

Klasisifikasi Tingkat Kemiskinan Di Indonesia Menggunakan Algoritma Naives Bayes

Suci Mulyani¹, Afril Efan Pajri², Fikram^{1*}

¹ Program Studi Ilmu Komputer, Universitas Muhammadiyah Bima, Bima, Indonesia

² Program Studi Sistem Komputer, Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro, Indonesia

* Email Korespondensi: fikramum41@gmail.com

Abstrak: Abstrak penelitian ini dilatarbelakangi oleh urgensi klasifikasi tingkat kemiskinan yang akurat sebagai landasan pengambilan keputusan dan penyaluran bantuan yang tepat sasaran. Penelitian ini bertujuan untuk membangun model klasifikasi tingkat kemiskinan menggunakan algoritma Naïve Bayes yang diimplementasikan dalam *software* RapidMiner. Metode penelitian yang digunakan meliputi pengumpulan data dari berbagai sumber, pra-pemrosesan data untuk membersihkan dan mentransformasi data, implementasi algoritma Naïve Bayes untuk membangun model klasifikasi, serta evaluasi kinerja model menggunakan metrik akurasi, presisi, dan *recall*. Hasil penelitian menunjukkan bahwa model yang dibangun mencapai akurasi sebesar 96.12%, dengan kemampuan identifikasi yang baik pada kelompok tidak miskin (presisi 100%) dan kelompok miskin (*recall* 100%). Meskipun demikian, presisi pada kelompok miskin masih perlu ditingkatkan (75%) untuk meminimalkan kesalahan klasifikasi dan memastikan bantuan yang diberikan lebih efektif dan tepat sasaran.

Kata Kunci: Klasifikasi Kemiskinan, Naïve Bayes, RapidMiner, Akurasi, Machine Learning.

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



1. Pendahuluan

Kemiskinan merupakan salah satu permasalahan mendasar yang dihadapi oleh banyak negara berkembang, termasuk Indonesia. Data dari Badan Pusat Statistik (BPS) menunjukkan bahwa tingkat kemiskinan di Indonesia masih menjadi isu signifikan, dengan jutaan penduduk hidup di bawah garis kemiskinan. Kondisi ini tidak hanya berdampak pada aspek ekonomi, tetapi juga memengaruhi kualitas hidup masyarakat secara keseluruhan, seperti akses terhadap pendidikan, kesehatan, dan infrastruktur dasar. Oleh karena itu, diperlukan upaya yang sistematis dan berbasis data untuk mengidentifikasi pola-pola kemiskinan guna mendukung pengambilan kebijakan yang lebih efektif [1].

Dalam beberapa tahun terakhir, pendekatan berbasis pembelajaran mesin (*machine learning*) telah menjadi metode yang banyak digunakan untuk memecahkan berbagai masalah sosial-ekonomi, termasuk dalam klasifikasi tingkat kemiskinan. Salah satu algoritma yang populer adalah Naïve Bayes, sebuah metode probabilistik yang sederhana namun sangat efektif untuk menangani masalah klasifikasi. Algoritma ini bekerja berdasarkan Teorema Bayes dengan asumsi independensi antar fitur, sehingga memungkinkan pengolahan data yang kompleks dengan efisiensi tinggi. Keunggulan utama dari Naïve Bayes terletak pada kemampuannya untuk menangani dataset besar dan memberikan hasil klasifikasi yang akurat meskipun

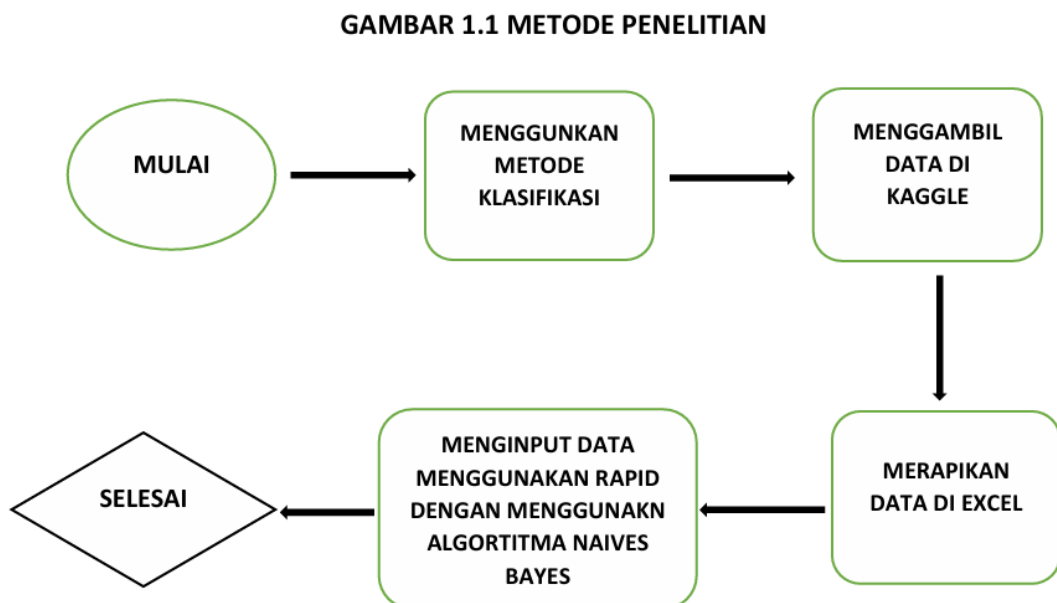
menggunakan model probabilistik yang relatif sederhana [2], [3]. Penelitian sebelumnya menunjukkan bahwa algoritma Naïve Bayes telah berhasil diterapkan dalam berbagai studi terkait pengentasan kemiskinan. Misalnya, penelitian oleh Nurmayanti menunjukkan bahwa algoritma ini mampu mengklasifikasikan status ekonomi rumah tangga dengan akurasi tinggi berdasarkan indikator-indikator seperti pendapatan bulanan, jumlah anggota keluarga, tingkat pendidikan kepala keluarga, dan akses terhadap fasilitas dasar [4]. Selain itu, penelitian lain di Papua mencatat bahwa penggunaan Naïve Bayes dalam klasifikasi rumah tangga miskin menghasilkan akurasi sebesar 80% [5]. Penelitian ini juga menunjukkan bahwa atribut seperti jenis kelamin, pendidikan kepala rumah tangga, dan sumber air minum berkontribusi signifikan terhadap klasifikasi status ekonomi [6].

Penelitian ini bertujuan untuk mengklasifikasikan tingkat kemiskinan di Indonesia menggunakan algoritma Naïve Bayes dengan memanfaatkan data sosial-ekonomi yang relevan. Dengan pendekatan ini, diharapkan dapat dihasilkan model prediktif yang mampu memberikan wawasan mendalam mengenai distribusi kemiskinan di berbagai wilayah Indonesia. Model ini tidak hanya akan membantu dalam pemetaan tingkat kemiskinan secara lebih akurat tetapi juga dapat menjadi alat bantu bagi pemerintah dan pemangku kepentingan lainnya dalam merancang kebijakan pengentasan kemiskinan yang lebih terarah dan berbasis data.

Melalui penelitian ini, kontribusi signifikan diharapkan dapat diberikan dalam pengembangan metode analisis data untuk mendukung pembangunan berkelanjutan di Indonesia. Dengan memanfaatkan teknologi berbasis machine learning seperti Naïve Bayes, potensi pemanfaatan data besar (big data) dalam konteks sosial-ekonomi dapat dioptimalkan untuk menciptakan solusi inovatif terhadap permasalahan kemiskinan.

2. Metode

Metode yang saya gunakan adalah metode klasifikasi yang menggunakan algoritma naives bayes untuk mengelompokkan data tingkat kemiskinan di Indonesia. Dengan tujuan untuk mengetahui tingkat kemiskinan di Indonesia di berbagai macam provinsi yang ada di Indonesia. Adapun langkah penelitian yang dilakukan dapat dilihat pada gambar 1:



Gambar 1: Metode penelitian

2.1. Rapid Miner

Perangkat lunak intelijen lain yang terbuka adalah Rapidminer (open source). Salah satu alat untuk data mining, text mining dan analisis prediktif disebut RapidMiner. Sekitar 500 operator data mining, termasuk yang untuk input, output, preprocessing, dan visualisasi,

tersedia di RapidMiner. RapidMiner adalah perangkat lunak yang beroperasi secara independen dan dapat dimasukkan ke dalam produk lain untuk melakukan analisis data dan penambangan. Karena RapidMiner ditulis dalam bahasa Java, itu kompatibel dengan semua sistem operasi [7].

2.2. Klasifikasi

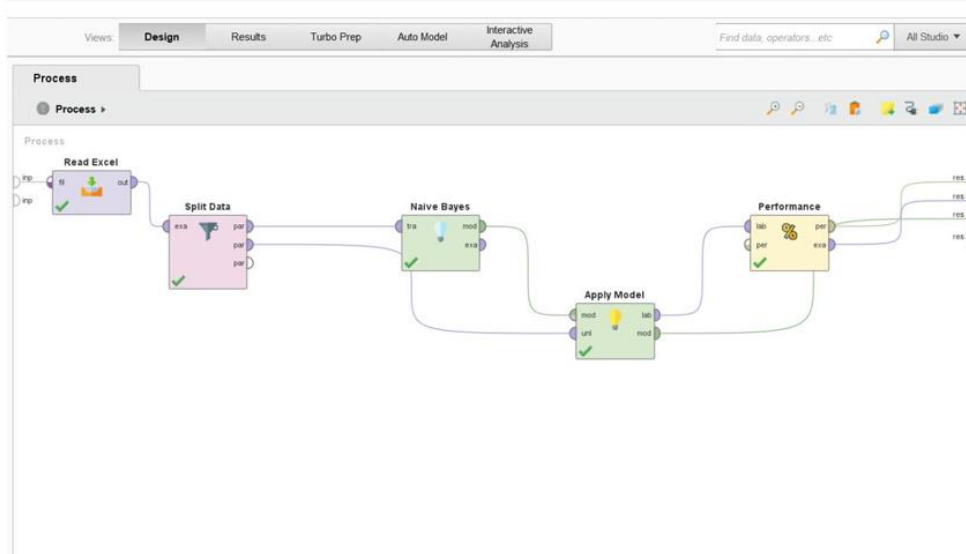
Proses penemuan model (atau fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui. Algoritma klasifikasi yang banyak digunakan secara luas, yaitu Decision/classification trees, Bayesian classifiers/ Naïve Bayes classifiers, Neural networks, Analisa Statistik, Algoritma Genetika, Rough sets, k-nearest neighbor, Metode Rule Based, Memory based reasoning, dan Support vector machines [8].

2.2. Algoritma Naive Bayes

Algoritma Naïve Bayes adalah metode klasifikasi berbasis probabilitas yang sederhana namun efektif, bekerja dengan menggunakan Teorema Bayes untuk memprediksi peluang suatu kelas berdasarkan atribut yang diamati. Dalam penelitian ini, data yang digunakan diperoleh dari berbagai sumber, termasuk survei sosial-ekonomi dari Badan Pusat Statistik (BPS) dan data sekunder yang relevan. Sebelum analisis, data akan melalui tahap pra-pemrosesan, yang mencakup pembersihan data, transformasi atribut kategorikal menjadi numerik, dan normalisasi data. Fitur-fitur yang digunakan dalam klasifikasi mencakup pendapatan bulanan, tingkat pendidikan kepala keluarga, jumlah anggota keluarga, status pekerjaan, dan akses terhadap fasilitas dasar [9].

3. Hasil Penelitian

Hasil dari penelitian mengenai klasifikasi tingkat kemiskinan di Indonesia menggunakan algoritma Naïve Bayes. Pemodelan dilakukan dengan menggunakan *software* RapidMiner, sebuah platform analisis data yang memungkinkan visualisasi proses dan implementasi berbagai algoritma *machine learning*, proses pemodelan dengan rapid miner dapat ditunjukkan pada gambar 1.



Gambar 1. Pemodelan dengan Rapid Miner

Gambar 1 menunjukkan alur pemodelan yang digunakan dalam penelitian ini. Proses dimulai dengan pembacaan data dari file Excel, diikuti dengan pembagian data menjadi dua bagian, yaitu data pelatihan (*training set*) dan data pengujian (*testing set*). Selanjutnya, algoritma Naïve Bayes diterapkan pada data pelatihan untuk membangun model klasifikasi. Model yang telah dilatih kemudian diterapkan pada data pengujian untuk memprediksi tingkat kemiskinan. Terakhir, kinerja model

dievaluasi menggunakan metrik evaluasi yang sesuai untuk mengukur seberapa baik model tersebut mampu mengklasifikasikan tingkat kemiskinan.

Row No.	Klasifikasi Kemiskinan	prediction(K...	confidence(...	confidence(...	Provinsi	Kab/Kota	Persentase ...	Rata-rata La...	Pengeluaran...	In...
1	No	No	1.000	0.000	ACEH	Bireuen	13.250	9.290	8867	72
2	No	No	1.000	0.000	ACEH	Kota Banda Ac...	7.610	12.830	16891	85
3	No	No	1.000	0.000	ACEH	Kota Langsa	10.960	11.120	12067	77
4	No	No	1.000	0.000	SUMATERA U...	Labuhan Batu	8.740	9.250	11212	72
5	No	No	1.000	0.000	SUMATERA U...	Karo	8.790	10	12412	74
6	No	No	1	0	SUMATERA U...	Deli Serdang	4.010	10.100	12291	75
7	No	No	0.958	0.042	SUMATERA U...	Pakpak Bharat	9.350	9.140	8254	67
8	No	No	1.000	0.000	SUMATERA U...	Padang Lawas...	9.920	9.380	10055	70
9	No	No	1.000	0.000	SUMATERA U...	Kota Sibolga	12.330	10.410	11540	73
10	No	No	1.000	0.000	SUMATERA B...	Pesisir Selatan	7.920	8.270	9270	70
11	No	No	1.000	0.000	SUMATERA B...	Sijunjung	6.800	8.120	10389	67
12	No	No	1.000	0.000	SUMATERA B...	Kota Solok	3.120	11.040	12168	78

Gambar 2. Transformasi Data

Gambar 2 menampilkan *ExampleSet* hasil transformasi data setelah proses *Apply Model*. Ini menunjukkan hasil prediksi model Naïve Bayes beserta tingkat kepercayaan (*confidence*) untuk setiap instance data. Kolom-kolom seperti "Klasifikasi Kemiskinan," "prediction(Klasifikasi Kemiskinan)," kolom *confidence*, serta informasi geografis (Provinsi dan Kab/Kota) dan fitur-fitur lainnya (Persentase, Rata-rata, Pengeluaran) memberikan gambaran detail mengenai hasil klasifikasi dan faktor-faktor yang memengaruhi prediksi.

4. Evaluasi

Evaluasi kinerja model dalam bentuk matriks kebingungan (*confusion matrix*) dan metrik akurasi. Dari gambar ini, kita dapat melihat akurasi keseluruhan model, presisi, dan *recall* untuk setiap kelas (Miskin dan Tidak Miskin). Hasil akurasi ini akan dianalisis lebih lanjut untuk memahami seberapa baik model mampu mengklasifikasikan tingkat kemiskinan.

accuracy: 96.12%

	true No	true Yes	class precision
pred. No	87	0	100.00%
pred. Yes	4	12	75.00%
class recall	95.60%	100.00%	

Gambar 3. Akurasi Prediksi

Gambar 3. Menunjukkan Nilai akurasi sebesar 96.12% menunjukkan bahwa model Naïve Bayes yang diimplementasikan di RapidMiner memiliki kemampuan yang sangat baik dalam mengklasifikasikan tingkat kemiskinan. Matriks kebingungan memberikan detail tambahan, di mana 87 instance yang sebenarnya "Tidak Miskin" berhasil diprediksi sebagai "Tidak Miskin" (True Negatives), dan 12 instance yang sebenarnya "Miskin" berhasil diprediksi sebagai "Miskin" (True Positives). Namun, terdapat 4 instance yang sebenarnya "Miskin" tetapi salah diprediksi sebagai "Tidak Miskin" (False Negatives), sementara tidak ada instance yang sebenarnya "Tidak Miskin" yang salah diprediksi sebagai "Miskin" (False Positives). Class recall sebesar 95.60% untuk kelas "Tidak Miskin" menunjukkan bahwa model berhasil mengidentifikasi sebagian besar instance "Tidak Miskin", tetapi masih ada sedikit yang

terlewatkan. Sementara itu, class recall untuk kelas "Miskin" adalah 100%, menunjukkan bahwa model berhasil mengidentifikasi seluruh instance yang sebenarnya "Miskin".

5. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan menggunakan RapidMiner dan algoritma Naïve Bayes, dapat disimpulkan bahwa model memiliki kemampuan yang sangat baik dalam mengklasifikasikan tingkat kemiskinan dengan akurasi 96.12%. Model ini sangat andal dalam mengidentifikasi kelompok masyarakat yang tidak miskin, dengan presisi mencapai 100%. Selain itu, model juga mampu mengidentifikasi seluruh kelompok masyarakat yang miskin (recall 100%), meskipun terdapat potensi untuk mengurangi kesalahan klasifikasi dalam memprediksi kelompok masyarakat miskin (presisi 75%) agar bantuan yang diberikan lebih tepat sasaran. Dengan demikian, algoritma Naïve Bayes dapat menjadi alat yang efektif dalam membantu pemerintah dan organisasi lain dalam mengidentifikasi dan menargetkan populasi yang membutuhkan intervensi dan program bantuan.

6. References

- [1] W. P. Nurmayanti, "Penerapan Naive Bayes Dalam Mengklasifikasikan Masyarakat Miskin Di Desa Lepak," *Geodika J. Kaji. Ilmu Dan Pendidik. Geogr.*, Vol. 5, No. 1, Pp. 123–132, Jun. 2021, Doi: 10.29408/Geodika.V5i1.3430.
- [2] D. Latipah, M. A. Mardiansah, And E. Adityarini, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes : Studi Kasus Di Desa Jati Mulya Kecamatan Sepatan Timur," 2024.
- [3] A. C. Putri, F. E. Hariyanto, N. L. E. Andini, And Z. C. S. Zulkarnaen, "Klasifikasi Rumah Tangga Miskin Di Provinsi Papua Tahun 2017 Menggunakan Metode Naive Bayes," *J. Sains Mat. Dan Stat.*, Vol. 7, No. 1, P. 89, Mar. 2021, Doi: 10.24014/Jsms.V7i1.11924.
- [4] C. Fuadi Ahmad, N. Suarna, And G. Dwilestari, "Klasifikasi Data Kemiskinan Menggunakan Metode Naïve Bayes Untuk Mengetahui Tingkat Kemiskian Studi Kasus: Desa Karangasem Kecamatan Leuwimunding Majalengka," *J. Inform. Dan Teknol. Inf.*, Vol. 2, No. 2, Pp. 203–208, Nov. 2023, Doi: 10.56854/Jt.V2i2.190.
- [5] A. Fatmawati, A. Irma Purnamasari, And I. Ali, "Implementasi Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sosial," *Jati J. Mhs. Tek. Inform.*, Vol. 8, No. 1, Pp. 745–750, Mar. 2024, Doi: 10.36040/Jati.V8i1.8714.
- [6] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, Vol. 10, No. 2, Pp. 160–165, Aug. 2018, Doi: 10.33096/Ilkom.V10i2.303.160-165.
- [7] "45090-Article Text-123090-1-10-20210314.Pdf."
- [8] A. Pebdika, R. Herdiana, And D. Solihudin, "Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima Pip," *Jati J. Mhs. Tek. Inform.*, Vol. 7, No. 1, Pp. 452–458, Mar. 2023, Doi: 10.36040/Jati.V7i1.6303.
- [9] A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," Vol. 2, No. 3, 2015.