



IMPLEMENTATION OF NAIVE BAYES METHOD IN CLASSIFICATION OF PROSPECTIVE KARTU KELUARGA SEJAHTERA

**Hendri¹, Rachman Komarudin², Dony Oscar³, Agus Salim⁴,
Nurul Afni⁵, Yana Iqbal Maulana^{6*}**

Informatika¹, Sistem Informasi^{2,3,4,5}, Teknologi Informasi⁶

Fakultas Teknologi Informasi^{1,2}, Fakultas Teknik dan Informatika^{3,4,5,6}

Universitas Nusa Mandiri^{1,2}, Universitas Bina Sarana Informatika^{3,4,5,6}

hendri.hed@nusamandiri.ac.id¹, rachman.rck@nusamandiri.ac.id²,
dony.dor@bsi.ac.id³, agus.salim@bsi.ac.id⁴, nurul.nrf@bsi.ac.id⁵,
yana.yim@bsi.ac.id⁶

Received: March 30, 2022. **Revised:** May 10, 2022. **Accepted:** May 28, 2022. **Issue Period:** Vol.6 No.1 (2022), Pp. 265-272

Abstrak: Kartu Keluarga Sejahtera merupakan bantuan bagi keluarga kurang mampu yang bertujuan untuk membantu masyarakat miskin yang diharapkan bisa meringankan beban masyarakat. Desa sumberjaya merupakan salah satu desa yang memiliki total penduduk sebanyak 73.631 jiwa, Ketentuan dalam penentuan status penerima Kartu Keluarga Sejahtera berdasarkan observasi serta wawancara dengan Kepala Urusan Perencanaan Desa Sumberjaya terkait proses penerima Kartu Keluarga Sejahtera yang masih kurang optimal dalam proses pendataanya sehingga masih terdapat data yang kurang tepat dalam penerimaannya. Oleh karena itu perlu dilakukan penilaian dengan menggunakan metode naive bayes untuk memberikan klasifikasi terhadap proses penerima Kartu Keluarga Sejahtera yaitu dengan menentukan Nama kepala keluarga, Umur, Jenis Kelamin, Status Perkawinan, Pendidikan, Pekerjaan, Pendapatan, Jumlah Tanggungan, Jenis Bangunan, dan Status Pemilik Rumah. Diharapkan data yang dihasilkan dari proses ini bisa digunakan sebagai bahan evaluasi terhadap proses penerima atau bukan penerima Kartu Keluarga Sejahtera.

Kata kunci: Penerapan Metode Naive Bayes; Klasifikasi Kartu Keluarga Sejahtera

Abstract: Kartu Keluarga Sejahtera is an assistance for underprivileged families that aims to help the poor, which is expected to ease the burden on society. Sumberjaya Village is one of the villages that has a total population of 73,631 people. The provisions in determining the status of Kartu Keluarga Sejahtera recipient are based on observations and interviews with the Head of the Sumberjaya Village Planning Affairs regarding the process of receiving Kartu Keluarga Sejahtera which is still not optimal in the data collection process so that there is still data Inaccurate in its acceptance, therefore it is necessary to conduct an assessment using the Naive Bayes method to give a classification to the process of receiving Kartu Keluarga Sejahtera, namely by determining age, gender, marital status, education, occupation, income, number of dependents, type of building and owner status. House. It is hoped that the data generated from this process can be used as material for evaluating the process of receiving / not receiving Kartu Keluarga Sejahtera.



DOI: 10.52362/jisicom.v6i1.821

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



Keywords: *Application of the Naive Bayes Method, Classification of a Kartu Keluarga Sejahtera*

I. PENDAHULUAN

Peran pemerintah diperlukan, untuk mengingat peraturan yang telah dibuat sebagai acuan di dalam kemajuan pembangunan. Program yang mengenai tentang mengurangi kemiskinan, yang dilakukan oleh pemerintah yaitu dengan penerbitan peraturan presiden republik indonesia Nomer 15 tahun 2010, yang berisikan mengenai cara mengurangi kemiskinan. Sejak diterbitkan peraturan presiden, dibentuklah sebuah tim nasional percepatan penanggulangan kemiskinan (TNP2K) di pusat, anggota anggota dari TNP2K berasal dari pemerintah, dunia usaha, dan lain nya. Pada tingkatan provinsi dan juga tingkat kabupaten, telah dibentuk sebuah tim koordinasi penanggulangan kemiskinan (TKPK). Dalam hal ini banyak program yang telah dilaksanakan dari pemerintah untuk meningkatkan kesejahteraan di berbagai bidang untuk memenuhi kebutuhan kehidupan masyarakat kurang mampu (miskin) [1]

Kemiskinan merupakan salah satu permasalahan yang ada di negara berkembang baik di indonesia maupun di luar negeri. Kartu Keluarga Sejahtera sendiri merupakan program pemerintah Presiden Jokowi Widodo yang merupakan bantuan bagi keluarga kurang mampu yang bertujuan untuk membantu masyarakat miskin yang diharapkan bisa meringankan beban masyarakat.

Kecamatan Tambun Selatan merupakan Kabupaten Bekasi yang terdiri dari 10 Desa dengan total penduduk 417.203 jiwa, Desa Sumberjaya merupakan salah satu desa yang memiliki total penduduk 73.631 jiwa dari total tersebut data yang kurang tepat dalam penentuan penerima Kartu Keluarga Sejahtera, sehingga terdapat kesalahan dalam penyaluran bantuannya, oleh karena itu perlu adanya metode dalam penentuan status Kartu Keluarga Sejahtera untuk membantu masyarakat miskin yang ada di wilayah Desa Sumberjaya. Berdasarkan dari uraian permasalahan yang ada maka peneliti akan menggunakan Metode Naive Bayes dalam klasifikasi penerima Kartu Keluarga Sejahtera, sehingga diharapkan dapat membantu pihak desa sumberjaya dalam mengelola data masyarakat untuk menentukan proses klasifikasi penerima Kartu Keluarga Sejahtera.

1.1. Kajian Pustaka

1. Keluarga Keluarga Sejahtera

Kartu Keluarga Sejahtera adalah program pemeritahan Presiden Jokowi Widodo yang merupakan bantuan dibidang ekonomi bagi keluarga kurang mampu yang sebelumnya merupakan program PKH(Program Keluarga Harapan) yang diharapkan kartu ini dapat meringankan beban masyarakat. Fungsi lain dari kartu keluarga sejahtera ini adalah sebagai penanda penyandang masalah kesejahteraan sosial (PMKS). [2]

2. Data Mining

Data mining merupakan pengumpulan informasi yang menarik dari beberapa data besar dengan menggunakan pengumpulan pola-pola yang ada di dalam data tersebut kemudian proses ini dapat dijadikan informasi. Data mining juga dapat disebut sebagai sistem pengolahan data yang sangat besar, yang memberikan peranan dalam beberapa bidang di dunia yaitu bidang keuangan, industri, transportasi, cuaca dan teknologi[3]

3. Naive Bayes



DOI: 10.52362/jisicom.v6i1.821

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



e-ISSN : 2597-3673 (Online) , p-ISSN : 2579-5201 (Printed)

Vol.6 No.1, June 202

Journal of Information System, Informatics and Computing

Website/URL: <http://journal.stmikjayakarta.ac.id/index.php/jisicom>

Email: jisicom@stmikjayakarta.ac.id , jisicom2017@gmail.com

Naive Bayes merupakan metode yang membagi permasalahan ke dalam sebuah kelas-kelas berdasarkan ciri-ciri persamaan dan perbedaan dengan menggunakan statistik yang bisa memprediksi probabilitas sebuah kelas. [4]

Metode probabilistik dan statistika yang di kemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa yang akan datang berdasarkan referensi kejadian waktu sebelumnya kemudian dikenal sebagai Teorema Bayes.[5]



DOI: 10.52362/jisicom.v6i1.821

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

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

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

Untuk menjelaskan metode Naive Bayes, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut, Karena itu, metode Naive Bayes di atas disesuaikan sebagai berikut:

$$P(C|F1 \dots Fn) = \frac{P(C)P(F1 \dots Fn|C)}{P(F1 \dots Fn)}$$

Dimana Variabel C merepresentasikan kelas, sementara variabel $F1 \dots Fn$

Merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (Posterior) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik-karakteristik sampel pada kelas C (disebut juga *likelihood*), dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (disebut juga *evidence*) Karena itu, rumus di atas dapat pula ditulis secara sederhana sebagai berikut :

$$Posterior = \frac{prior \times likelihood}{evidence}$$

4. Pengkelompokan Data Mining

Beberapa kelompok teknik data mining berdasarkan tugas yang dapat dilakukan yaitu [6] :

a. Deskripsi

Para peneliti terkadang mencoba untuk menemukan pola yang terdapat pada data.

b. Estimasi

Didalam estimasi hampir sama dengan klasifikasi, perkiraan nilai variabel biasanya lebih ke arah numerik atau kategorial

c. Klasifikasi

Pada klasifikasi variabel merupakan kategori, contohnya kita seperti dapat menentukan klasifikasi umur yang dibagi kedalam tiga kelas yaitu Dewasa, Lansia, dan Manula

d. Asosiasi

Menentukan antara keterkaitan pola peristiwa pada waktu yang sama

5. Penelitian Terkait

Penelitian ini membahas tentang metode naive bayes dalam menentukan Klasifikasi Status Kesejahteraan Rumah Tangga di Kabupaten Siak Menggunakan Algoritma Naive Bayes Classifier. Dimana yang menjelaskan mengenai status kesejahteraan di kabupaten siak dimana penelitian ini dapat digunakan secara efektif. Hasil akurasi yang didapatkan dalam menentukan Klasifikasi Status Kesejahteraan Rumah Tangga yaitu memiliki rata-rata sebanyak 53%. [7] . Penelitian ini merupakan upaya dalam penanggulangan kemiskinan yang ada pada Kecamatan Tibawa dalam Klasifikasi Masyarakat Miskin menggunakan Metode Naive Bayes dapat disimpulkan berdasarkan hasil pengujian *confusion matrix* memperoleh hasil akurasi sebanyak 73% dan nilai *Precision* sebanyak 92% serta *Recall* 86%. [8]

Penelitian Penerapan Data Mining untuk Mengklasifikasi Penerima Bantuan PKH Desa Wae Jare Menggunakan metode Naive Bayes bertujuan guna mengklasifikasi apakah Desa tersebut penerima atau



DOI: 10.52362/jisicom.v6i1.821

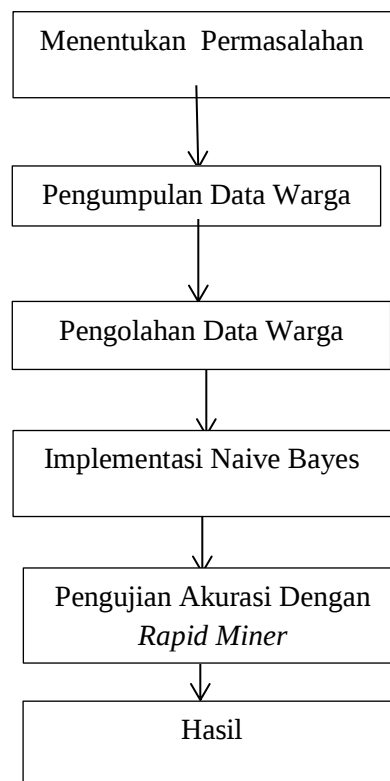
Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

bukan penerima bantuan PKH selain itu metode Naive Bayes di dalam penelitian ini terdapat 210 data dan menghasilkan akurasi 82,14% . [9]

II. METODE DAN MATERI

2.1. Tahapan Penelitian

Tahapan yang dilakukan dalam penelitian ini untuk menentukan status penerimaan Kartu Keluarga Sejahtera dengan menggunakan metode Naive Bayes. Berikut adalah langkah-langkah yang digunakan dalam tahapan proses yang akan dilaksanakan :



Gambar 1. Bagan Kerangka Tahap Penelitian

2.2. Instrumen Penelitian

Instrumen penelitian yang digunakan dalam penelitian ini berupa Data kualitatif dimana data tersebut merupakan data yang tidak menggunakan angka melainkan sebuah data yang berisi informasi. Informasi data tersebut merupakan data Warga Desa Sumberjaya Tambun Selatan.

2.3. Metode Pengumpulan Data

Dalam pengumpulan data-data yang dilakukan untuk penelitian informasi data diperoleh dengan melakukan beberapa hal yaitu :

1. Observasi

Melakukan observasi yang dilakukan secara langsung di kantor Desa Sumberjaya Tambun Selatan



DOI: 10.52362/jisicom.v6i1.821

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

2. Wawancara

Melakukan wawancara dengan Kepala Urusan kantor Desa Sumberja mengenai masalah sertamelakukan pengumpulan data dari kantor Desa Sumberjaya pada proses penerimaan Kartu Keluarga Sejahtera.

3. Studi Kepustakaan

Melakukan pengumpulan data melalui buku serta jurnal terkait dengan permasalahan yang diangkat sebagai acuan dalam penyelesaiannya.

2.4. Menentukan Populasi

Populasi merupakan objek dan subyek yang mempunyai kualitas dan karakteristik tertentu sehingga dapat menghasilkan suatu kesimpulan. Adapun populasi dari penelitian ini adalah wilayah Desa Sumberjaya Tambun Selatan.[10]

III. PEMBAHASA DAN HASIL

3.1. Menentukan Kriteria Penerima Kartu Keluarga Sejahtera

Dalam penentuan penerima kartu keluarga sejahtera dan bukan penerima kartu keluarga sejahtera terdapat beberapa kriteria yang digunakan di dalam penentuannya berdasarkan hasil observasi serta wawancara langsung oleh Kepala Urusan Perencanaan Kantor Desa Sumberjaya.

Kriteria yang digunakan dalam nilai perhitungan penerima kartu keluarga sejahtera adalah : Umur, Jenis Kelamin, Status Perkawinan, Status Pendidikan Terakhir, Pekerjaan, Pendapatan, Jumlah Tanggungan, Jenis Bangunan, Status Pemilik Rumah.

3.2. Perhitungan Tabel Probabilitas

Pada perhitungan probabilitas kelas yaitu dengan mentransformasi data set yang diolah menjadi data training kemudian menghitung jumlah seluruh data yang dikategorikan menjadi 2 label yaitu : penerima dan bukan penerima. Berikut adalah hasil dari perhitungan tabel probabilitas kelas :

Tabel Probabilitas			
Label	Jumlah Data	Jumlah Seluruh Data	Hasil
Penerima	402	600	0,669230769
Bukan Penerima	198	600	0,330769231

Gambar 2. Tabel Probabilitas Kelas

$P(C_i)$

$P(\text{Penerima}) = 402/600 = 0,669230769$

$P(\text{Bukan Penerima}) = 198/600 = 0,330769231$

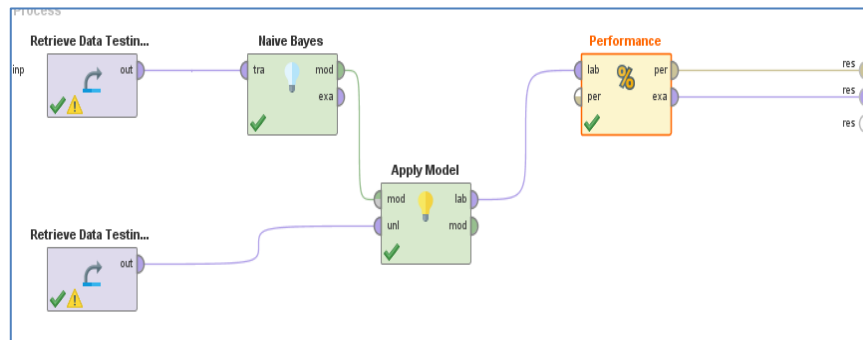
3.3 Hasil Perhitungan



DOI: 10.52362/jisicom.v6i1.821

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

Berdasarkan dari hasil penelitian yang telah dilakukan mengenai penerapan metode *Naive Bayes* dalam memprediksi penerima Kartu Keluarga Sejahtera dapat disimpulkan bahwa metode ini dapat digunakan dalam penentuan nilai penerima Kartu Keluarga Sejahtera dimana dalam penentuannya terdapat 9 kategori kriteria penting di dalamnya yaitu Umur, Jenis Kelamin, Status perkawinan, Status Pendidikan, Pekerjaan, Pendapatan, Jumlah Tanggungan, Jenis Bangunan serta Status Pemilik rumah. dengan menggunakan metode ini salahsatunya dapat dipahami dan menghasilkan hasil akurasi yang tepat berdasarkan data uji yang dilakukan dengan menggunakan 600 Kepala Keluarga untuk data training dan 120 Kepala Keluarga untuk data testing yang dipilih secara acak dengan menggunakan *Rapid Miner* maka hasilnya adalah sebagai berikut :



Gambar 3. Desain Model Rapid Miner

accuracy: 99.17%			
	true Penerima	true Bukan Penerima	class precision
pred. Penerima	80	1	98.77%
pred. Bukan Penerima	0	39	100.00%
class recall	100.00%	97.50%	

Gambar 4. Hasil

Pada perhitungan tersebut maka dapat dihasilkan akurasi penerima sebesar 100% dan bukan penerima 97,50%

IV. KESIMPULAN

Berdasarkan dari hasil penelitian yang telah dilakukan mengenai penerapan metode *Naive Bayes* dalam memprediksi penerima Kartu Keluarga Sejahtera dapat disimpulkan bahwa metode ini dapat digunakan dalam penentuan nilai penerima Kartu Keluarga Sejahtera dimana dalam penentuannya terdapat 9 kategori kriteria penting di dalamnya yaitu Umur, Jenis Kelamin, Status perkawinan, Status Pendidikan, Pekerjaan, Pendapatan, Jumlah Tanggungan, Jenis Bangunan serta Status Pemilik rumah. Dengan menggunakan metode ini salah satunya dapat dipahami dan menghasilkan hasil akurasi yang tepat berdasarkan data uji yang dilakukan dengan menggunakan 600 Kepala Keluarga untuk data training dan 120 Kepala Keluarga untuk data testing yang dipilih secara acak, dari pengujian yang dilakukan dengan aplikasi Rapid Miner hasil uji tingkat akurasi yang didapat sebesar 99% dan 1% error.



DOI: 10.52362/jisicom.v6i1.821

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



REFERENSI

- [1] A. Anggleni, "Implementasi Kebijakan Program Kartu Keluarga Sejahtera (KKS) dalam Meningkatkan Kesejahteraan Masyarakat Miskin di Kelurahan Sekip Jaya Kecamatan Kemuning Kota Palembang," *J. PPS UNISTI*, vol. 1, no. 1, pp. 24–39, 2018.
- [2] M. A. Suhendra, D. Ispriyanti, and S. Sudarno, "Ketepatan Klasifikasi Pemberian Kartu Keluarga Sejahtera Di Kota Semarang Menggunakan Metode Regresi Logistik Biner Dan Metode Chaid," *J. Gaussian*, vol. 9, no. 1, pp. 64–74, 2020.
- [3] A. N. Putri, "Penerapan Naive Bayesian Untuk Perankingan Kegiatan Di Fakultas Tik Universitas Semarang," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 8, no. 2, p. 603, 2017.
- [4] S. Lestari and M. Badrul, "Implementasi Klasifikasi Naive Bayes Untuk Prediksi," vol. 7, no. 1, pp. 8–16, 2020.
- [5] I. W. Supriana, "Pengembangan Sistem Identifikasi Keluarga Miskin Di Kabupaten Tabanan Dengan Menggunakan Metode Naïve Bayes Dalam Pengentasan Kemiskinan," *JST (Jurnal Sains dan Teknol.*, vol. 7, no. 1, p. 102, 2018.
- [6] C. D. L. Daniel T.Larose, *Discovering Knowledge in Data : An Introduction to Data Mining*, 2nd, illustr ed. John Wiley & Sons, 2014, 2014.
- [7] A. Tarigan, E. Wahyudi, and J. Adhiva, "Klasifikasi Status Kesejahteraan Rumah Tangga di Kabupaten Siak Menggunakan Algoritma Naive Bayes Classifier," no. November, pp. 187–196, 2019.
- [8] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 160–165, 2018.
- [9] M. Algoritma, B. Bri, and U. Anduonohu, "Penerapan Data Mining untuk Mengklasifikasi Pola Nasabah," vol. 1, no. 1, pp. 22–28, 2019.
- [10] Sugiyono, "Metode Penelitian Pendidikan pendekatan Kuantitatif, Kualitatif dan R&D.," in *METODE PENELITIAN ILMIAH*, 2014.

