengemukakan teorema Bayes dimana metode ini dapat melakukan prediksi untuk peluang di masa mendatang dengan berdasarkan pengalaman di masa sebelumnya. Algoritma Naïve Bayes ini dapat membantu melakukan prediksi angka kelahiran [11].

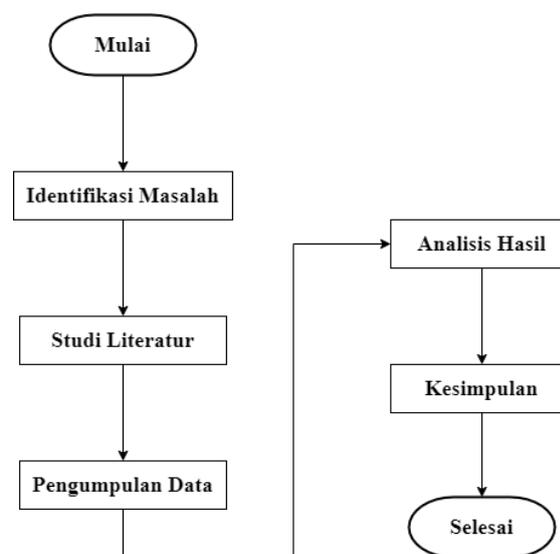
Untuk mengatasi tinggi rendahnya angka kelahiran bayi di Desa Randau Jekak, dibutuhkan suatu analisis untuk lembaga atau instansi Desa Randau untuk menekan laju pertumbuhan penduduk pada tahun 2022 ini yaitu dengan metode klasifikasi, karena metode ini dapat membedakan kelas data berdasarkan model, yang tujuannya agar dapat memperkirakan tingkat kelahiran di Desa Randau pada tahun 2022 berada di tingkat tinggi atau rendah di lihat dari objek yang label nya tidak diketahui, oleh karena itu algoritma naïve bayes dapat digunakan untuk memprediksi angka kelahiran di masa yang akan datang karena algoritma ini termasuk ke dalam algoritma yang sederhana dan dengan menggunakan data yang sangat banyak sebagai informasi dalam

pengumpulan kelompok data, dan dengan teknik data mining.

Tujuan dari penelitian ini yaitu agar dapat membantu lembaga seperti Kantor Desa Randau Jekak dalam pengambilan keputusan bagaimana nanti kedepannya dalam rangka menekan laju pertumbuhan di Desa tersebut.

2. METODE PENELITIAN

Langkah-langkah yang digunakan untuk mendukung pembuatan Penelitian “Prediksi Angka Kelahiran Bayi Dengan Algoritma Klasifikasi Naïve Bayes Di Desa Randau” ini dapat dilihat pada gambar 1 sebagai berikut :



Gambar 1. Metode penelitian

2.1. Identifikasi Masalah

Tahap identifikasi masalah adalah tahap awal untuk menentukan masalah yang berkaitan pada angka kelahiran. Berdasarkan permasalahan yang sudah dibahas dalam pendahuluan, untuk mengatasi permasalahan dalam melakukan prediksi angka kelahiran diperlukan suatu analisis di kantor desa untuk mengetahui apakah angka fertilitas di Desa Randau Jekak berada di tingkat tinggi atau rendah.

2.2. Studi Literatur

Studi literatur dilakukan untuk menemukan teori, konsep dan metode yang relevan dengan definisi masalah. Penulis melakukan pengumpulan studi literatur dengan mempelajari beberapa jurnal, buku, serta referensi dari berbagai media untuk mempelajari hal yang terkait dengan penelitian ini.

2.3. Pengumpulan Data

Setelah melakukan studi literatur, Penulis melakukan pengumpulan data yang di dapatkan dari data penduduk di desa Randau Jekak yang diambil dari Poskesdes Desa Randau dan Kantor Kepala

Desa. Data yang di diperoleh langsung dari Kantor Poskesdes Desa Randau Jekak dan Kantor Desa Randau Jekak melalui wawancara dengan pihak Poskesdes dan Kantor Desa Randau Jekak.

Data yang diperoleh kemudian diolah menggunakan algoritma Naive Bayes. Kemudian diuji dengan tools RapidMiner menggunakan Performance yang berfungsi sebagai validasi. Dalam penelitian ini data yang digunakan meliputi dari data training dan data testing, dimana data testing nantinya akan diolah dengan algoritma Naive Bayes untuk mendapatkan hasil klasifikasi.

2.3.1. Sumber Data

Sumber data yang digunakan dalam penelitian ini adalah data yang di dapatkan langsung dari data kependudukan di Desa Randau melalu Kantor Desa.

Adapun sampel data kependudukan di Desa Randau pada tahun 2011-2020 terdapat pada tabel berikut.

Tabel 1. Sampel Data Kependudukan Desa Randau Jekak Tahun 2011-2020

Tahun	L	P	JP	AP	AK	WUS
2011	690	620	1310	64	40	339
2012	1023	993	2016	60	61	390
2013	1.637	1.580	3.217	72	100	500
2014	1.872	1.769	3.641	68	110	490
2015	1.900	1.865	3.765	65	113	450
2016	2.001	1.995	3.996	74	69	511
2017	2.081	2.051	4.132	75	100	610
2018	2.218	2.107	4.325	69	97	655
2019	2.520	2.400	4.920	77	151	599
2020	2.620	2.580	5.200	89	80	662

Dengan L adalah Laki-laki, P adalah Perempuan, JP adalah jumlah penduduk, AP adalah angka pernikahan, AK adalah angka kelahiran dan WUS adalah Wanita Usia Subur.

2.4. Metode yang Digunakan

Metode yang akan digunakan untuk melakukan pengklasifikasian terhadap tingkat angka kelahiran di Desa Randau Jekak adalah dengan menggunakan teknik data mining untuk algoritma Naïve Bayes.

2.4.1. Data Mining

Data mining merupakan sebuah proses dalam memisahkan suatu data yang sangat besar (big data) dengan menggunakan sebuah algoritma tertentu dan teknik pengolahan data yang ada dalam bidang statistika, machine learning, dan sistem basis data. Data mining dapat menganalisis data dengan banyak persepektif yang nantinya akan mendapatkan sebuah kesimpulan menjadi informasi yang akan digunakan untuk mempermudah analisis data [12].

Definisi lain juga mengatakan bahwa data mining adalah sebuah teknik dalam mengumpulkan serta menggunakan data historis untuk menentukan keteraturan pola dan hubungan data di dalam pola yang sangat besar. Dari beberapa definisi di atas dapat

ditarik kesimpulan bahwa Data Mining merupakan suatu proses atau teknik pengumpulan dari data yang sangat besar dimana data tersebut nantinya akan di proses agar mendapatkan informasi-informasi yang berguna untuk proses analisis data.

Dari beberapa definisi yang telah disampaikan terdapat beberapa kelompok dalam data mining, sebagai berikut [13]:

1. Deskripsi

Dalam proses pendeskripsian biasanya Peneliti sering mencoba untuk menggambarkan tren atau pola tersembunyi dalam data.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, hanya saja variabel targetnya numerik bukan kategorik.

3. Prediksi

Prediksi hampir sama dengan estimasi dan klasifikasi, tetapi prediksi menunjukkan hasil yang belum terjadi atau akan terjadi di masa depan.

4. Klasifikasi

Dalam klasifikasi terdapat kategori misalnya kita ingin mengklasifikasikan tingkat penggunaan listrik rumah tangga yaitu penggunaanya tinggi, sedang atau rendah.

5. Klustering

Klustering lebih ke arah pengelompokkan data, pengamatan, terhadap kasus yang memiliki kemiripan dalam kelasnya.

6. Asosiasi

Asosiasi lebih ke arah mengidentifikasi hubungan antara berbagai peristiwa dalam satu waktu.

Data mining sebagai rangkaian proses, dapat dibagi menjadi beberapa tahapan-tahapan dalam data mining. Tahap-tahap tersebut dapat bersifat interaktif, terlibat langsung atau dengan perantara knowledge base:

1. Pembersihan Data (Data Cleaning)

Pembersihan data merupakan proses menghilangkan atau membersihkan noise dan data yang tidak relevan dan data yang tidak konsisten.

2. Integrasi Data (Data Integration)

Integrasi data ini merupakan proses menggabungkan data dari beberapa database yang kemudian digabungkan dalam sebuah database baru.

3. Seleksi Data (Data Selection)

Dalam tahap seleksi data, hanya data yang sesuai untuk dianalisis yang akan digunakan.

4. Transformasi Data (Data Transformation)

Data yang didapatkan yang sudah diubah dan digabungkan kedalam format yang sudah sesuai untuk di proses ke dalam data mining.

2.4.2. Algoritma Naïve Bayes

Algoritma Naïve Bayes merupakan algoritma yang memungkinkan untuk mengklasifikasikan variabel tertentu dengan menggunakan metode statistik dan probabilitas [14]. Naïve Bayes merupakan suatu metode yang menggunakan salah satu ilmu matematika yaitu teori probabilitas untuk

mencari peluang terbesar dari pengklasifikasian yang memungkinkan, dengan cara memeriksa frekuensi setiap pengklasifikasian dalam data latih (data training). Pengklasifikasian Naive Bayes mengasumsikan bahwa ada atau tidak adanya beberapa karakteristik dari sebuah kelas tidak ada hubungannya dengan karakteristik kelas lainnya. Persamaan untuk teorema Bayes dengan tipe data nominal adalah [15]:

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (1)$$

Keterangan :

X : Sampel data yang memiliki kelas (label) yang tidak diketahui

H : Hipotesa bahwa X adalah data kelas (label)

P(H) : Peluang dari hipotesa H

P(X) : Peluang dari data sampel yang diamati

P(X|H) : Peluang dari data sampel X bila diasumsikan bahwa hipotesa benar

Rumus Naive Bayes tipe data numerik [16]:

Distibusi Gaussian = $g(x, \mu, \sigma) =$

$$\frac{1}{\sqrt{2\pi} \cdot \sigma} \exp \frac{-(x-\mu)^2}{2\sigma^2} \quad (2)$$

$$\text{Mean} = \mu = \frac{\sum_{i=1}^n x_i}{n} \quad (3)$$

$$\text{Standar Deviasi} = \sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1}} \quad (4)$$

Dengan demikian Naive Bayes Classifier dapat didefinisikan sebagai metode klasifikasi yang berdasarkan rekayasa probabilitas dan teorema bayes bahwa setiap variabel keputusan atau parameter adalah bersifat bebas (independence) sehingga keberadaan setiap variabel tidak ada hubungannya dengan keberadaan atribut yang lain.

Keuntungan dalam penggunaan algoritma klasifikasi Naive Bayes ini adalah bahwa metode Naive Bayes ini hanya membutuhkan jumlah data training yang kecil untuk menentukan estimasi parameter yang diperlukan untuk proses klasifikasi. Terdapat 4 proses dalam tahapan proses Algoritma Naive Bayes ini, yaitu [17], [18] :

1. Menghitung jumlah kelas / label
2. Menghitung jumlah kasus per kelas
3. Kali kan semua variabel kelas.
4. Membandingkan hasil per kelas

2.5. Analisis Hasil

Analisis data yang digunakan dalam penelitian ini menggunakan data kuantitatif beserta teknik analisis data untuk aplikasi RapidMiner dan teknik data mining untuk Algoritma Naive Bayes.

2.6. Kesimpulan

Pada tahap ini kesimpulan akan dibuat berdasarkan studi yang telah dilakukan. Apakah itu selaras dengan tujuan dan harapan yang diinginkan atau belum. Informasi yang dihasilkan dari aplikasi RapidMiner akan menampilkan dan memberikan informasi prediksi angka kelahiran kepada Kantor Desa Randau Jekak, kemudian Kantor Desa dapat mengambil keputusan yang baik dan tepat dalam mengatur laju pertumbuhan penduduk dengan hasil dari penelitian ini.

3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan implementasi data mining dengan algoritma Naive Bayes, dimana data yang diperoleh adalah data kependudukan di desa Randau Jekak.

Dari data yang sudah di dapat tahap selanjutnya adalah dengan mendefinisikan data yang akan digunakan sebagai data training dan data testing, kemudian akan dihitung prediksi atau probabilitasnya menggunakan algoritma Naive Bayes.

Tabel 2. *Data Training*

Tahun	JP	AP	AK	WUS	Keterangan
2011	1.310	64	40	339	Tinggi
2012	2.016	60	61	390	Tinggi
2013	3.217	72	100	500	Tinggi
2014	3.641	68	110	490	Tinggi
2015	3.765	65	113	450	Tinggi
2016	3.996	74	69	511	Rendah
2017	4.132	75	100	610	Rendah
2018	4.325	69	97	655	Rendah
2019	4.920	77	151	599	Tinggi
2020	5.200	89	80	662	Rendah

Setelah menentukan data training selanjutnya yang akan dilakukan adalah menentukan data testing yang akan digunakan sebagai berikut.

Tabel 3. *Data Testing*

Tahun	JP	AP	AK	WUS
2021	5513	75	175	700

Setelah selesai menentukan data testing, selanjutnya yang akan dilakukan adalah melakukan perhitungan menggunakan Algoritma Naive Bayes.

3.1. Perhitungan Menggunakan tools RapidMiner

RapidMiner merupakan sebuah tools yang digunakan dalam machine learning, data mining, text mining dan predictive analytics [Muslehatin & Ibnu, 2017]. Rapid Miner adalah software atau perangkat lunak yang digunakan untuk mengolah data. RapidMiner dapat memudahkan pengguna untuk menghitung data dalam jumlah besar (big data) dengan menggunakan operator yang memiliki fungsi pengolahan data. Data tersebut kemudian akan dihubungkan ke node operator di RapidMiner untuk melihat hasilnya. Hasil yang ditampilkan oleh

quickminer juga bisa dalam bentuk visual dengan grafik, tabel, dll. [14].

Setelah menentukan data testing dan data training, selanjutnya yang dilakukan adalah mengimport data yang sudah dikumpulkan ke dalam RapidMiner. Berikut adalah data yang akan diimport ke RapidMiner.

Tabel 4. Data yang di import ke RapidMiner

Tahun	JUP	AP	AK	WUS	Keterangan
					Label
					Binominal
2011	1.310	64	40	339	Tinggi
2012	2.016	60	61	390	Tinggi
2013	3.217	72	100	500	Tinggi
2014	3.641	68	110	490	Tinggi
2015	3.765	65	113	450	Tinggi
2016	3.996	74	69	511	Rendah
2017	4.132	75	100	610	Rendah
2018	4.325	69	97	655	Rendah
2019	4.920	77	151	599	Tinggi
2020	5.200	89	80	662	Rendah

3.2. Proses Perhitungan Probabilitas Menggunakan RapidMiner

Hasil perhitungan nilai P (Ci) dari kelas (label) Keterangan menggunakan tools RapidMiner adalah sebagai berikut.

SimpleDistribution

```
Distribution model for label attribute Keterangan

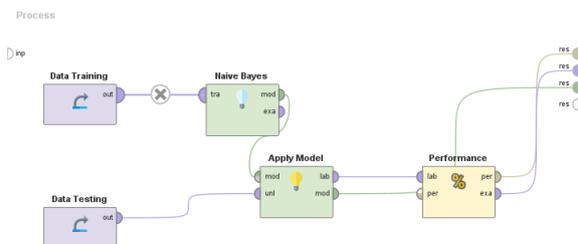
Class Tinggi (0.600)
5 distributions

Class Rendah (0.400)
5 distributions
```

Gambar 2. Hasil Perhitungan Menggunakan tools RapidMiner

3.3. Proses Prediksi Data Testing Menggunakan RapidMiner

Setelah melakukan perhitungan terhadap probabilitas, kemudian melakukan Pemodelan Data Testing menggunakan RapidMiner adalah sebagai berikut.



Gambar 3. Hasil Pemodelan Data Testing

Berdasarkan hasil perhitungan menggunakan RapidMiner terlihat bahwa angka kelahiran di Desa Randau Jekak pada tahun 2021 “tinggi”.

...	Ket...	prediction(K...	confi...	con...	Tahun	Jumlah Pend...	Angka Perni...	Angka Kelah...	Wanita Us...
1	Tinggi	Tinggi	1.000	0.000	2021	5513	75	175	700

Gambar 4. Hasil Prediksi Data Testing angka kelahiran menggunakan tools RapidMiner

Berdasarkan hasil penelitian yang telah di dapatkan, akan menghasilkan parameter dari setiap atribut yang tertinggi dan terendah. Berikut merupakan hasil parameter yang telah di dapat menggunakan tools RapidMiner.

Attribute	Parameter	Tinggi	Rendah
Tahun	mean	2014	2017.750
Tahun	standard deviation	2.828	1.708
Jumlah Penduduk	mean	3144.833	4413.250
Jumlah Penduduk	standard deviation	1298.314	541.591
Angka Pernikahan	mean	67.667	76.750
Angka Pernikahan	standard deviation	6.088	8.578
Angka Kelahiran	mean	95.833	86.500
Angka Kelahiran	standard deviation	39.716	14.617
Wanita Usia Subur	mean	461.333	609.500
Wanita Usia Subur	standard deviation	91.051	69.592

Gambar 5. Hasil Perhitungan Parameter Menggunakan RapidMiner

4. DISKUSI

Dalam penelitian yang dilakukan oleh Ranti Pancasasti dan Enis Khaerunisa dalam jurnalnya yang berjudul “Analisis Dampak Laju Pertumbuhan Penduduk Terhadap Aspek Kependudukan Berwawasan Gender Pada Urban Area Di Kota Serang” menjelaskan bahwa dengan lajunya pertumbuhan penduduk di kota Serang berakibat pada pemukiman yang semakin padat, semakin banyaknya limbah rumah tangga serta fungsi hutan yang berubah. Jadi untuk memenuhi kebutuhan industri, perumahan, dan lain-lain. Lajunya pertumbuhan penduduk ini juga dapat berdampak pada terjadinya rawan pangan, hal ini terjadi karena meningkatnya laju pertumbuhan penduduk yang begitu pesat sehingga kebutuhan pangan meningkat [19].

Yohanni Syahra, Ismawardi Santoso, dan Rini Kustini dalam jurnalnya yang berjudul “Implementasi Data Mining Untuk Prediksi Angka Kelahiran Bayi Pada Desa Sibolangit Menggunakan Multi Regresi” menyimpulkan bahwa tingginya angka kelahiran dapat mempengaruhi kualitas penduduk serta gizi di daerah tersebut. Dengan naiknya angka kelahiran di Desa tersebut berakibat pada banyaknya anak yang kurang gizi di Desa Sibolangit yang sampai sekarang belum mengalami perbaikan meskipun sudah terdapat sedikit penurunan angka kelahiran di desa tersebut dan dengan menggunakan metode multi regresi dapat membantu dalam mendata angka kelahiran bayi karena penggunaan metode multi regresi ini adalah dengan cara melakukan prediksi pada suatu variable kontinu

dengan mengansumsikan sebuah model linier atau non linier [20].

Oleh karena itu, Penulis ingin mengkaji mengenai permasalahan kali ini dengan menambah jumlah data dan dengan data yang berbeda yang akan di teliti menggunakan data training dan data testing untuk melihat seberapa tingkat keakuratan Algoritma Naïve Bayes dalam memprediksi angka kelahiran di Desa Randau Jekak. Diharapkan penelitian ini dapat membantu lembaga atau instansi di Desa Randau Jekak dalam pengambilan sebuah keputusan dalam rangka menekan lajunya angka kelahiran bayi setiap tahunnya.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil pengujian serta analisa data yang sudah dibahas, maka dapat menarik kesimpulan bahwa Data Mining sangat berguna dalam memprediksi angka kelahiran untuk tahun berikutnya dalam rangka menekan laju pertumbuhan penduduk, selain itu Data Mining juga dapat membantu pihak atau Lembaga di Desa Randau Jekak dalam memprediksi angka kelahiran di Desa Randau Jekak untuk tahun mendatang. Selain itu, hasil dalam memprediksi angka fertilitas atau angka kelahiran bayi di Desa Randau Jekak pada tahun 2021 dengan menggunakan tools Rapidminer dengan kelas (label) keterangan yang sama yaitu bernilai 'Tinggi'.

Dari Hasil penelitian ini, pihak Kantor Desa di Randau Jekak dapat melakukan evaluasi untuk laporan data kependudukan serta sebagai acuan dalam penekanan jumlah angka kelahiran di tahun mendatang serta pihak Kantor Desa Randau Jekak dapat bekerja sama dengan petugas kesehatan dalam menekan laju pertumbuhan penduduk, seperti sosialisasi Keluarga Berencana (KB) dan sebagainya.

Adapun saran untuk pengembangan penelitian serupa selanjutnya adalah sebagai berikut.

Penentuan data training dapat mempengaruhi hasil klasifikasi, maka dari itu untuk penelitian selanjutnya diharapkan untuk menambah atribut atau data training yang lebih lengkap karena penentuan tingkat akurasi yang di dapat ditentukan oleh data training.

Akan lebih baik jika menggunakan perbandingan dari beberapa software serupa seperti Weka, R, dan sebagainya untuk mendapatkan hasil yang akurat.

DAFTAR PUSTAKA

- [1] M. Idris, "Implementasi Data Mining Dengan Algoritma Naive Bayes Untuk Memprediksi Angka Kelahiran," *J. Pelita Inform.*, vol. 7, no. 3, pp. 421–428, 2019, [Online]. Available: <https://ejurnal.stmik-budidarma.ac.id/index.php/pelita/article/view/1154>
- [2] M. G. Sadewo, A. P. Windarto, and D. Hartama, "Penerapan Datamining Pada Populasi Daging Ayam Ras Pedaging Di Indonesia Berdasarkan Provinsi Menggunakan K-Means Clustering," *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 2, no. 1, pp. 60–67, 2017, doi: 10.30743/infotekjar.v2i1.164.
- [3] A. Arsandi, D. R. I. Ismiyati, and F. Hermawan, "Dampak Pertumbuhan Penduduk Terhadap Infrastruktur Di Kota Semarang," *J. Karya Tek. Sipil S1 Undip*, vol. 6, no. 4, pp. 1–14, 2017.
- [4] Jefi, "Prediksi Bayi Lahir Secara Prematur Dengan Menggunakan Metode C . 45 Berbasis Particle Swarm Optimization Pada Klinik Umi," *Indones. J. Netw. Secur.*, vol. 8, no. 2, pp. 1–7, 2019.
- [5] N. I. Parihah, S. Hartini, and J. Siregar, "Prediksi Angka Kelahiran Bayi Pada Desa Tridaya Sakti Dengan Menggunakan Algoritma Naive Bayes," *J. Students' Res. Comput. Sci.*, vol. 1, no. 2, pp. 77–88, 2020, doi: 10.31599/jsrscs.v1i2.423.
- [6] N. R. Indraswari and Y. I. Kurniawan, "Aplikasi Prediksi Usia Kelahiran Dengan Metode Naive Bayes," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 9, no. 1, pp. 129–138, 2018, doi: 10.24176/simet.v9i1.1827.
- [7] L. . T. Nike Syafitri , Sutardi, "Aplikasi Forecasting Mengenai Angka Kelahiran di Kota Kendari Menggunakan Metode Regresi Linear Berganda (Studi Kasus: Dinas Kesehatan Kota Kendari)," *Inform. Jur. Tek. Tek. Fak. Oleo, Univ. Halu*, vol. 4, no. 1, pp. 31–38, 2018.
- [8] H. Zahab, Y. Dharmawan, S. Winarni, B. Biostatistika, and F. K. Masyarakat, "Hubungan Antara Perilaku Pacaran Remaja Dan Pernikahan Di Bawah Usia 20 Tahun Terhadap Angka Kelahiran Menurut Kelompok Umur 15-19 Tahun Di Indonesia (Analisis Data Sekunder Sdki 2012)," *J. Kesehat. Masy.*, vol. 5, no. 3, pp. 56–65, 2017.
- [9] F. S. D. Purba, H. S. Tambunan, I. S. Saragih, and ..., "Implementasi Algoritma K-Medoids Clustering Dalam Pengelompokan Angka Kelahiran Di Kota Pematang Siantar," *Pros. Semin. ...*, vol. 2, pp. 124–131, 2020, [Online]. Available: <http://tunasbangsa.ac.id/seminar/index.php/enaris/article/view/152>
- [10] D. K. Indahsari and Y. I. Kurniawan, "Aplikasi Prediksi Usia Kelahiran Dengan Metode K-Nearest Neighbor," *J. Kebidanan*, vol. XI, no. 100, pp. 1–14, 2019.
- [11] A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya

- Penggunaan Listrik Rumah Tangga,” *Creat. Inf. Technol. J.*, vol. 2, no. 3, pp. 207–217, 2019, [Online]. Available: <https://books.google.co.id/books?id=8NKdDwAAQBAJ&lpq=PP1&ots=YfxJP3Ocpe&q=Data Mining Algoritma dan Implementasi dengan Pemrograman PHP&lr&pg=PR4#v=onepage&q=Data Mining Algoritma dan Implementasi dengan Pemrograman PHP&f=false>
- [12] I. K. Syahputra, F. A. Bachtiar, and S. A. Wicaksono, “Implementasi Data Mining untuk Prediksi Mahasiswa Pengambil Mata Kuliah dengan Algoritme Naive Bayes,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 5902–5910, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id./index.php/j-ptiik/article/view/3464>
- [13] C. Fadlan, S. Ningsih, and A. P. Windarto, “Penerapan Metode Naïve Bayes Dalam Klasifikasi Kelayakan Keluarga Penerima Beras Rastra,” *J. Tek. Inform. Musirawas*, vol. 3, no. 1, p. 1, 2018, doi: 10.32767/jutim.v3i1.286.
- [14] A. F. B. Watratan, A. Puspita, and D. Moeis, “Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia,” *J. Appl. Comput. Sci. Technol. (Jacost)*, vol. 1, no. 1, pp. 7–14, 2020.
- [15] A. Z. A. M. Bajabir, “Penerapan metode naive bayes untuk prediksi menentukan karyawan tetap pada pt. ysp industries indonesia,” 2018.
- [16] Suntoro J, “Data Mining Algoritma dan Implementasi dengan Pemrograman PHP,” *PT Elex Media Komputindo*, 2019, [Online]. Available: <https://books.google.co.id/books?id=8NKdDwAAQBAJ&lpq=PP1&ots=YfxJP3Ocpe&q=Data Mining Algoritma dan Implementasi dengan Pemrograman PHP&lr&pg=PR4#v=onepage&q=Data Mining Algoritma dan Implementasi dengan Pemrograman PHP&f=false>
- [17] A. Muzaki and A. Witanti, “Sentiment Analysis of the Community in the Twitter To the 2020 Election in Pandemic Covid-19 By Method Naive Bayes Classifier,” *J. Tek. Inform.*, vol. 2, no. 2, pp. 101–107, 2021, doi: 10.20884/1.jutif.2021.2.2.51.
- [18] A. Alwi, P. Studi, T. Informatika, F. Teknik, and U. M. Ponorogo, “the Concept of Naive Bayes and Its Simple Use for Prediction Final Konsep Naive Bayes Dan Penggunaannya Secara Sederhana,” vol. 3, no. 1, pp. 133–140, 2022.
- [19] R. Pancasasti and E. Khaerunisa, “Analisis Dampak Laju Pertumbuhan Penduduk Terhadap Aspek Kependudukan Berwawasan Gender Pada Urban Area Di Kota Serang,” *Tirtayasa Ekon.*, vol. 13, no. 1, p. 130, 2018, doi: 10.35448/jte.v13i1.4231.
- [20] Y. Syahra, I. Santoso, and R. Kustini, “Implementasi Data Mining Untuk Prediksi Angka Kelahiran Bayi Pada Desa Sibolangit Menggunakan Multi Regresi,” *Semin. Nas. Sains Teknol. Inf.*, no. 1, pp. 687–690, 2019.