

IMPLEMENTASI SISTEM PENDUKUNG KEPUTUSAN PENENTUAN KELAYAKAN CALON MITRA BADAN PUSAT STATISTIK MENGGUNAKAN METODE NAIVE BAYES

IMPLEMENTATION OF DECISION SUPPORT SYSTEM FOR DETERMINING THE ELIGIBILITY OF POTENTIAL PARTNERS OF CENTRAL BUREAU OF STATISTICS USING NAIVE BAYES METHOD

Melani Vebtania Nomseo^{1§}

¹Universitas Nusa Cendana dan Jl. Matani Raya, Nusa Tenggara Timur [mnomseo@gmail.com]

[§]*Corresponding Author*

Received 20th May 2024; Accepted 19th Nov 2024; Published 30th Jun 2025;

Abstrak

Badan Pusat Statistik (BPS) adalah lembaga resmi yang bertanggung jawab untuk menyelenggarakan kegiatan statistik di tingkat Nasional di Indonesia. Penelitian ini bertujuan untuk mengetahui nilai akurasi yang baik dan tepat pada kriteria-kriteria yang mempengaruhi kelayakan calon mitra BPS. Metode pengklasifikasian yang digunakan yaitu metode naïve bayes. Hasil dari penelitian ini menunjukkan kriteria yang mempengaruhi kelayakan calon mitra BPS diterima adalah umur dari 18-30 tahun dan hasil prediksi perhitungan Naive Bayes dengan menggunakan RapidMiner dengan jumlah data yang diuji 1.611 data, Jumlah data yang diprediksi benar "Diterima" 660 data, Jumlah data yang diprediksi benar "Tidak Diterima" 403 data, Jumlah data yang diprediksi salah "Diterima" 452 data, Jumlah data yang diprediksi salah "Tidak Diterima" 96 data. Dengan akurasi sebesar 65.98% dan tingkat eror sebesar 34.02%. Maka sangat cocok untuk memprediksi peluang dimasa depan.

Kata Kunci: BPS, Naïve Bayes, Nilai akurasi, Prediksi

Abstract

The Central Statistics Agency (BPS) is the official institution responsible for carrying out statistical activities at the national level in Indonesia. This research aims to determine good and precise accuracy values for the criteria that influence the suitability of prospective BPS partners. The classification method used is the Naive Bayes Method. The results of this research show that the criteria that influence the suitability of prospective BPS partners to be accepted are age from 18-30 years and the predicted results of Naive Bayes calculations using RapidMiner with the amount of data tested being 1.611 data, the number of data to be correctly predicted "Accepted" 660 data, the amount of data predicted correctly "Not Accepted" 403 data, Number of data predicted incorrectly "Accepted" 452 data, Number of data predicted incorrectly "Not Accepted" 96 data. With an accuracy of 65.98 % and an error rate of 34.02 %. So it is very suitable for predicting future opportunities.

Keywords: BPS, Naïve Bayes, Accuration value, Prediction

1. Pendahuluan

Badan Pusat Statistik (BPS) adalah lembaga resmi yang bertanggung jawab untuk menyelenggarakan kegiatan statistik di tingkat Nasional di Indonesia. BPS mengumpulkan data statistik melalui sensus, survei, dan sumber data lainnya sesuai dengan perkembangan ilmu pengetahuan dan teknologi. Dalam kegiatan BPS memerlukan mitra untuk bekerjasama maka dilakukanlah perekrutan calon mitra oleh BPS. Calon mitra harus mengetahui kriteria-kriteria apa saja yang dinilai dalam perekrutan kelayakan calon mitra karena banyaknya jumlah pelamar namun belum tentu memenuhi kriteria dalam suatu instansi [1]. Oleh karena itu dalam setiap tahapan perekrutan calon mitra BPS harus diikuti sesuai tahapan yang ada diantaranya seleksi administrasi, tes kompetensi, dan seleksi akhir.

Permasalahan muncul pada saat pelaksanaan penilaian kelayakan calon mitra BPS. Berdasarkan data penilaian kelayakan calon mitra BPS terdapat temuan diantaranya calon mitra yang dinyatakan diterima di BPS tidak sesuai kriteria-kriteria yang diberikan seperti sebagian calon mitra yang diterima di BPS tidak berdomisili di wilayah pendaftaran. Secara umum, munculnya perilaku di atas dipengaruhi oleh faktor internal dan faktor eksternal. Faktor internal berasal dari diri pegawai itu sendiri seperti kepentingan diri sendiri karena calon mitra yang mendaftar adalah keluarganya maka diterima. Sedangkan faktor eksternal berasal dari luar diri pegawai misalnya calon mitra yang mendaftar adalah mitra senior yang sudah sangat sering mengikuti kegiatan BPS sehingga calon

mitra tersebut diterima tanpa melihat kriteria yang ada [3].

BPS Kabupaten Kupang belum pernah melakukan penilaian calon mitra BPS menggunakan Metode Naive Bayes. Penilaian kelayakan calon mitra selama ini dilakukan secara subjektif oleh koordinator fungsi berdasarkan pengamatan langsung dan masukan dari pegawai organik BPS lainnya. Penilaian secara subjektif tentunya tidak ideal dijadikan acuan penilaian kelayakan calon mitra karena dasar yang digunakan setiap orang berbeda-beda. Untuk mempermudah BPS dalam menentukan dan mengklasifikasikan kelayakan calon mitra BPS dengan baik, maka dalam penelitian ini menerapkan sebuah metode pengklasifikasian untuk mengatasi permasalahan tersebut. Metode yang digunakan adalah metode Naive Bayes.

Metode Naive Bayes adalah salah satu metode klasifikasi dalam statistika dan machine learning yang berdasarkan pada Teorema Bayes Sa'adawiyah [5]. Dalam konteks pengambilan keputusan kelayakan mitra, konsep dasar Naive Bayes melibatkan perhitungan probabilitas kondisional. Konsep Dasar Naive Bayes adalah Teorema Bayes, Asumsi Naif [6], probabilitas Kelayakan Mitra. Tulisan ini bertujuan untuk mengetahui Kriteria dan nilai akurasi penentuan kelayakan calon mitra Badan Pusat Statistik. Dari hasil yang diperoleh penelitian ini dapat memberikan masukan kepada pihak BPS Kabupaten Kupang agar tercapainya calon mitra yang potensial.

2. Landasan Teori

2.1. Metode Naïve Bayes

Naive Bayes adalah algoritma pembelajaran terawasi berdasarkan Teorema Bayes yang digunakan untuk menyelesaikan masalah klasifikasi menggunakan metode probabilitas. Naive Bayes diusulkan oleh ilmuwan Inggris Thomas Bayes, yang secara khusus memprediksi peluang masa depan berdasarkan eksperimen sebelumnya, itulah sebabnya disebut Teorema Bayes. Teorema Naive Bayes dapat dinyatakan sebagai berikut [7]:

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

Keterangan:

X	: Data kelas yang tidak diketahui
C	: Hipotesis data-data X adalah kelas tertentu
$P(C X)$	: Probabilitas hipotesis C berdasarkan keadaan (<i>posterior probability</i>)
$P(C)$	: Probabilitas hipotesis C (sebelum kemungkinan)
$P(X C)$	: Probabilitas X berdasarkan kondisi pada hipotesis C
$P(X)$	: Probabilitas X

Selanjutnya penjabaran $P(X_1, \dots, X_n|C)$ dapat disederhanakan menjadi Persamaan:

$$P(X_1, \dots, X_n|C) = P(X_1|C)P(X_2|C), \dots, P(X_n|C)$$

$$P(X_1, \dots, X_n|C) = \prod_{i=1}^n P(X_i|C)$$

Dalam metode Naive Bayes diperlukan data latih dan data uji yang ingin diklasifikasikan, semakin banyak data latih yang dilibatkan, semakin baik hasil prediksi yang diberikan. Menghitung $P(C)$ yang

merupakan probabilitas prior untuk setiap kelas C yang akan dihasilkan menggunakan Persamaan:

$$P(C) = \frac{SI}{S}$$

Keterangan:

SI	: Jumlah data dari kategori C
S	: Jumlah total data

2.2. Distribusi Posterior

Distribusi posterior adalah fungsi kepadatan bersyarat θ jika diketahui nilai pengamatan X . Dapat ditulis sebagai berikut:

$$f(\theta|X) = \frac{f(\theta, X)}{f(x)} \quad (2.2)$$

Jika θ kontinu, distribusi posterior θ dapat dinyatakan dengan rumus fungsi kepadatan. Fungsi kepadatan bersyarat dari variabel acak tunggal, jika diketahui bahwa nilai variabel acak kedua hanyalah fungsi dari kepadatan sambungan dua variabel acak yang dibagi dalam fungsi kepadatan n marjinal variabel acak tersebut. Tetapi fungsi densitas bersama $f(\theta, X)$ dan fungsi kepadatan marginal $f(X)$.

3. Hasil Dan Pembahasan

Sebelum proses perhitungan, terlebih dahulu dilakukan pengumpulan data. Pengumpulan data dilakukan secara tidak langsung atau data diperoleh dari pihak lain yaitu Badan Pusat Statistik Kabupaten Kupang.

Data yang akan digunakan adalah data Kriteria calon mitra yaitu umur (X_1), Pendidikan (X_2), Alamat Tempat Tinggal (X_3), Pernah Mengikuti Sensus/Survei Sebelumnya (X_4), Nilai Ujian (X_5). Data ini diperoleh dari Badan Pusat Statistik Kabupaten Kupang sebanyak 1.611 data, yang diterima sebagai mitra BPS sebanyak 1.112 orang dan yang tidak diterima sebanyak 499 orang.

3.1. Perhitungan Naïve Bayes

1. Perhitungan Probabilitas Prior ($P(C)$)

Data akan diproses dengan Metode Naïve Bayes dengan beberapa tahapan sehingga perlu menentukan besaran nilai dari data, yaitu 1.611 data yang digunakan, diketahui kelas C_0 (diterima) sebanyak 1.112 data, dan kelas C_i (tidak diterima) sebanyak 499 data.

Perhitungan Probabilitas Prior	Hasil
$P(C_0)$	0.690255
$P(C_i)$	0.309745
Total	100

Berdasarkan hasil perhitungan diatas menghasilkan dua kelas yaitu kelas diterima (C_0) dengan nilai probabilitas 0.690255 atau 69,0255% dan kelas tidak diterima (C_i) dengan nilai probabilitas 0.309745 atau 30,9745% sehingga ketika kedua kelas tersebut dijumlahkan akan menghasilkan nilai 100 %.

2. Perhitungan Probabilitas Posterior X Bersyarat $C(P(X|C))$

Untuk menghitung probabilitas posterior bersyarat maka digunakan jumlah kriteria dengan kelas “Diterima”, kemudian bagi dengan probabilitas kelas “Diterima”. Begitu juga dengan mencari probabilitas untuk kelas “Tidak diterima”.

X_1	Jumlah Kejadian		Probabilitas	
	Diterima	Tidak Diterima	Diterima	Tidak Diterima
Umur				
18-30	833	412	0.749101	0.825651
31-50	276	86	0.248201	0.712345
>50	3	1	0.002698	0.002004
Jumlah	1.112	499	1	1

Pendidikan	Diterima	Tidak diterima	Diterima	Tidak diterima
SD-SMP	1	0	0.000899	0
SMA-D1/D2/D3	478	192	0.429856	0.427617
D4/S1	626	303	0.562950	0.563474
S2	7	4	0.006295	0.008909
Jumlah	1.112	499	1	1

Alamat	Diterima	Tidak diterima	Diterima	Tidak diterima
Perkotaan	477	283	0.428957	0.567134
Pedesaan	513	182	0.461331	0.068136
Pinggiran Kota	122	34	0.109712	0.068136
Jumlah	1.112	499	1	1

Nilai Ujian	Diterima	Tidak diterima	Diterima	Tidak diterima
00-30	69	319	0.062050	0.639279
31-50	176	31	0.158273	0.062124
51-70	315	61	0.283273	0.12224
71-90	457	70	0.410971	0.140281
>90	95	18	0.085431	0.036072
Jumlah	1.112	499	1	1

3. Menghitung Probabilitas Akhir Untuk Setiap Kelas

Untuk menghitung probabilitas akhir pada setiap kelas yaitu dengan memasukan semua data calon mitra BPS Kabupaten Kupang yang ada kedalam satu kelas yang sama, selanjutnya dari masing-masing atribut dan nilai probabilitas kelas tersebut dikalikan. Dari kedua hasil yang sudah ditentukan pada tiap kelas, kemudian bandingkan nilai yang paling tinggi. Jika kelas “Diterima” bernilai paling tinggi maka hasilnya “Diterima” begitu pula sebaliknya.

4. Kasus Perhitungan Naïve Bayes

Contoh kasus perhitungan Naive Bayes dengan

variabelnya X sebagai berikut:

Umur X1 = 18-30

Pendidikan X2 = SD-SMP

Alamat X3 = Perkotaan

Mengikuti Sensus X4 = 0

Nilai Ujian X5 = 00-30

apakah calon mitra tersebut bisa diterima di BPS atau tidak?

Penyelesaian:

Tahap 1. Menghitung jumlah kelas P(C)

Jumlah data "Diterima" dibagi jumlah total data

$$P(C0) = \frac{1.112}{1.611} = 0.690255$$

Jumlah data "Tidak Diterima" dibagi jumlah total data.

$$P(Ci) = \frac{499}{1.611} = 0.309746$$

Tahap 2. Menghitung jumlah kasus yang sama dengan kelas yang sama yaitu:

$$P(X|C) \quad P(X|C0)$$

$$P(C0) = 1.112$$

$$P(X1|C0) = \frac{833}{1.112} = 0.749101$$

$$P(X2|C0) = \frac{1}{1.112} = 0.000899$$

$$P(X3|C0) = \frac{477}{1.112} = 0.428957$$

$$P(X4|C0) = \frac{537}{1.112} = 0.482914$$

$$P(X5|C0) = \frac{69}{1.112} = 0.062050$$

Mengalikan semua hasil atribut "Diterima"

$$P(X1, \dots, Xn) = P(X1|C), P(X2|C), \dots, P(Xn|C)$$

$$\begin{aligned} P(X1|Ci) &= 0.749101 \cdot 0.000899 \cdot 0.428957 \\ &\quad \cdot 0.482914 \cdot 0.062050 \\ &= 0.000009 \end{aligned}$$

Berdasarkan jumlah perhitungan semua hasil atribut yang diterima ketika dikalikan maka

diperoleh hasil akhirnya yaitu 0.000009.

$$P(Ci) = 499$$

$$P(X1|Ci) = \frac{412}{499} = 0.825651$$

$$P(X2|Ci) = \frac{303}{499} = 0$$

$$P(X3|Ci) = \frac{182}{499} = 0.567134$$

$$P(X4|Ci) = \frac{129}{499} = 0.741483$$

$$P(X5|Ci) = \frac{70}{499} = 0.639279$$

$$P(X1|Ci) = 0.825651 \cdot 0 \cdot 0.567134 \cdot 0.741483 \cdot 0.639279 = 0.000000$$

Berdasarkan jumlah perhitungan semua hasil atribut yang tidak diterima ketika dikalikan maka diperoleh hasil akhirnya yaitu 0.000000.

Tahap ke 3. Membandingkan nilai kelas "Diterima (C0)" dan "Tidak Diterima (Ci)" untuk kelas diterima (C0) dengan cara mengalikan $P(X|C0)$ dengan $P(C0)$

$$P(C0|X) = P(X|C0) \cdot P(C0)$$

$$= (0.000009) \cdot (0.690255) = 0.000006$$

Berdasarkan perbandingan perhitungan jumlah semua hasil atribut yang diterima dan tidak diterima ketika dikalikan dengan hasil akhir atribut diterima maka diperoleh hasil 0.000006. Kemudian untuk kelas "Tidak Diterima (Ci)" adalah dengan cara mengalikan $P(X|Ci)$ dengan $P(Ci)$:

$$P(Ci|X) = P(X|Ci) \cdot P(Ci)$$

$$= (0.000000) \cdot (0.309746)$$

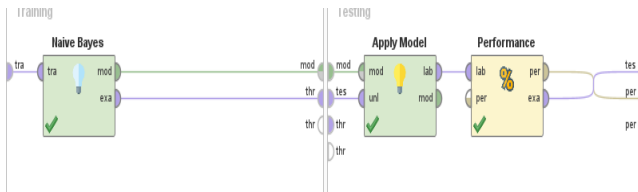
$$= 0.000000$$

Berdasarkan hasil perhitungan di atas dapat diketahui bahwa nilai $P(X|C0)$ adalah 0.000006 sedangkan nilai $P(X|Ci)$ adalah 0.000000. Artinya nilai $P(X|C0)$ lebih besar dari pada nilai

$P(X|C_i)$ sehingga dapat disimpulkan bahwa untuk kasus perhitungan Naive Bayes masuk kedalam kelas "Diterima" yang artinya dapat diterima sebagai Calon Mitra di BPS.

3.2. Perhitungan Rapidminer Studio

Perhitungan ini dilakukan untuk memprediksi apakah Calon Mitra Badan Pusat Statistik tersebut diterima sebagai Mitra BPS atau tidak diterima dengan menggunakan Naive Bayes. Untuk mencari keakuratan validasi data yang digunakan. Berikut merupakan proses validasi data menggunakan rapidminer studio



Dengan Menggunakan RapidMiner studio didapatkan hasil prediksi bahwa jumlah data yang diprediksi benar"diterima" adalah 660 orang, jumlah data yang diprediksi benar"tidak diterima" adalah 403 orang, jumlah data yang diprediksi salah"diterima" adalah 452 orang, jumlah data yang diprediksi salah"tidak diterima" adalah 96 orang. Dari data yang diteliti jumlah calon mitra yang diterima di BPS hanya sebanyak 1.112 orang berarti tidak sesuai dengan hasil prediksi dari aplikasi RapidMiner studio dikarenakan formasi yang disiapkan di BPS hanya 1.112 orang saja sehingga dari 660 orang hanya diterima sebanyak 1.112 orang dan sudah disesuaikan dengan honor yang akan diberikan oleh masing-masing orang yang akan bekerja di BPS. Untuk menghitung nilai akurasi sebagai berikut:

Jumlah data yang di uji: 1.611

Jumlah data yang diprediksi benar"Diterima": 660

Jumlah data yang diprediksi benar "Tidak Diterima" :403

Jumlah data yang diprediksi salah "Diterima": 452

Jumlah data yang diprediksi salah "Tidak Diterima": 96

$$\begin{aligned} \text{Accuracy} &= \frac{\text{Jumlah data yang diprediksi benar}}{\text{Jumlah data yang diuji}} \cdot 100\% \\ &= \frac{(660 + 403)}{(660 + 403 + 452 + 96)} \cdot 100\% \\ &= 65.98\% \end{aligned}$$

$$\begin{aligned} \text{Error} &= \frac{\text{Jumlah data yang diprediksi salah}}{\text{Jumlah data yang diuji}} \cdot 100\% \\ &= \frac{(452 + 96)}{(660 + 403 + 452 + 96)} \cdot 100\% \\ &= 34.02\% \end{aligned}$$

Dari hasil perhitungan tersebut dapat disimpulkan bahwa klasifikasi dengan menggunakan metode naive bayes untuk menentukan penerimaan calon mitra badan pusat statistik (BPS) menghasilkan tingkat Accuracy sebesar 65.98% dan tingkat error 34.02%.

4. Kesimpulan Dan Saran

Berdasarkan analisis dan hasil pembahasan yang telah dipaparkan mengenai Penentuan Kelayakan Calon Mitra Badan Pusat Statistik di BPS Kabupaten Kupang dengan menggunakan Metode Naive Bayes, maka diperoleh kesimpulan bahwa Dari 5 kriteria yang digunakan, kriteria yang mempengaruhi penerimaan calon mitra BPS di Kabupaten Kupang adalah umur dimana kisaran umur yang diterima sebagai mitra di BPS

Kabupaten Kupang adalah umur 18-30 tahun yang paling tepat diterima menjadi mitra di BPS Kabupaten Kupang dan Hasil prediksi perhitungan Naive Bayes dengan menggunakan Rapid Miner Studio dengan jumlah data yang diuji sebanyak 1.611 data, dengan nilai akurasi sebesar 65.98% dan tingkat eror sebesar 34.02%. Maka dari perhitungan tersebut sangat cocok diterapkan dalam memprediksi peluang dimasa depan sehingga memudahkan para pegawai BPS dalam menentukan penerimaan calon mitra BPS Kabupaten Kupang.

5. Ucapan Terima Kasih

Ucapan terimakasih yang sebesar-besarnya kepada Universitas Nusa Cendana yang turut membantu, membimbing dan mendukung pelaksanaan kegiatan penelitian ini.

Daftar Pustaka

- [1] Y. Yumhi, "Implementasi Rekrutmen Di Bps Kabupaten Lebak (Studi Kasus Rekrutmen Tenaga Sensus)," *The Asia Pacific Journal of Management Studies*, vol. 7, no. 3, 2020.
- [2] S. Mubarak, S. Zauhar, E. Setyowati, and S. Suryadi, "Policy implementation analysis: exploration of george edward iii, marilee s grindle, and mazmanian and sabatier theories in the policy analysis triangle framework," *JPAS (Journal of Public Administration Studies)*, vol. 5, no. 1, pp. 33–38, 2020.
- [3] D. Nofriansyah and S. Defit, *Multi Criteria Decision Making (MCDM) pada sistem pendukung keputusan*. Deepublish, 2017.
- [4] S. Baker *et al.*, "Teaching Academic Content and Literacy to English Learners in Elementary and Middle School. IES Practice Guide. NCEE 2014-4012.," *What Works Clearinghouse*, 2014.
- [5] I. M. A. A. Pramana, I. W. Sudiarsa, and P. G. S. C. Nugraha, "Penerapan Algoritma Naïve Bayes Untuk Prediksi Penjualan Produk Terlaris Pada CV Akusara Jaya Abadi," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 10, no. 4, 2023.
- [6] E. Manalu, F. A. Sianturi, and M. R. Manalu, "Penerapan Algoritma Naive Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan Dan Jumlah Pemesanan Pada Cv. Papadan Mama Pastries," *Jurnal Mantik Penusa*, vol. 1, no. 2, pp. 16–21, 2017.
- [7] N. Maulidah, R. Supriyadi, D. Y. Utami, F. N. Hasan, A. Fauzi, and A. Christian, "Prediksi Penyakit Diabetes Melitus Menggunakan Metode Support Vector Machine dan Naive Bayes," *Indonesian Journal on Software Engineering (IJSE)*, vol. 7, no. 1, pp. 63–68, 2021.