

Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Dalam Memprediksi Angka Kelahiran

Saveria Marbun¹, Nurcahyo Budi Nugroho², Rico Imanta Ginting³

^{1,2} Sistem Informasi, STMIK Triguna Dharma

³ Teknik Komputer, STMIK Triguna Dharma

Email: ¹saveriamarbun97@gmail.com, ²nurcahyobn@gmail.com, ³icoversi90@gmail.com

Email Penulis Korespondensi: saveriamarbun97@gmail.com

Abstrak

Data Mining merupakan kegiatan yang mencakup pengumpulan data, pemakaian data historis yang menjelaskan keteraturan, pola, dan hubungan didalam set data berukuran besar. Angka kelahiran merupakan suatu hal yang dapat mempengaruhi peningkatan pertumbuhan penduduk. Saat ini yang menjadi permasalahan yaitu tidak semua daerah mengalami keseimbangan jumlah penduduk dan melonjaknya jumlah penduduk di kecamatan pakkat memang perlu diperhatikan. Jumlah kelahiran bayi di Kecamatan Pakkat setiap tahunnya semakin bertambah.

Dari permasalahan yang dialami maka dibutuhkan suatu solusi untuk dapat memecahkan masalah tersebut yaitu dengan cara menggunakan metode. Metode yang digunakan pada saat ini adalah metode Naïve Bayes untuk melakukan perhitungan dalam menentukan tinggi rendahnya angka kelahiran pertahun di Kecamatan Pakkat. Metode Naïve Bayes adalah sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan.

Hasil dari penelitian ini adalah: pertama Menganalisa data sistem yang akan dibuat dengan cara mengumpulkan informasi dari berbagai pihak perusahaan melalui wawancara dan observasi, kedua Untuk membuat perancangan sistem yaitu dengan cara mengetahui kebutuhan sistem dalam segmentasi data desa. Berdasarkan hasil penelitian, sebelum dilakukan perancangan sistem maka terlebih dahulu dirancang menggunakan Unified Modelling Language (UML) dan melakukan pembuatan database, ketiga Dalam membangun sistem terhadap desain sistem yang sudah dirancang adalah dengan melakukan pengkodean dari sistem input, proses dan output menggunakan bahasa pemrograman yang dibutuhkan program dan sistem aplikasi dijalankan dengan melakukan perhitungan metode dengan cara menguji coba sistem untuk meminimalisir kesalahan terhadap aplikasi.

Kata Kunci: Data Mining, Naïve Bayes, Kecamatan Pakkat, Angka Kelahiran

1. PENDAHULUAN

Angka kelahiran merupakan suatu hal yang dapat mempengaruhi peningkatan pertumbuhan penduduk. Jumlah penduduk yang besar merupakan beban bagi pembangunan. Menurut teori Malthus menyatakan bahwa pertumbuhan penduduk yang besar bukanlah kesejahteraan yang didapatkan melainkan kemelaratan yang akan ditemui apabila jumlah angka kelahiran penduduk tidak dikendalikan sesegera mungkin. Jumlah angka kelahiran di Kecamatan Pakkat semakin bertambah disetiap tahun nya. Maka dari itu dengan adanya Data Mining menggunakan Algoritma Naïve Bayes dapat mempermudah dalam membantu perhitungan untuk memprediksi angka kelahiran di Kecamatan Pakkat. Data Mining dalam memprediksi angka kelahiran bertujuan untuk mengetahui jumlah angka kelahiran pada tahun yang akan mendatang dengan menggunakan Algoritma Naïve Bayes[1]. Dengan digunakan nya Algoritma Naïve bayes ini diharapkan dapat menyelesaikan permasalahan yang terjadi di Kecamatan Pakkat dalam menangani serta mengatasi perhitungan angka kelahiran bayi dan dapat membantu pihak di Kecamatan Pakkat dalam mengatur pertumbuhan jumlah penduduk di tahun-tahun yang akan mendatang[2].

Hasil yang diperoleh dari data yang sudah diambil dapat dihitung dengan data mining dengan menggunakan Algoritma Naïve Bayes sehingga menghasilkan sebuah informasi yang dapat digunakan sebagai acuan dalam mengetahui jumlah angka kelahiran [3]. Persebaran penduduk secara optimal berdasarkan keseimbangan jumlah penduduk dengan daya dukung oleh kualitas masyarakat dan kualitas lingkungan yang merupakan harapan pemerintah untuk setiap daerahnya. Namun yang menjadi permasalahan adalah tidak semua daerah memiliki keseimbangan jumlah penduduk dan semakin melonjaknya jumlah penduduk terutama di Kecamatan Pakkat memang perlu diperhatikan.

Para pengambil keputusan berusaha untuk memanfaatkan gudang data yang sudah dimiliki untuk menggali informasi yang berguna dalam membantu mengambil keputusan, hal ini mendorong munculnya berbagai cabang ilmu baru untuk mengatasi masalah atau pola penting dan menarik data dalam jumlah yang besar yang disebut dengan Data Mining[4]. Pada proses prediksi tersebut diterapkan metode Algoritma Naïve Bayes yang menggambarkan hubungan antara dua variabel atau lebih.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Metodologi Penelitian merupakan suatu teknik atau langkah yang dilakukan untuk mengumpulkan informasi atau data yang lebih spesifik rancangan penelitian yang akan dibuat dengan menggunakan beberapa teknik pengumpulan data. Berikut ini merupakan metode penelitian yang digunakan dalam penelitian, yaitu:

- Pengumpulan data (data collecting), dengan melakukan pengumpulan suatu data terdapat beberapa hal yang perlu dilakukan yaitu : Observasi, dan wawancara.
- Pada studi pustaka, penelitian yang dilakukan memerlukan studi atau landasan teoritis dengan menggunakan jurnal-jurnal.

2.2 Penduduk

Penduduk merupakan sejumlah manusia yang menempati suatu daerah tertentu pada waktu tertentu. Pertumbuhan jumlah penduduk yang begitu besar di Indonesia masih menjadi persoalan terutama pada suatu wilayah karena jika semakin besar pertumbuhan dan jumlah penduduk tentu harus diikuti oleh penambahan fasilitas social yang dibutuhkan untuk menunjang kehidupan berikutnya [5].

2.3 Data Mining

Data mining dapat digunakan untuk pengelompokan Kasus dan Kematian COVID-19 di Asia Tenggara [6]. Data Mining merupakan proses yang mempekerjakan satu atau lebih teknik pembelajaran computer untuk menganalisis dan mengekstraksi pengetahuan secara otomatis [7]. Data mining mencakup pengumpulan data, pemakaian data historis yang menjelaskan keteraturan, pola, hubungan didalam set data ukuran besar [8]. Definisi lain dari data mining adalah proses yang akan mempekerjakan computer untuk menganalisis dan mengekstrak pengetahuan secara otomatis atau serangkaian proses untuk menggali beberapa nilai tambahan pada suatu kumpulan yang berupa pengetahuan yang selama ini tidak diketahui secara manual [9]. Pada banyak kasus, ini dapat digambarkan sebagai tahap preprocessing pengenalan pola [10].

2.4 Algoritma Naïve Bayes

Metode Naïve Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan dengan menjumlahkan frekuensi dan kombinasi dari dataset yang diberikan[11]. Algoritma menggunakan teorema bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas[12].

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

Keterangan :

- X : Sampel data yang memiliki kelas (label) yang tidak diketahui
 H : Hipotesa X adalah data kelas (label) P(H) : Peluang dari hipotesa
 P(X) : peluang dari data sampel yang diamati P(X|H) : peluang dari data sampel X bila diasumsikan bahwa hipotesa benar

adapun alur dari metode naïve bayes sebagai berikut[13] :

- Menghitung nilai peluang kasus baru dari setiap hipotesa
- Menghitung nilai akumulasi peluang dari setiap kelas P(X|Ci)
- Menghitung nilai P(X|Ci) x P(Ci)
- Menentukan kelas dari kasus baru

3. HASIL DAN PEMBAHASAN

3.1 Penerapan Algoritma Naïve Bayes

Identifikasi data dilakukan setelah data terkumpul dengan kebutuhan sistem. Berikut kasus data mining dengan pengelompokan sampel data penduduk kecamatan pakkat dengan menggunakan algoritma *naïve bayes*[12].

Tabel 1. Sampel Data Kependudukan Kecamatan Pakkat

No	Nama Desa	Jumlah Laki-laki	Jumlah perempuan	Jumlah Penduduk	Angka Pernikahan	Jumlah WUS	Angka Kelahiran	Keterangan
1	Pakkat Hauagong	1.784	2019	3.803	25	1200	20	Tinggi
2	Hauagong	432	402	834	30	200	15	Tinggi
3	Ambobi Paranginan	422	408	830	16	240	10	Rendah

4	Tukka Dolok	717	737	1.454	18	1.351	13	Tinggi
5	Peadungdung	449	448	897	15	352	10	Tinggi
6	Purba Baringin	628	640	1.268	14	456	13	Rendah
7	Purba Bersatu	780	703	1.483	12	521	11	Tinggi
8	Purba Sianjur	283	310	593	13	102	9	Rendah
9	Banuarea	257	315	572	20	150	17	Rendah
10	Panggugunan	331	302	633	19	168	15	Tinggi
11	Pulogodang	420	398	818	18	258	10	Rendah
12	Sipagabu	344	394	738	17	236	12	Rendah
13	Manalu	557	574	1.131	16	247	9	Rendah
14	Karya	578	588	1.166	18	291	10	Rendah
15	Sijarango	440	416	856	24	156	18	Tinggi
16	Sijarango 1	355	415	770	10	1278	5	Rendah
17	LumbanTongatonga	612	547	1.159	16	248	8	Rendah
18	Rura Tanjung	995	974	1.969	14	675	10	Tinggi
19	Rura Aek Sopang	719	705	1.424	20	548	16	Rendah
20	Siambaton	181	173	354	17	120	6	Rendah
21	Siambaton Pahe	430	418	848	25	256	11	Rendah
22	Parmonangan	353	362	715	20	124	15	Tinggi
23	Laksa	1.921	2.001	3.922	16	140	13	Tinggi
24	Pulogiring	337	410	747	17	125	12	Rendah
25	Martonabala	1.938	1.967	3.905	19	220	15	Rendah
26	Hariara nauli	300	364	664	20	150	10	Rendah
27	Songkalmihuta	358	315	673	18	270	14	Rendah
28	Simatabo	360	220	580	20	310	14	Rendah
29	Pasarambasang	356	168	524	13	235	12	Rendah
30	Huta Ginjang	420	367	787	16	170	12	Rendah

1. Menghitung nilai peluang kasus baru dari setiap hipotesa dengan label yang ada $P(XK|Ci)$

A. Memprediksi angka kelahiran pada desa “Pakkat Hauagong”

Tabel 2. Prediksi Desa Pakkat Hauagong

Nama Desa	Kriteria 1	Kriteria 2	Kriteria 3	Kriteria 4
Pakkat Hauagong	Diatas 3.803 Dibawah 3.803	Diatas 25 Dibawah 25	Diatas 1.200 Dibawah 1.200	Diatas 20 Dibawah 20

Menghitung nilai $P(XK|Ci)$

- $P(\text{Kriteria 1} = \text{“Dibawah 3.803”} \mid \text{Klasifikasi} = \text{“Rendah”})P(\text{Kriteria 1} = 27/30 = 0,9$
- $P(\text{Kriteria 1} = \text{“Diatas 3.803”} \mid \text{Klasifikasi} = \text{“Tinggi”})P(\text{Kriteria 1} = 3/30 = 0,1$
- $P(\text{Kriteria 2} = \text{“Dibawah 25”} \mid \text{Klasifikasi} = \text{“Rendah”})P(\text{Kriteria 2} = 13/30 = 0,4$
- $P(\text{Kriteria 2} = \text{“Diatas 25”} \mid \text{Klasifikasi} = \text{“Tinggi”})P(\text{Kriteria 2} = 2/30 = 0,0$
- $P(\text{Kriteria 3} = \text{“Dibawah 1.200”} \mid \text{Klasifikasi} = \text{“Rendah”})P(\text{Kriteria 3} = 27/30 = 0,9$
- $P(\text{Kriteria 3} = \text{“Diatas 1.200”} \mid \text{Klasifikasi} = \text{“Tinggi”})P(\text{Kriteria 3} = 3/30 = 0,1$
- $P(\text{Kriteria 4} = \text{“Dibawah 20”} \mid \text{Klasifikasi} = \text{“Rendah”})P(\text{Kriteria 4} = 29/30 = 0,1$
- $P(\text{Kriteria 4} = \text{“Diatas 20”} \mid \text{Klasifikasi} = \text{“Tinggi”})P(\text{Kriteria 4} = 1/30 = 0,1$

2. Hitung nilai akumulasi peluang $P(X|Ci)$ untuk setiap kelas(label)

$$P(X|\text{Klasifikasi} = \text{“Tinggi”}) = 0,1 * 0,2 * 0,1 * 0,3 = 0,0006$$

$$P(X|\text{Klasifikasi} = \text{“Rendah”}) = 0,9 * 0,8 * 0,7 * 0,9 = 0,4536$$

3. Hitung nilai $P(X|Ci) * P(Ci)$

$$(P(X|\text{Klasifikasi} = \text{“Tinggi”}) * P(\text{Klasifikasi} = \text{“Angka Kelahiran”})) = 0,0006 * 10/30 = 0,0002$$

$$P(X|Ci) * P(Ci)$$

$$(P(X|\text{Klasifikasi} = \text{“Rendah”}) * P(\text{Klasifikasi} = \text{“Angka Kelahiran”})) = 0,4536 * 20/30 = 0,6666$$

Berdasarkan perhitungan akhir dengan mengalikan nilai peluang dari kasus yang diangkat, kita melihat bahwa nilai $P(X|\text{Keterangan} = \text{“Tinggi”})$ lebih rendah dari $P(X|\text{Keterangan} = \text{“Rendah”})$, maka :

Tabel 3. Hasil Prediksi Pada Desa Pakkat Hauagong

Nama Desa	Kriteria 1	Kriteria 2	Kriteria 3	Kriteria 4	Keterangan
Pakkat Hauagong	Diatas 3.803 Dibawah 3.803	Diatas 25 Dibawah 25	Diatas 1.200 Dibawah 1.200	Diatas 20 Dibawah 20	Rendah

B. Memprediksi angka kelahiran pada desa “Pangguguan”

Tabel 4. Prediksi Desa Pangguguan

Nama Desa	Kriteria 1	Kriteria 2	Kriteria 3	Kriteria 4
Pangguguan	Diatas 633 Dibawah 633	Diatas 16 Dibawah 16	Diatas 168 Dibawah 168	Diatas 9 Dibawah 9

Menghitung Nilai P(XK|Ci)

- P (Kriteria 1 = “Dibawah 633” | Klasifikasi = “Rendah”)P (Kriteria 1 = 5/30 =0,1
- P (Kriteria 1 = “Diatas 633” | Klasifikasi = “Tinggi”)P (Kriteria 1 = 25/30 = 0,8
- P (Kriteria 2 = “Dibawah 10” | Klasifikasi = “Rendah”)P (Kriteria 2 = 16/30 =0,5
- P (Kriteria 2 = “Diatas 10” | Klasifikasi = “Tinggi”)P (Kriteria 2 = 21/30 = 0,7
- P (Kriteria 3 = “Dibawah 150” | Klasifikasi = “Rendah”)P (Kriteria 3 = 7/30 =0,2
- P (Kriteria 3 = “Diatas 168” | Klasifikasi = “Tinggi”)P (Kriteria 3 = 9/30 = 0,3
- P (Kriteria 4 = “Dibawah 5” | Klasifikasi = “Rendah”)P (Kriteria 4= 19/30 = 0,6
- P (Kriteria 4 = “Diatas 5” | Klasifikasi = “Tinggi”)P (Kriteria 4 = 11/30 = 0,3

Hitung nilai akumulasi peluang P(X|Ci) untuk setiap kelas(label)

$$P(X|Klasifikasi = \text{“Tinggi”}) = 0,8 * 0,7 * 0,3 * 0,3 = 0,0504$$

$$P(X|Klasifikasi = \text{“Rendah”}) = 0,1 * 0,5 * 0,2 * 0,6 = 0,006$$

Hitung nilai P(X|Ci) * P (Ci)

$$(P(X|Klasifikasi = \text{“Tinggi”}) * P(Klasifikasi = \text{“Angka Kelahiran”})) = 0,0504 * 10/30 = 0,0168$$

$$P(X|Ci) * P (Ci)$$

$$(P(X|Klasifikasi = \text{“Rendah”}) * P(Klasifikasi = \text{“Angka Kelahiran”})) = 0,006 * 20/30 = 0,004$$

Berdasarkan perhitungan akhir dengan mengalikan nilai peluang dari kasus yang diangkat, kita melihat bahwa nilai P(X|Keterangan “Tinggi”) lebih tinggi dari P(X|Keterangan “Rendah”), maka :

Tabel 5. Hasil Prediksi Pada Desa Pangguguan

Nama Desa	Kriteria 1	Kriteria 2	Kriteria 3	Kriteria 4	Keterangan
Pangguguan	Diatas 633 Dibawah 633	Diatas 18 Dibawah 18	Diatas 168 Dibawah 168	Diatas 10 Dibawah 10	Tinggi

C. Memprediksi angka kelahiran pada desa “Ambobi Paranginan”

Tabel 6. Prediksi Ambobi Paranginan

Nama Desa	Kriteria 1	Kriteria 2	Kriteria 3	Kriteria 4
Ambobi Paranginan	Diatas 830 Dibawah 830	Diatas 16 Dibawah 16	Diatas 240 Dibawah 240	Diatas 10 Dibawah 10

Menghitung Nilai P(XK|Ci)

- P (Kriteria 1 = “Dibawah 830” | Klasifikasi = “Rendah”)P (Kriteria 1 = 14/30 =0,4
- P (Kriteria 1 = “Diatas 830” | Klasifikasi = “Tinggi”)P (Kriteria 1 = 15/30 = 0,5
- P (Kriteria 2 = “Dibawah 16” | Klasifikasi = “Rendah”)P (Kriteria 2 = 14/30 =0,4
- P (Kriteria 2 = “Diatas 16” | Klasifikasi = “Tinggi”)P (Kriteria 2 = 11/30 = 0,3
- P (Kriteria 3 = “Dibawah 240” | Klasifikasi = “Rendah”)P (Kriteria 3 = 13/30 =0,4
- P (Kriteria 3 = “Diatas 240” | Klasifikasi = “Tinggi”)P (Kriteria 3 = 14/30 = 0,4
- P (Kriteria 4 = “Dibawah 10” | Klasifikasi = “Rendah”)P (Kriteria 4= 10/30 = 0,3
- P (Kriteria 4 = “Diatas 10” | Klasifikasi = “Tinggi”)P (Kriteria 4 = 17/30 = 0,5

Hitung nilai akumulasi peluang P(X|Ci) untuk setiap kelas(label)

$$P(X|Klasifikasi = \text{“Tinggi”}) = 0,5 * 0,4 * 0,4 * 0,3 = 0,024$$

$$P(X|Klasifikasi = \text{“Rendah”}) = 0,4 * 0,4 * 0,4 * 0,3 = 0,0192$$

Hitung nilai P(X|Ci) * P (Ci)

$$(P(X|Klasifikasi = \text{“Tinggi”}) * P(Klasifikasi = \text{“Angka Kelahiran”})) = 0,024 * 10/30 = 0,16$$

$$P(X|Ci) * P (Ci)$$

$$(P(X|Klasifikasi = \text{“Rendah”}) * P(Klasifikasi = \text{“Angka Kelahiran”})) = 0,0192 * 20/30 = 0,0128$$

Berdasarkan perhitungan akhir dengan mengalikan nilai peluang dari kasus yang diangkat, kita melihat bahwa nilai $P(X|Keterangan\ "Tinggi")$ lebih tinggi dari $P(X|Keterangan\ "Rendah")$, maka :

Tabel 7. Hasil Prediksi Pada Desa Pangguguan

Nama Desa	Kriteria 1	Kriteria 2	Kriteria 3	Kriteria 4	Keterangan
Ambobi Paranginan	Diatas 830 Dibawah 830	Diatas 16 Dibawah 16	Diatas 240 Dibawah 240	Diatas 10 Dibawah 10	Tinggi

Tabel 8. Hasil Pengujian Data angka kelahiran

No	Nama Desa	Angka Kelahiran
1	Pakkat Hauagong	Rendah
2	Pangguguan	Tinggi
3.	Ambobi Paranginan	Tinggi

3.2 Implementasi Sistem

Berikut ini merupakan hasil dari tampilan antarmuka sistem yang telah dibangun :

a. Tampilan Form Login

Berguna untuk mengamankan sistem dari user yang tidak bertanggung jawab. Berikut tampilan dari menu login :



Gambar 1. Tampilan form login

b. Tampilan Menu Utama

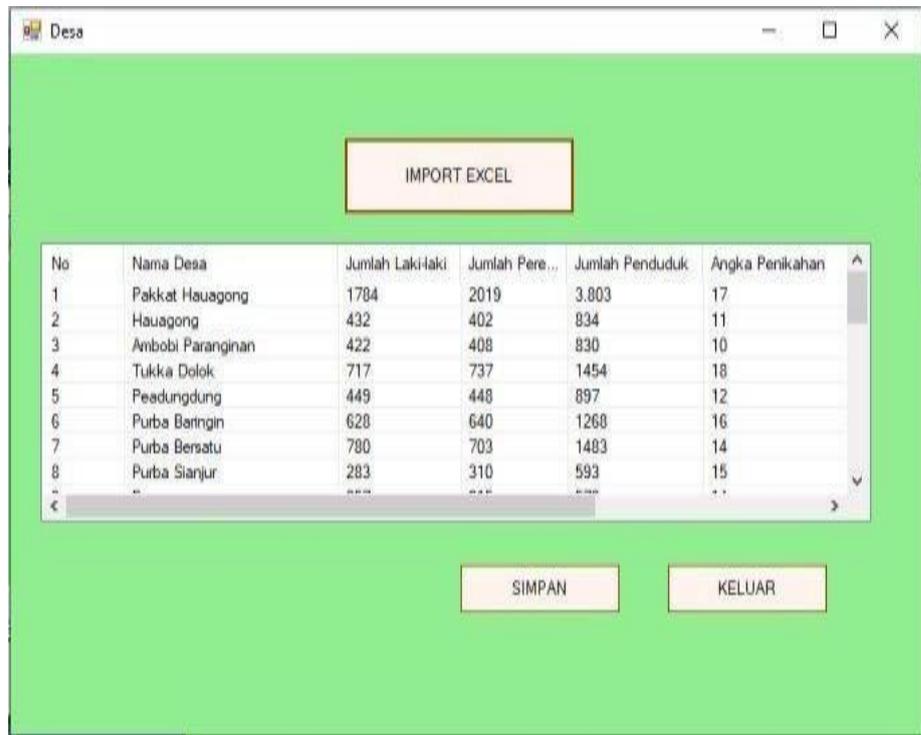
Berguna sebagai penghubung form yang berhubungan dengan data desa, data transaksi, proses dan laporan. Berikut ini adalah tampilan dari menu utama



Gambar 2. Tampilan Menu Utama

c. Menu desa

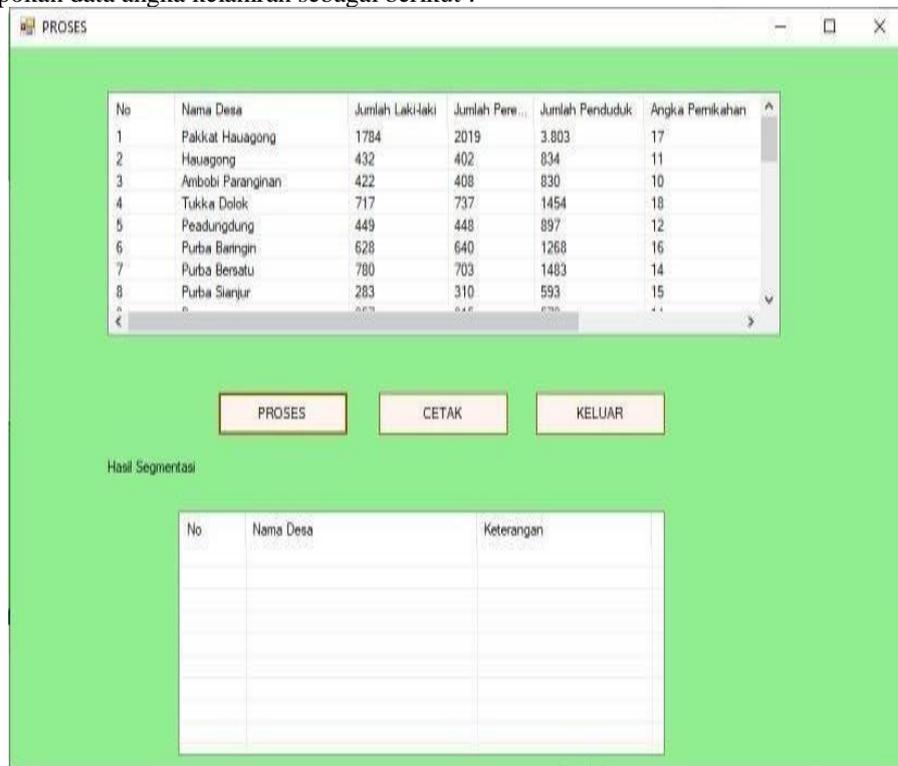
Digunakan untuk penolahan data pada desa. Berupa inport data dan simpan data. Berikut ini adalah tampilan dari menu desa.



Gambar 3. Menu Desa

d. Menu proses

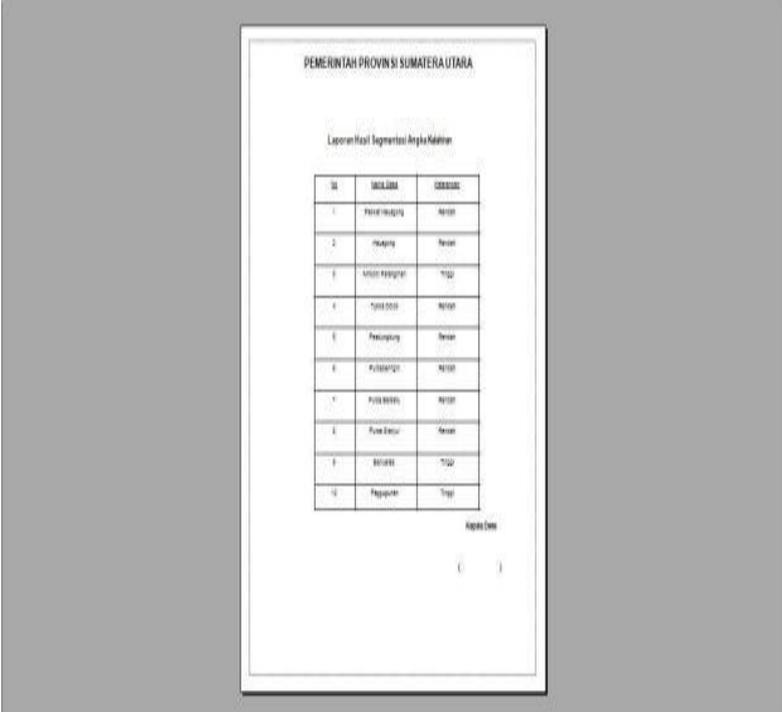
Tahap ini melakukan pengujian terhadap data baru untuk menguji keakuratan sistem. Adapun hasil dari proses program dalam pengelompokan data angka kelahiran sebagai berikut :



Gambar 4. Menu Proses

e. Hasil Laporan

Adapun tampilan dari hasil laporan dari proses program yaitu sebagai berikut :



No	Wilayah	Jumlah
1	Pesisir Barat	10000
2	Pesisir Tengah	10000
3	Kawasan Perbukitan	10000
4	Tapanuli Selatan	10000
5	Pesisir Utara	10000
6	Pesisir Tengah	10000
7	Pesisir Selatan	10000
8	Pesisir Barat	10000
9	Pesisir Tengah	10000
10	Pesisir Utara	10000

Gambar 5. Hasil Laporan

4. KESIMPULAN

Berdasarkan hasil pembahasan dengan menggunakan algoritma naïve bayes untuk memprediksi angka kelahiran yang telah dikemukakan, maka dapat diperoleh beberapa kesimpulan yaitu: dari hasil analisa yang dilakukan untuk memprediksi jumlah angka kelahiran maka dibutuhkan sebuah sistem yang mampu digunakan dalam memprediksi jumlah angka kelahiran. Maka dari itu, dirancang sebuah sistem yang mampu mengadopsi atau menerapkan metode naïve bayes sehingga dapat membantu pihak perusahaan khususnya admin untuk menentukan prediksi jumlah angka kelahiran dengan cepat dan efisien.

UCAPAN TERIMAKASIH

Terima kasih disampaikan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat dan karunia nya sehingga penulis masih dapat merasakan nafas kehidupan sehingga dapat menyelesaikan penelitian ini, juga kepada kedua orangtua yang selalu mendukung dalam setiap proses yang dilalui, terimakasih juga kepada para sahabat yang sudah mau diertpotkan dalam penyusunan penelitian ini, serta pihak-pihak yang telah mendukung terlaksananya penelitian ini.

DAFTAR PUSTAKA

- [1] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2, hal. 160–165, 2018, doi: 10.33096/ilkom.v10i2.303.160-165.
- [2] Y. E. Fadrial, "Algoritma Naive Bayes Untuk Mencari Perkiraan Waktu Studi Mahasiswa," *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 4, no. 1, hal. 20–29, 2021, doi:10.31539/intecom.v4i1.2219.
- [3] Y. Syahra, I. Santoso, dan R. Kustini, "Implementasi Data Mining Untuk Prediksi Angka Kelahiran Bayi Pada Desa Sibolangit Menggunakan Multi Regresi," *Semin. Nas. Sains Teknol. Inf.*, no. 1, hal. 687–690, 2019.
- [4] M. Idris, "Implementasi Data Mining Dengan Algoritma Naive Bayes Untuk Memprediksi Angka Kelahiran," *J. Pelita Inform.*, vol. 7, no. 3, hal. 421–428, 2019.
- [5] "PENGATURAN PENDUDUK PENDATANG DALAM AWIG-AWIG DESA

- PAKRAMAN Oleh: I Ketut Sudantra Fakultas Hukum Universitas Udayana,” no. 1597, hal. 1–17.
- [6] J. Hutagalung, N. L. W. S. R. Ginantra, G. W. Bhawika, W. G. S. Parwita, A. Wanto, and P. D. Panjaitan, “COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm,” *J. Phys. Conf. Ser.*, vol. 1783, no. 1, pp. 1–6, 2021, doi: 10.1088/1742-6596/1783/1/012027.
- [7] R. Girsang, E. F. Ginting, dan M. Hutasuhut, “Penerapan Algoritma C4 . 5 Pada Penentuan Penerima Program Bantuan Pemerintah Daerah,” vol. 1, hal. 449–459, 2022.
- [8] M. Hutasuhut, D. Octaviana, dan J. Halim, “Penerapan Data Mining dalam Menganalisa Pola Kelayakan Siswa Pada Kelas Unggulan Menggunakan Algoritma Iterative Dichotomiser 3 (ID3) pada,” vol. 18, no. 2, hal. 154–160, 2019.
- [9] Y. H. Syahputra and J. Hutagalung, “Superior Class to Improve Student Achievement Using the K-Means Algorithm,” *Sink. J. dan Penelit. Tek. Inform.*, vol. 7, no. 3, pp. 891–899, 2022.
- [10] S. Susanto dan D. Suryani, “Pengantar Data Mining Menggali Pengetahuan Dari Bongkahan Data,” *Buku*, hal. 1–3, 2010.
- [11] Rayuwati, Husna Gemasih, dan Irma Nizar, “IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSI TINGKAT PENYEBARAN COVID,” *Jurnal Ris. Rumpun Ilmu Tek.*, vol. 1, no. 1, hal. 38–46, 2022, doi: 10.55606/jurritek.v1i1.127.
- [12] Y. I. Kurniawan, “Perbandingan Algoritma Naive Bayes dan C.45 dalam Klasifikasi Data Mining,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 4, hal. 455, 2018, doi:10.25126/jtiik.201854803.
- [13] N. I. Parihah, S. Hartini, dan J. Siregar, “Prediksi Angka Kelahiran Bayi Pada Desa Tridaya Sakti Dengan Menggunakan Algoritma Naive Bayes,” *J. Students’ Res. Comput. Sci.*, vol. 1, no. 2, hal. 77–88, 2020, doi: 10.31599/jsrsc.v1i2.423.