

Data Mining Dengan Metode Naïves Bayes Classifier dalam Memprediksi Tingkat Kepuasan Pelayanan Dokumen Kependudukan

Susi Wirma^{1✉}

¹Independent Researcher

susiwirma@gmail.com

Abstract

Community satisfaction with population document services is very important in improving the quality of services as desired by the community at the Population and Civil Registration Service. The purpose of this study is to predict the level of community satisfaction with service quality and to find out the results of the accuracy that has been obtained using the Naïve Bayes algorithm. The processed data is population service data in 2022. From the results of the processed data, it can be used to improve the quality of population document services and can also be used to evaluate the performance of its services. This is done to determine the extent to which the value and quality of the services provided by the Population and Civil Registration Office of Pariaman City to the community. The Naïve Bayes method used in this study is to try to predict community satisfaction with service quality at the Department of Population and Civil Registration. The results of the satisfaction index with the service of population documents at the Department of Population and Civil Registration using the Naïve Bayes method get good results.

Keywords: Data Mining, Nave Bayes, Classification, satisfaction index, resident documents

Abstrak

Kepuasan masyarakat terhadap layanan dokumen kependudukan merupakan hal yang sangat penting dalam peningkatan kualitas layanan sesuai yang diinginkan oleh masyarakat pada Dinas Kependudukan dan Pencatatan Sipil. Tujuan dari Penelitian ini adalah untuk memprediksi tingkat kepuasan masyarakat terhadap kualitas layanan dan untuk mengetahui hasil dari akurasi yang telah didapatkan dengan menggunakan algoritma Naïve Bayes. Data yang diolah merupakan data layanan kependudukan Tahun 2022. Dari hasil data yang diolah, bisa digunakan untuk meningkatkan kualitas dalam pelayanan dokumen kependudukan dan juga bisa digunakan untuk mengevaluasi kinerja pelayanannya. Hal ini dilakukan untuk mengetahui sejauh mana nilai dan kualitas dari pelayanan yang diberikan oleh Dinas Kependudukan dan Pencatatan Sipil Kota Pariaman kepada Masyarakat. Metode Naïve Bayes yang digunakan dalam penelitian ini untuk mencoba memprediksi kepuasan masyarakat terhadap kualitas pelayanan pada Dinas Kependudukan dan Pencatatan Sipil. Hasil dari indeks kepuasan terhadap pelayanan dokumen kependudukan pada Dinas Kependudukan dan Pencatatan Sipil menggunakan metode Naïve Bayes mendapatkan hasil yang baik.

Kata kunci: Data Mining, Naïve Bayes, Klasifikasi, indeks kepuasan, dokumen kependudukan.

INFEB is licensed under a Creative Commons 4.0 International License.



1. Pendahuluan

Perkembangan teknologi informasi telah memberikan kontribusi pada cepatnya pertumbuhan jumlah data yang dikumpulkan dan disimpan dalam basis data yang berukuran besar [1]. Data tersebut dapat diekstrak menjadi sebuah pengetahuan dengan menggunakan Knowledge Discovery In Database (KDD). KDD adalah suatu proses menggali dan menganalisis sejumlah data dan menghasilkan informasi dan pengetahuan yang berguna [2]. KDD merupakan satu metode yang bisa diterapkan dalam melakukan proses Data Mining [3]. Data Mining adalah suatu proses analisa data untuk memperoleh informasi yang bermakna dan berguna dari data yang tersimpan dalam database [4]. Data Mining bisa juga diartikan sebagai suatu proses yang memakai satu atau lebih teknik pembelajaran yang dilakukan secara komputerisasi yang bertujuan untuk menganalisa dan

mengembangkan suatu ilmu pengetahuan dari data yang ada pada database [5].

Data Mining adalah suatu algoritma yang digunakan untuk menggali informasi yang berharga yang tersembunyi dalam database yang besar sehingga ditemukan suatu pola yang menarik yang sebelumnya tidak diketahui [6]. Salah satu algoritma yang akan digunakan pada Data Mining adalah algoritma Naïve Bayes Classifier [7]. Algoritma Naïve Bayes merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. [8]. Naïve Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes.

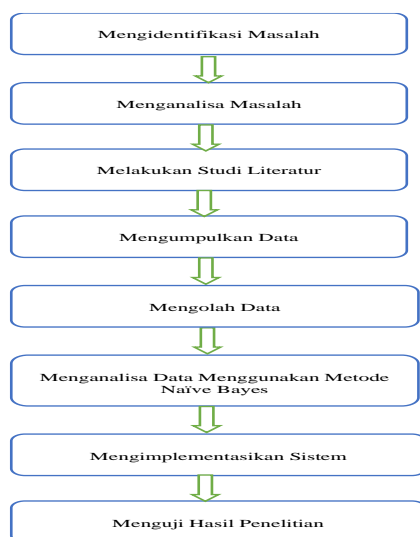
Pada penelitian terdahulu, di mana metode Naïve Bayes digunakan untuk memprediksi kondisi sarana air minum memiliki tingkat akurasi 85,05% [9]. Penelitian

yang lain dalam klasifikasi tingkat kepuasan siswa terhadap pembelajaran daring. Data yang digunakan merupakan hasil observasi terhadap siswa SMA (Sekolah Menengah Atas) di Kabupaten Tulungagung, di mana hasil klasifikasi menunjukkan bahwa algoritma Naïve Bayes cocok digunakan untuk mengukur tingkat kepuasan siswa terhadap pembelajaran daring dengan tingkat akurasi terbesarnya adalah 100% dengan nilai precision sebesar 100% dan nilai recallnya sebesar 100% sehingga bisa dikategorikan sebagai Good Classification [10].

Dinas Kependudukan dan Pencatatan Sipil merupakan suatu instansi pemerintahan yang bergerak dalam bidang pelayanan publik yang memberikan pelayanan terkait administrasi kependudukan [11]. Pelayanan Publik merupakan suatu tolak ukur kinerja pemerintah yang paling nyata [12]. Suatu pelayanan dikatakan memuaskan apabila harapan masyarakat terhadap suatu produk sesuai dengan kenyataan yang diterimanya [13]. Saat ini masih ada saja permasalahan yang terjadi pada proses layanan sehingga mencerminkan rasa tidak puas masyarakat terhadap kualitas pelayanan publik. Kepuasan pelayanan terhadap masyarakat yang ditetapkan dapat diukur melalui Indeks Kepuasan Masyarakat (IKM) [14]. IKM adalah data informasi tentang tingkat kepuasan masyarakat dari hasil pengukuran secara kualitatif dan kuantitatif tentang pendapat masyarakat dari pelayanan yang di dapat selama pelayanan publik dengan membandingkan antara harapan dan kebutuhan yang di dapat [15].

2. Metodologi Penelitian

Kerangka penelitian merupakan suatu bentuk kerangka kerja yang bisa digunakan sebagai pendekatan dalam memecahkan masalah. Adapun kerangka kerja yang penulis lakukan dapat dilihat pada Gambar 1.



Gambar 1. Kerangka Kerja Penelitian

Berdasarkan kerangka kerja penelitian pada Gambar 1, maka masing-masing langkah dapat diuraikan sebagai berikut:

a. Mengidentifikasi Masalah

Tahap ini adalah langkah awal dari penelitian yang akan dilakukan di mana peneliti akan membuat rumusan masalah dan menentukan batasan dari permasalahan yang diteliti agar lebih terarah.

b. Menganalisa Masalah

Pada tahap ini dilakukan untuk memahami permasalahan dengan ruang lingkup dan batasan yang telah ditentukan. Dengan melakukan analisa permasalahan yang sudah ditentukan, maka masalah tersebut dapat dengan mudah dipahami.

c. Melakukan Studi Literatur

Tahap selanjutnya dalam penelitian ini adalah melakukan studi literatur. Di mana studi literatur diperlukan sebagai bahan untuk menguatkan kajian penelitian, agar penelitian dapat dilakukan dengan baik dan mencapai tujuan yang diharapkan. Literatur diambil dari berbagai sumber seperti artikel, jurnal dan referensi lainnya untuk melengkapi penelitian ini.

d. Mengumpulkan Data

Dalam mengumpulkan data dan informasi, peneliti menggunakan data yang bersumber dari database kepuasan masyarakat dalam melakukan pelayanan dokumen kependudukan.

e. Mengolah Data

Setelah memperoleh data, selanjutnya dilakukan pengolahan data dengan menyeleksi data, membuang data yang tidak sesuai dan memperbaiki kesalahan data.

f. Menganalisa Data Menggunakan Metode Naïve Bayes

Pada tahap ini, data yang selesai diolah akan dianalisa dan diklasifikasikan menggunakan Microsoft Excel dengan metode Naïve Bayes berdasarkan ketentuan yang ada pada metode tersebut.

g. Mengimplementasikan Sistem

Pada tahap ini, penulis akan melakukan suatu perhitungan untuk memprediksi indek kepuasan masyarakat dalam pelayanan dokumen kependudukan dengan menggunakan software Microsoft Excel dan software Rapid Miner untuk pengolahan data.

3. Hasil dan Pembahasan

Data yang peneliti gunakan dalam penelitian ini adalah data yang diperoleh dari data Survey Kepuasan Masyarakat (SKM) Dinas Kependudukan dan Pencatatan Sipil Kota Pariaman Tahun 2022, dimana terdapat 9 unsur pelayanan yang digunakan dalam penilaiannya berdasarkan Kemenpan RB No. 16 Tahun

2014. Adapun unsur penilaian tersebut data dilihat pada Tabel 1.

Tabel 1. Unsur Pelayanan

No	Kode Pelayanan	Unsur Pelayanan
1	V1	Kesesuaian persyaratan administrasi
2	V2	Kemudahan prosedur pelayanan
3	V3	Ketepatan penyelesaian layanan dgn janji
4	V4	Biaya
5	V5	Kesesuaian hasil dgn data
6	V6	Kemampuan petugas memberi pelayanan
7	V7	Kesopanan & keramahan petugas
8	V8	Sarana dan Prasarana
9	V9	Mekanisme sarana pengaduan masyarakat

3.1. Data Traning

Data Training adalah data yang akan dipergunakan untuk melakukan perhitungan menggunakan metode Naïve Bayes. Pada penelitian ini data training yang gunakan sebanyak 50. Data training dapat dilihat pada Tabel 2 yang direkap dalam 10 data.

Tabel 2. Data Training

No	V1	V2	V3	V4	V5	V6	V7	V8	V9	Klasifikasi
1	Pasti Sesuai	Sangat Mudah	Tepat	Gratis	Kadang Sesuai	Mampu	Ramah	Sangat Baik	Dikelola dengan Baik	Puas
2	Kurang Sesuai	Sangat Mudah	Tepat	Gratis	Sangat Sesuai	Mampu	Ramah	Sangat Baik	Dikelola dengan Baik	Puas
3	Pasti Sesuai	Sangat Mudah	Kurang Tepat	Gratis	Kadang Sesuai	Mampu	Ramah	Sangat Baik	Dikelola dengan Baik	Puas
4	Pasti Sesuai	Sangat Mudah	Kurang Tepat	Gratis	Sangat Sesuai	Mampu	Ramah	Sangat Baik	Dikelola dengan Baik	Puas
5	Kurang Sesuai	Sangat Mudah	Kurang Tepat	Ada, Sukarela	Kadang Sesuai	Mampu	Ramah	Sangat Baik	Berfungsi Kurang Maksimal	Tidak Puas
6	Sesuai	Sangat Mudah	Tepat	Gratis	Sangat Sesuai	Cukup	Ramah	Sangat Baik	Dikelola dengan Baik	Puas
7	Kurang Sesuai	Sangat Mudah	Kurang Tepat	Gratis	Sangat Sesuai	Mampu	Ramah	Sangat Baik	Dikelola dengan Baik	Puas
8	Sesuai	Sangat Mudah	Kurang Tepat	Ada, Sukarela	Sangat Sesuai	Cukup	Ramah	Baik	Befungsi kurang maksimal	Puas
9	Kurang Sesuai	Berbelit belit	Tepat	Gratis	Sangat Sesuai	Cukup	Cukup	Baik	Dikelola dengan Baik	Tidak Puas
10	Sesuai	Berbelit Belit	Kurang Tepat	Ada, Sukarela	Kadang Sesuai	Mampu	Cukup	Baik	Dikelola dengan Baik	Puas

3.2. Perhitungan Menggunakan metode Naïve Bayes

Langkah pertama yang mesti dilakukan menghitung nilai Probabilitas prior $P(C_i)$ untuk setiap kelas. Kelas C0 (Puas) adalah sebanyak 8 data, sedangkan kelas C1 (Tidak Puas) sebanyak 2 data. Perhitungan probabilitas dapat dilihat pada Tabel 3.

Tabel 3. Probabilitas

	Probabilita	Hasil
Puas (C0)	8/10	0,8
Tidak Puas (C1)	2/10	0,2

Selanjutnya melakukan Probabilitas V bersyarat ($P(X|C_i)$), pada data latih sebanyak 10 data dengan menggunakan V sebagai vektor kepuasan seperti yang terlihat pada Tabel 1. Unsur Pelayanan. Sehingga $P(C_i)$ dapat di lihat pada Tabel 4.

Tabel 4. Klasifikasi Kesesuaian Pesyaratan

V1	Jumlah Keadian yang dipilih		Probabilitas	
Kesesuaian Persyaratan	Puas	Tidak Puas	Puas (C0)	Tidak Puas (C1)
Pasti Sesuai	4	1	0,5714285714	0,3333333333
Kurang sesuai	3	2	0,4285714286	0,6666666667
Jumlah	7	3	1	1

Tabel 5. Klasifikasi Prosedure Pelayanan

V2	Jumlah Kejadian yang dipilih		Probabilitas	
Prosedure	Puas	Tidak Puas	Puas (C0)	Tidak Puas (C1)
Sangat Mudah	7	1	0,875	0,5
Berbelit Belit	1	1	0,125	0,5
Jumlah	8	2	1	1

Tabel 6. Klasifikasi Waktu Pelayanan

V3	Jumlah Kejadian yang dipilih		Probabilitas	
Waktu Pelayanan	Puas	Tidak Puas	Puas (C0)	Tidak Puas (C1)
Tepat	3	1	0,375	0,5
Kurang Tepat	5	1	0,625	0,5
Jumlah	8	2	1	1

Tabel 7. Klasifikasi Biaya

V4	Jumlah Kejadian yang dipilih		Probabilitas	
Biaya	Puas	Tidak Puas	Puas (C0)	Tidak Puas (C1)
Gratis	6	1	0,75	0,5
Ada, Sukarela	2	1	0,25	0,5
Jumlah	8	2	1	1

Tabel 8. Klasifikasi Kesamaan Hasil dan Data

V5	Jumlah Kejadian yang dipilih		Probabilitas	
Kesamaan Hasil dan Data	Puas	Tidak Puas	Puas (C0)	Tidak Puas (C1)
Sangat Sesuai	5	1	0,625	0,5
Kadang Sesuai	3	1	0,375	0,5
Jumlah	8	2	1	1

Tabel 9. Klasifikasi Kemampuan Petugas memberi Layanan

V6	Jumlah Kejadian yang dipilih		Probabilitas	
Kemampuan Petugas	Puas	Tidak Puas	Puas (C0)	Tidak Puas (C1)
Mampu	6	1	0,75	0,5
Cukup	2	1	0,25	0,5
Jumlah	8	2	1	1

Tabel 10. Klasifikasi Kesopanan dan Keramahan Petugas

V7	Jumlah Kejadian yang dipilih		Probabilitas	
Kesopanan dan Keramahan	Puas	Tidak Puas	Puas (C0)	Tidak Puas (C1)
Ramah	7	1	0,75	0,5
Cukup	1	1	0,25	0,5
Jumlah	8	2	1,00	1,00

Tabel 11. Klasifikasi Sarana dan Prasarana

V8	Jumlah Kejadian yang dipilih		Probabilitas	
Sarana dan Prasarana	Puas	Tidak Puas	Puas (C0)	Tidak Puas (C1)
Sangat Baik	6	1	0,8571428571	0,3333333333
Baik	1	2	0,1428571427	0,6666666667
Jumlah	7	3	1	1

Tabel 12. Klasifikasi Mekanisme Sarana Pengaduan

V9	Jumlah Kejadian yang dipilih		Probabilitas	
Mekanisme Sarana Pengaduan	Puas	Tidak Puas	Puas (C0)	Tidak Puas (C1)
Berfungsi Kurang Maksimal	1	1	0,125	0,5
Dikelola dengan Baik	7	1	0,875	0,5
Jumlah	8	2	1	1

Setelah ditemukan probabilitas pada masing-masing kriteria selanjutnya menghitung jumlah probabilitas.

$$P(\text{Puas}[C0])$$

$$=0,5714285714*0,875*0,375*0,75*0,375*0,75*0,75*0,857142857*0,875$$

$$=0,022247318$$

$$P(\text{Tidak Puas}[C1])$$

$$=0,4285714286*0,875*0,625*0,25*0,375*0,75*0,75*0,8571428571*0,125$$

$$=0,001324245$$

$$P(X|\text{Puas}[C0]) \times P(\text{Puas}[C0])$$

$$=0,8*0,022247318$$

$$=0,017797854$$

$$P(X|\text{Tidak Puas}[C1]) \times P(\text{Tidak Puas}[C1])$$

$$=0,2*0,001324245$$

$$=0,000264849$$

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan dengan menggunakan metode Naïve Baye Classifier, nilai probabilitas prediksi indeks kepuasan masyarakat diperoleh dengan hasil 0,000264849 yang diklasifikasikan Puas. Dengan demikian dapat disimpulkan bahwa Pelayanan pada Dinas Kependudukan dan Pencatatan Sipil Kota Pariaman sudah dilakukan dengan baik.

Daftar Rujukan.

- [1] Wu, W. T., Li, Y. J., Feng, A. Z., Li, L., Huang, T., Xu, A. D., & Lyu, J. (2021). Data mining in clinical big data: the frequently used databases, steps, and methodological models. *Military Medical Research*, 8(1), 1-12. <https://doi.org/10.1186/s40779-021-00338-z>
- [2] Ghazal, M. M., & Hammad, A. (2022). Application of knowledge discovery in database (KDD) techniques in cost overrun of construction projects. *International Journal of Construction Management*, 22(9), 1632-1646. <https://doi.org/10.1080/15623599.2020.1738205>

- [3] Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37-37. DOI: <https://doi.org/10.1609/aimag.v17i3.1230>
- [4] Gan, W., Lin, J. C. W., Chao, H. C., & Zhan, J. (2017). Data mining in distributed environment: a survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(6), e1216. <https://doi.org/10.1002/widm.1216>
- [5] Na'am, J. (2017). Edge Detection on Objects of Medical Image with Enhancement multiple Morphological Gradient (EmMG) Method. *4th Proc. EECSI*. 23-24 Sep. 2017. Yogyakarta: Indonesia. <http://dx.doi.org/10.1109/EECSI.2017.8239085>
- [6] Farhana, S. (2021). Classification of Academic Performance for University Research Evaluation by Implementing Modified Naive Bayes Algorithm. *Procedia Computer Science*, <https://doi.org/10.1016/j.procs.2021.10.077>
- [7] Saritas, M. M., & Yasar, A. (2019). Performance analysis of ANN and Naive Bayes classification algorithm for data classification. *International Journal of Intelligent Systems and Applications in Engineering*, 7(2), 88-91. DOI: <https://doi.org/10.18201/ijisae.2019252786>
- [8] Safitri, A. R., & Muslim, M. A. (2020). Improved accuracy of naive bayes classifier for determination of customer churn uses smote and genetic algorithms. *Journal of Soft Computing Exploration*, 1(1), 70-75. DOI: <https://doi.org/10.52465/josce.v1i1.5>
- [9] Yulias, N., & Widiyanto, S. R. (2021). Prediction of drinking water facility conditions using the Naive Bayes Algorithm. *Jurnal Mantik*, 4(4), 2599-2603. DOI: <https://doi.org/10.35335/mantik.Vol4.2021.1190.pp2599-2603>
- [10] Amillina, I., & Qoiriah, A. (2021). Penerapan Algoritma Naïve Bayes dalam Klasifikasi Tingkat Kepuasan Siswa terhadap Pembelajaran Daring. *Jurnal Ilmiah Teknologi Informasi dan Robotika*, 3(2), 16-23. DOI: <https://doi.org/10.33005/jifti.v3i2.59>
- [11] Rudiansyah, Y. A., Argenti, G., & Febriantini, K. (2022). Kualitas pelayanan administrasi kependudukan pada masa pandemi covid 19 di dinas kependudukan dan pencatatan sipil. *Kinerja*, 18(4), 513-520. DOI: <http://dx.doi.org/10.29264/jkin.v18i4.9837>
- [12] Oktariyanda, T. A. INOVASI APLIKASI SAMSAT RAME (SAMSAT RAMAH MERAKYAT) DALAM OPTIMALISASI PELAYANAN DI KANTOR BERSAMA SAMSAT GRESIK. <https://doi.org/10.26740/publika.v9n2.p185-200>
- [13] Marlius, D., & Putriani, I. (2020). Kepuasan Nasabah PT. Bank Rakyat Indonesia Unit Tapan Cabang Painan Dilihat dari Kualitas Layanan Customer Service. *Jurnal Pundi*, 3(2), 111-122. DOI: <https://doi.org/10.31575/jp.v3i2.151>
- [14] Damayanti, L. D., Suwena, K. R., & Haris, I. A. (2019). Analisis kepuasan masyarakat terhadap pelayanan publik berdasarkan indeks kepuasan masyarakat (IKM) Kantor Kecamatan Sawan Kabupaten Buleleng. *Jurnal Pendidikan Ekonomi Undiksha*, 11(1), 21-32. DOI: <https://doi.org/10.23887/jjpe.v11i1.20048>