



## Menganalisa Tingkat Sosial Ekonomi Calon Pengantin Pria Yang Terdapat Di Kabupaten Bondowoso Dengan Metode Algoritma *Naïve Bayes Classifier*

Tito Alif Firdaus\*, Lutfi Ali Muharom, Daryanto

Teknik Informatika, Universitas Muhammadiyah Jember

Email: [titoalif23@gmail.com](mailto:titoalif23@gmail.com)\*, [lutfi.muharom@unmuhjember.ac.id](mailto:lutfi.muharom@unmuhjember.ac.id), [daryanto@unmuhjember.ac.id](mailto:daryanto@unmuhjember.ac.id)

### ABSTRAK

Pernikahan sebuah cara yang dipilih Allah SWT untuk manusia agar berkembang dan melestarikan hidupnya. Keharmonisan sebuah keluarga yaitu seluruh anggota keluarga dapat berhubungan secara serasi dan seimbang, saling memuaskan dan memenuhi kebutuhan anggota keluarga yang lainnya dan memperoleh pemuasan atas segala kebutuhannya agar dapat merasakan kebahagiaan. Pernikahan dini terjadi dikarenakan pergaulan bebas yang tidak terkendali, seperti hamil di luar nikah dan alasan ekonomi. Penerapan algoritma *Naive Bayes Classifier* merupakan sebuah algoritma atau cara untuk mencari dan membuat keputusan masalah tersebut dengan memperhitungkan berbagai macam faktor yang ada di dalam lingkup masalah. Algoritma *Naïve Bayes* sangat menarik untuk dibahas karena dari penelitian sebelumnya algoritma ini mampu memanfaatkan data training untuk menghasilkan probabilitas setiap kriteria untuk kelas yang berbeda seperti melakukan prediksi minat beli pelanggan terhadap kartu internet yang baru akan diluncurkan dengan melakukan prediksi berdasarkan kartu internet yang sudah beredar sebelumnya. Untuk penelitian ini memiliki data sebanyak 238 dan di proses secara klasifikasi data menggunakan algoritma *Naïve Bayes Gaussian* menghasilkan angka persentase di akurasi 57%, presisi 20,8% dan recall 86%. Sehingga sebagai calon pengantin pria dapat dikatakan layak di klasifikasi akurasi dan recall, sedangkan di klasifikasi presisi calon pengantin pria masih belum cukup untuk dikatakan layak dalam hal sosial ekonominya.

**Kata Kunci:** Pernikahan Dini, DPPKB, Metode *Naïve Bayes*

### ABSTRACT

*Marriage is a way chosen by Allah SWT for humans to develop and preserve their lives. Harmony in a family means that all family members can relate in harmony and balance, mutually satisfy and fulfill the needs of other family members and obtain satisfaction of all their needs in order to feel happiness. Early marriage occurs due to uncontrolled promiscuity, such as pregnancy out of wedlock and economic reasons. The application of the Naive Bayes Classifier algorithm is an algorithm or way to search and make decisions on the problem by taking into account various factors within the scope of the problem. The Naïve Bayes algorithm is very interesting to discuss because from previous research this algorithm is able to utilize training data to produce probabilities for each criterion for different classes, such as predicting customer buying interest in new internet cards that will be launched by making predictions based on internet cards that have been circulating previously. For this study, we had 238 data and processed the data classification using the Naïve Bayes Gaussian algorithm resulting in percentage figures of 57% accuracy, 20.8% precision and 86% recall. So as a prospective groom, he can be said to be worthy in the accuracy and recall classification, while in the precision classification the prospective groom is still not enough to be said to be worthy in terms of his socio-economic status*

**Keywords:** Early Marriage, DPPKB, *Naïve Bayes Method*.

## 1. PENDAHULUAN

Pernikahan sebuah cara yang dipilih Allah SWT untuk manusia agar berkembang dan melestarikan hidupnya. Demi menjaga kehormatan dan martabat manusia, Allah mengadakan hukumsesuai dengan martabatnya, sehingga hubungan laki-laki dan perempuan diatur secara terhormat dan berdasarkan rasa saling meridhoi dengan ijab kabul yang dihadiri oleh keluarga dan saksi yang menyaksikan bahwa pasangan laki-laki dan perempuan tersebut sudah menjadi keluarga yang harmonis.

Keharmonisan sebuah keluarga yaitu seluruh anggota keluarga dapat berhubungan secara serasi dan seimbang, saling memuaskan dan memenuhi kebutuhan anggota keluarga yang lainnya dan memperoleh pemuasan atas segala kebutuhannya agar dapat merasakan kebahagiaan (Suharman, 2014). Di beberapa daerah, terutama di Kabupaten Bondowoso yang terletak di Provinsi Jawa Timur, terdapat beberapa lokasi yang di tuju seperti Desa Tlogosari, Wringin, Botolinggo, Maesan, dan

Cermee. Pernikahan dini terjadi dikarenakan pergaulan bebas yang tidak terkendali, seperti hamil di luar nikah dan alasan ekonomi. Perempuan yang masih belum cukup umur yang melakukan pernikahan dini sering dipaksa putus sekolah karena dapat dikatakan status sosialnya yang lebih rendah di keluarga, seorang suami yang tidak bisa mengontrol reproduksinya akan berdampak pada kesehatan perempuan muda yang melakukan pernikahan dini.

Di Kabupaten Bondowoso, jumlah kasus pernikahan di usia yang belum cukup memumpuni atau anak yang belum berumur 20 tahun dalam hitungan waktu 5 tahun belakang mengalami kenaikan dan penurunan. Menurut data dari SEKDA-DPPKB (Dinas Pemberdayaan Perempuan dan Keluarga Berkualitas) Kabupaten Bondowoso, Kasus usia wanita pertama kali menikah dibawah umur 20 tahun pada tahun 2017 sejumlah 2157 orang, dan pada tahun 2018 mengalami kenaikan menjadi 2670 orang, namun pada tahun 2019 mengalami penurunan menjadi 1999 orang dan terus menurun di tahun 2020 menjadi 1131 orang. Penurunan kasus ini tidak lepas dari usaha yang dilakukan oleh DPPKB Kabupaten Bondowoso seperti peningkatan dan penguatan sosialisasi untuk pencegahan perkawinan dini, melakukan kerjasama (MoU) bersama lintas sektoral dan lain-lainnya. Maka dengan adanya penelitian ini diharapkan kesadaran akan pentingnya pengaruh kesehatan terhadap calon pengantin itu sendiri bertambah dan dapat mengurangi jumlah angka kasus pernikahan dini pada tahun-tahun berikutnya.

## 2. KAJIAN PUSTAKA

### A. Calon Pengantin

Materi pembekalan yang telah disampaikan pada kursus calon pengantin, kepala KUA beserta jajarannya yang dilaksanakan di KUA Dukun pada tanggal 02 Juli 2021 sangatlah penting untuk para calon pengantin yang akan mengawali sebuah pernikahan dengan harus memahami ilmunya. Jika semua calon pengantin mampu mengaplikasikan dalam menjalankan pernikahannya, maka sangat bermanfaat sekali untuk pernikahan agar membentuk keluarga yang sakinah. Materi yang diberikan dalam bimbingan para calon pengantin juga harus sesuai dengan tujuan pernikahan yakni membentuk keluarga yang sakinah (Dina, 2021). Bimbingan untuk calon pengantin sangat berperan besar dalam langkah awal untuk memahami pernikahan karena bimbingan tersebut berisi tentang membangun rumah tangga yang tidak hanya untuk individu calon pengantin dan pasangannya saja, melainkan untuk anak keturunan mereka (Aminah, 2022).

### B. *Naïve Bayes Classifier*

*Naive Bayes Classifier* merupakan metode yang cocok untuk klasifikasi biner dan *multiclass*. Algoritma *Naive Bayes Classifier* merupakan salah satu algoritma yang terdapat pada teknik klasifikasi *Naive bayes* berdasarkan asumsi sederhananya nilai atribut secara kondisional yang saling bebas jika diberi nilai output. Diberi nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu atau bentuk umum dari teorema Bayes sebagai berikut :

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

Keterangan :

Data dengan class yang belum diketahui

Hipotesis data X merupakan suatu class spesifik

$P(Y|X)$  : Probabilitas akhir bersyarat (*conditional probability*) suatu hipotesis Y terjadi jika diberikan bukti (*evidence*) X terjadi.

$P(Y)$  : Probabilitas awal (*priori*) hipotesis Y terjadi tanpa mengandung bukti apapun.

$P(X|Y)$  : Probabilitas sebuah bukti X terjadi akan mempengaruhi hipotesis Y.

$P(X)$  : Probabilitas awal (*priori*) bukti X terjadi tanpa mengandung hipotesis / bukti yang lain

Jika  $X$  merupakan vektor masuk berupa fitur dan  $Y$  merupakan label kelas, maka *Naïve Bayes* ditulis  $P(Y|X)$ . Notasi tersebut merupakan probabilitas kelas  $Y$  didapatkan setelah fitur  $X$  diamati. Notasi ini disebut juga probabilitas akhir untuk  $Y$ , sedangkan  $P(Y)$  disebut probabilitas awal (*prior probability*)  $Y$ .

$$P(X|Y) = \prod_{i=1}^n P(X_i|Y) \quad (2)$$

Keterangan:

$P(X|Y)$  : Probabilitas data dengan vektor  $X$  pada kelas  $Y$ .

$P(Y)$  : Probabilitas awal kelas  $Y$ .

$\prod_{i=1}^n P(X_i|Y)$  : Probabilitas independen kelas  $Y$  dari semua fitur dalam vektor  $X$

### C. K-Fold Cross Validation

*Cross-validation* dapat juga disebut sebagai estimasi rotasi yang dimaksud adalah Teknik validasi model untuk menilai bagaimana hasil statistik analisis akan menggeneralisasikan kumpulan data yang independen. Dalam teknik ini diutamakan karena digunakan untuk memprediksi model dan juga memperkirakan keakuratan sebuah model prediktif dimana Ketika jalan nya praktikum. Dalam teknik validasi silang adalah salah satunya K-fold cross, memecah data  $K$  yang mana bagian dari dataset dengan ukuran yang sama. Dimana dalam penggunaan k-fold validation untuk menghilangkan bias pada dataset tersebut. Dalam pelatihan dan pengujian sebanyak  $K$  kali. Dalam percobaan pertama, subset  $S_1$  Pelatihan dan pengujian dilakukan sebanyak  $k$  kali. Pada percobaan pertama, subset  $S_1$  dilakukan sebagai data pengujian data set lainnya yang dilakukan sebagai data pelatihan, pada percobaan kedua subset  $S_1, S_3, \dots, S_k$  menjadi data pelatihan dan  $S_2$  menjadi data pengujian, dan begitu seterusnya (Tempola et al., 2018).

## 3. METODE PENELITIAN

### A. Jenis dan Sumber Data

Data pada penelitian merupakan data sekunder yang diperoleh dari kantor BKKBN kabupaten bondowoso pada tahun 2021. Data yang digunakan yaitu data kuantitatif berupa data calon pengantin dengan variabel yang digunakan yaitu jenis kelamin pria, pekerjaan, sumber penghasilan, dan nilai penghasilan. Pada proses pengumpulan data ini diperoleh sejumlah 238 data.

### B. Metode Penelitian

Metode analisis yang digunakan untuk mencapai tujuan penelitian dalam penulisan Tugas Akhir ini diuraikan sebagai berikut:

1. Mengumpulkan data kemudian diinput ke dalam sebuah file bertipe csv (*microsoft excel*)
2. Melakukan proses *preprocessing* data untuk menyiapkan data agar bisa diolah. Proses yang dilakukan pada tahap ini berupa *Data Cleaning* dan *Data Transformation*
3. Mengidentifikasi data dengan melakukan analisis deskriptif pada data
4. Hasil *preprocessing* data diinput ke dalam database program *RStudio*
5. Klasifikasi dengan *Naïve Bayes Classifier*
6. Melakukan evaluasi model dengan melihat pengujian *K-Fold Cross Validation* dengan  $k = 2, 4, 5, 8$
7. Interpretasi dan Kesimpulan

## 4. HASIL DAN PEMBAHASAN

### A. Preprocessing Data

Tahap *preprocessing data* dilakukan sebelum data tersebut digunakan pada proses pengujian. Pada proses ini akan dilakukan pembersihan data atribut yang dianggap tidak perlu atau dianggap sebagai noise. Tidak semua atribut dalam tabel data digunakan, hanya atribut yang dianggap berpengaruh saja yang akan digunakan dalam penelitian. Pada data Calon Pengantin penentuan label diberikan pada kategori Resiko Rendah (0), Resiko Tinggi (1). Atribut yang akan digunakan pada proses pengujian

yaitu kondisi sosial ekonomi pria diantaranya Pekerjaan, Sumber Penghasilan, Nilai penghasilan, Umur dan pendidikan.

a) Hasil pengujian menggunakan *Naive bayes Gaussian* K 2-fold Cross Validation

Akurasi, presisi dan recall diperoleh dari hasil pengujian Algoritma *Naive Bayes Gaussian* dengan menggunakan 2-fold. Berikut merupakan gambaran pengujian pada 2-fold skenario :

Tabel :

Data Uji	
Data Latih	

  

Skenario 1	Data Uji	Data Latih
Skenario 2	Data Latih	Data Uji

Perhitungan Akurasi, Presisi, dan Recall :

$$\text{Akurasi} = \frac{0,33 + 0,82}{2} \times 100\% = \frac{1,15}{2} \times 100\% = 0,57 \times 100\% = 57\%$$

$$\text{Presisi} = 0,19 \times 100\% = 19\%$$

$$\text{Recall} = 0,86 \times 100\% = 86\%$$

b) Hasil pengujian menggunakan *naive bayes Gaussian* K 4-fold Cross Validation

Dari total 200 data dalam pengujian menggunakan k-fold 4 skenario kedua dengan perhitungan pada Algoritma *naive bayes gaussian* dilakukan menggunakan nilai akurasi yaitu sebesar 37,75 %, presisi 12,5 % dan recall 67 %. Berikut merupakan gambaran pada 4-fold skenario :

Tabel :

Skenario 1				
Skenario 2				
Skenario 3				
Skenario 4				

Perhitungan Akurasi, Presisi, dan Recall :

$$\text{Akurasi} = \frac{0,2 + 0,32 + 0,17 + 0,82}{4} \times 100\% = \frac{1,51}{4} \times 100\% = 0,3775 \times 100\% = 37,75\%$$

$$\text{Presisi} = \frac{0,17 + 0,16 + 0,17}{4} \times 100\% = \frac{0,5}{4} \times 100\% = 0,125 \times 100\% = 12,5\%$$

$$\text{Recall} = \frac{1 + 1 + 0,71}{4} \times 100\% = \frac{2,71}{4} \times 100\% = 0,67 \times 100\% = 67\%$$

c) Hasil Pengujian Menggunakan *Naive Bayes Gaussian* K 5-fold Cross Validation

Dari total 200 data dalam pengujian menggunakan k-fold 5 skenario kedua dengan perhitungan pada Algoritma *naive bayes gaussian* dilakukan menggunakan nilai akurasi yaitu sebesar 35,8 %, presisi 20,8 % dan recall 52,6 %.Berikut merupakan gambaran pada 5-fold skenario.

Tabel :

Skenario 1					
Skenario 2					
Skenario 3					
Skenario 4					
Skenario 5					

Perhitungan Akurasi, Presisi, dan Recall:

$$\text{Akurasi} = \frac{0,15 + 0,31 + 0,31 + 0,21 + 0,81}{5} \times 100\% = \frac{1,79}{5} \times 100\% = 0,358 \times 100\% = 35,8\%$$

$$\text{Presisi} = \frac{0,15 + 0,16 + 0,19 + 0,19}{5} \times 100\% = \frac{1,04}{5} \times 100\% = 0,208 \times 100\% = 20,8\%$$

$$\text{Recall} = \frac{1 + 0,8 + 0,83}{5} \times 100\% = \frac{2,63}{5} \times 100\% = 0,526 \times 100\% = 52,6\%$$

d) Hasil Pengujian Menggunakan *Naïve Bayes Gaussian* K 8-fold Cross Validation

Dari total 200 data dalam pengujian menggunakan k-fold 8 skenario kedua dengan perhitungan pada Algoritma *naïve bayes gaussian* dilakukan menggunakan nilai akurasi yaitu sebesar 28,1 %, presisi 13,1 % dan recall 79 %. Berikut merupakan gambaran pada 8-fold skenario keempat :

Tabel :

Skenario 1								
Skenario 2								
Skenario 3								
Skenario 4								
Skenario 5								
Skenario 6								
Skenario 7								
Skenario 8								

Perhitungan Akurasi, Presisi, dan Recall :

$$\text{Akurasi} = \frac{0,15 + 0,25 + 0,3 + 0,25 + 0,25 + 0,25 + 0,2 + 0,8}{8} \times 100\%$$

$$= \frac{2,25}{8} \times 100\%$$

$$= 0,281 \times 100\%$$

$$= 28,1\%$$

$$\text{Presisi} = \frac{0,15 + 0,16 + 0,125 + 0,21 + 0,21 + 0,2}{8} \times 100\%$$

$$= \frac{1,055}{8} \times 100\%$$

$$= 0,131 \times 100\%$$

$$= 13,1\%$$

$$\text{Recall} = \frac{1 + 1 + 0,66 + 0,66 + 1 + 1 + 1}{8} \times 100\% = \frac{6,32}{8} \times 100\% = 0,79 \times 100\% = 79\%$$

## 5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian pada Klasifikasi Sosial Ekonomi Calon Pengantin Pria menggunakan Algoritma *Naïve Bayes Gaussian* maka dapat disimpulkan bahwa :

- Dari percobaan yang dilakukan, algoritma *Naïve Bayes Gaussian* mampu mengklasifikasikan data Sosial Ekonomi calon pengantin pria dengan tingkat keberhasilan sekitar 57% pada pengaturan 2 fold, dari segi presentase akurasi yang di dapat maka calon pengantin pria dapat dikatakan layak dalam hal pekerjaan, sumber penghasilan dan nilai penghasilan.
- Untuk hasil nilai presisi di 5 fold mendapatkan nilai sebesar 20,8%. Bahwasannya untuk calon pengantin pria bisa dikatakan masih belum cukup banyak seorang pria yang bisa dikatakan layak sebagai calon pengantin.
- Dalam kasus ini, algoritma *Naïve Bayes Gaussian* mencapai tingkat recall tertinggi sebesar 86% pada pengaturan 2 fold. Artinya, dari semua calon pengantin pria yang sebenarnya termasuk dalam kategori Sosial Ekonomi tertentu, algoritma berhasil mengidentifikasi 86% dari mereka.

## 6. DAFTAR PUSTAKA

- Nurainun, Yusuf, A.M. (2022) Analisis Tingkat Kesiapan Menikah Calon Pengantin. *Jurnal Ilmu Pendidikan*, 4(2), 2110–2115. <https://doi.org/10.31004/edukatif.v4i2.2345>.
- Najah, U., Desyanty, E.S., dan Widiyanto, E. (2021). Kontribusi Program Pembinaan Calon Pengantin Terhadap Kesiapan Berumah Tangga Bagi Masyarakat Kota Malang. *Aksara: Jurnal Ilmu Pendidikan Nonformal*, 7(3), 1303-1312. <https://doi.org/10.37905/aksara.7.3.1303-1312.2021>.
- Eska, J. (2016). Penerapan Data Mining Untuk Prekdiksi Penjualan Wallpaper Menggunakan Algoritma C4.5 STMIK Royal Ksianan. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 2, 9–13.
- Aminah, S. (2022). Efektivitas Bimbingan Calon Pengantin Dalam Membentuk Keluarga Harmonis Di Kua Kecamatan Dukun. *Journal UNUGIRI*, 5(1), 19-28.
- Khaerani, S. N. (2019). Faktor Ekonomi Dalam Pernikahan Dini Pada Masyarakat Sasak Lombok. *Qawwam*, 13(1), 1–13. <https://doi.org/10.20414/qawwam.v13i1.1619>
- Messerli, F. H., & Devereux, R. B. (1983). Introduction: Left ventricular hypertrophy- Good orevil? *The American Journal of Medicine*, 75(3 Part A), 1–3. [https://doi.org/10.1016/0002-9343\(83\)90110-9](https://doi.org/10.1016/0002-9343(83)90110-9)
- Najah, U., Desyanty, E. S., & Widiyanto, E. (2021). Kontribusi Program Pembinaan Calon Pengantin Terhadap Kesiapan Berumah Tangga Bagi Masyarakat Kota Malang.
- Nur Khormarudin, A. (2016). Teknik Data Mining: Algoritma K-Means Clustering. *Jurnal Ilmu Komputer*, 1–12. <https://ilmukomputer.org/category/datamining/>
- Nur Khormarudin, A. (2016). Teknik Data Mining: Algoritma K-Means Clustering. *Jurnal Ilmu Komputer*, 1–12. <https://ilmukomputer.org/category/datamining/>
- Saputra, H. K. (2018). Analisis Data Mining Untuk Pemetaan Mahasiswa Yang Membutuhkan Bimbingan Dan Konseling Menggunakan Algoritma *Naïve Bayes Classifier*. *Jurnal Teknologi Informasi Dan Pendidikan*, 11(1), 14–26. <https://doi.org/10.24036/tip.v11i1.104>.