

Algoritma Genetika untuk Optimalisasi Klasifikasi Kepuasan Pelayanan e-KTP

Castaka Agus Sugianto^{1st}

Program Studi Teknik Informatika
Politeknik TEDC Bandung
Cimahi, Indonesia
castaka@poltektedc.ac.id

Tri Herdiawan Apandi^{2nd}

Manajemen Informatika
Politeknik Negeri Subang
Subang, Indonesia
h.apandi@gmail.com

Abstract— Pelayanan publik adalah segala kegiatan pelayanan yang dilaksanakan oleh penyelenggara pelayanan publik sebagai upaya pemenuhan kebutuhan penerima pelayanan, maupun dalam rangka pelaksanaan ketentuan peraturan perundang – undangan. Salah satu pelayanan public yang paling sering dikeluhkan adalah E-KTP menurut data sisa perekaman masih banyak yang belum melakukan perekaman data tersebut. Untuk melihat kepuasan masyarakat terhadap pelayanan public perlu dibuatnya klasifikasi keluhan masyarakat. Keluhan dari masyarakat dikumpulkan dalam bentuk pertanyaan yang telah disebar sebelumnya. Data yang dikumpulkan akan diklasifikasi menggunakan metode Naïve Bayes ditambah dengan optimasi menggunakan Algoritma Genetik. Hasil klasifikasi akan didapat nilai akurasi. Akurasi akan dibandingkan pengaruh optimasinya. Pengujian akurasi menunjukkan Naïve Bayes ditambah optimasi menggunakan Algoritma Genetik dengan pemilihan Roulette Wheel memiliki nilai akurasi yang lebih baik dengan nilai akurasi 85,64% dibandingkan dengan pemilihan skema *default* dengan nilai akurasi 83,03.

Kata Kunci –Naïve Bayes, Optimasi, Algoritma Genetik.

I. PENDAHULUAN

Pelayanan publik adalah segala kegiatan pelayanan yang dilaksanakan oleh penyelenggara pelayanan publik sebagai upaya pemenuhan kebutuhan penerima pelayanan, maupun dalam rangka pelaksanaan ketentuan peraturan perundang - undangan [1]. Penyelenggaraan Pelayanan Publik ini tidak akan bisa dicapai secara maksimal apabila aparatur pemerintah tidak bekerja secara optimal, harus diimbangi dengan upaya optimalisasi kinerja aparatur pemerintah melakukannya secara konsisten dengan memperhatikan segala kebutuhan dan harapan masyarakat.

Berdasarkan surat edaran Menteri Dalam Negeri (Mendagri) nomor 471/1768/SJ bahwa sisa perekaman di Kabupaten Bandung Barat berjumlah 7.974 jiwa harus diselesaikan selambat-lambatnya pada 30 September 2016. Sementara, sisa pencetakan E-KTP berjumlah 27.507 jiwa yang harus selesai pada 31 Desember 2016 dari wajib E-KTP berjumlah 1.014.742 jiwa. [2].

Salah satu dari kecataman yang mengalami masalah perekaman data di kabupaten bandung barat adalah kecamatan Batujajar. Kecamatan tersebut memiliki beberapa masalah yang dikeluhkan masyarakat menurut hasil observasi sementara diantaranya adalah waktu pelayanan yang *relative* lama dan tidak ada batasan waktu yang pasti, masih banyak warga yang belum paham terhadap prosedur pelayanan yang diberikan, kurangnya sosialisasi tentang pembuatan E-KTP sehingga warga tidak mengurus kartu identitas tersebut, keterlambatan petugas dalam pembuatan E-KTP, kurangnya fasilitas dan kelengkapan pembuatan E-KTP seperti komputer, *finger print* yang kurang memadai sehingga menimbulkan antrian jika yang mengajukan permohonan pembuatan E -KTP banyak.

Dengan adanya masalah yang telah dijabarkan diatas maka daripada itu diperlukan proses melihat tingkat kepuasan pelayanan E-KTP lebih awal. Untuk itu diperlukan cara mengelolah tingkat kepuasan pelayanan dengan cepat, salah satunya dengan metode prediksi pada data mining. Banyak penelitian yang

telah dilakukan sebelumnya dan algoritma naive bayes ialah merupakan salah satu algoritma yang terbaik dibandingkan algoritma lainnya seperti: *logistic regression*, *neural network*, *random forest*, *decision tree*, *support vector machine* dan *k-nearest neighbor*. Naive bayes merupakan algoritma klasifikasi yang cukup sederhana selain daripada itu mudah dalam penerapannya[3] sehingga algoritma ini sangat efektif apabila diuji menggunakan dataset yang tepat, terutama apabila naive bayes digabungkan dengan seleksi fitur, maka naive bayes dapat mengurangi *redundant* pada data [4], selain dapat di gabungkan dengan teknik *clustering* seperti *k-means*[5]. Naive bayes terbukti memiliki akurasi yang tinggi dibandingkan dengan SVM [6] dan kecepatan yang cepat saat diaplikasikan ke dalam database dengan data yang besar [7].

Dalam mengukur tingkat kepuasan masyarakat terhadap pelayanan memerlukan beberapa indikator, adapun indikator yang digunakan dalam penelitian ini adalah indikator yang berasal dari Direktorat Jenderal Sumber Daya dan Perangkat Pos dan Informatika (Ditjen SDPPI) yang berjumlah 17 item dan dijadikan 44 sub pertanyaan, untuk mengetahui indikator yang paling berpengaruh dalam menentukan tingkat kepuasan masyarakat terhadap pelayanan E-KTP di Kec. Batujajar diperlukan metode pemilihan fitur. Pemilihan atribut/fitur merupakan tugas penting dalam berbagai domain data mining yang merupakan salah satu teknik yang paling sering dan penting dalam pra-pengolahan data, dan telah menjadi komponen yang sangat diperlukan dalam proses pembelajaran mesin [8]. Penulis telah mempertimbangkan berbagai teknik pemilihan fitur untuk berbagai aplikasi di berbagai domain. Pemilihan fitur adalah prosedur untuk memilih indikator yang optimal dari yang tersedia [9].

Naive bayes merupakan algoritma klasifikasi data mining yang menganggap setiap atribut tidak saling berhubungan. Maka dari itu perlu menggunakan algoritma genetika (GA) untuk mengoptimalkan algoritma naive bayes dalam memilih atribut-atribut atau variabel yang harus digunakan sehingga diharapkan dapat meningkatkan nilai akurasi. Algoritma Genetika adalah kelas paling populer dari algoritma evolusioner yang menggunakan gen dan kromosom untuk mewakili individu (yang sesuai dengan solusi)[10]. Dimana tugasnya adalah untuk menemukan seperangkat variabel yang optimal.

Algoritma genetika cocok untuk masalah berskala besar [11]. Perbandingan antara Algoritma genetika dan pemilihan model penuh (*support vector machine* dan *a particle swarm model selection*) pada masalah klasifikasi, hasilnya menunjukkan bahwa Algoritma genetika memberikan kinerja yang lebih

baik pada masalah dengan dimensi yang tinggi dan set pelatihan yang besar [11].

II. PENELITIAN TERKAIT

Terdapat beberapa riset yang telah dilakukan oleh banyak peneliti lain sebelumnya yang berkaitan dengan kepuasan pelayanan, seperti yang akan dijelaskan dibawah ini :

Rezha, Rochmah dan Siswidiyanto yang judulnya “Analisis Pengaruh Kualitas Pelayanan Publik terhadap Kepuasan Masyarakat (Study Tentang Pelayanan Perekaman Kartu Tanda Penduduk Elektronik (E-KTP) di Kota Depok)”. Kualitas Pelayanan yang terdiri dari bukti fisik, *reliabilitas*, daya tanggap, jaminan, dan empati berpengaruh signifikan terhadap kepuasan masyarakat yang kualitas pelayanan memberikan pengaruh secara berarti terhadap kepuasan masyarakat yang menerima pelayanan perekaman e-KTP. Bahwa 75,8% variabel Kepuasan masyarakat akan dipengaruhi oleh kualitas pelayanan, sedangkan sisanya 24,2% variabel Kepuasan masyarakat akan dipengaruhi oleh variabel-variabel yang lain yang tidak dibahas dalam penelitian ini [12].

Hamta dalam penelitiannya yang berjudul “Analisis Penerapan Data Mining Dalam Mengukur Tingkat Kepuasan Masyarakat Pada Pelayanan Samsat Batam”. Dari hasil analisis *data mining* dengan menunjukkan tingkat kepuasan masyarakat pada pelayanan SAMSAT Batam sebesar 94.04 %, dengan validitas hasil penelitian 79.6%. Dan penerapan *data mining* menunjukkan bahwa variabel Tanggung Jawab dan Keamanan memiliki persentase tertinggi dalam memberikan kepuasan masyarakat dengan tingkat nilai 96%, dan Kedisiplinan diposisi terendah dari variabel lainnya yakni 92% [13].

Adapun penelitian yang dilakukan Maris dengan judul “Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5”. Pengimplementasian metode C4.5 menggunakan data pelanggan dapat digunakan untuk menentukan kepuasan pelanggan. *Ratio data training* yang digunakan mempengaruhi nilai akurasi pada setiap percobaan. Pada percobaan ke- 1 nilai akurasi adalah 86% dengan *data training* 70% dan *data testing* 30%. Pada percobaan ke-2 nilai akurasi adalah 93% dengan *data training* 90% dan *data testing* 10%. Dengan hasil ini pengujian yang terbaik yaitu dengan akurasi 93% [14].

Buani dalam penelitiannya “Optimasi Algoritma Naive Bayes dengan Menggunakan Algoritma Genetika untuk Prediksi Kesuburan (*Fertility*)” Penelitian ini membuktikan bahwa gaya hidup dapat mempengaruhi tingkat kesuburan (*Fertility*). Menambahkan Algoritma Genetika untuk meningkatkan hasil prediksi yang dilakukan oleh Naive Bayes terbukti dapat meningkatkan tingkat

akurasi prediksi kesuburan(Fertility) dengan akurasi sebesar 99,33% [15].

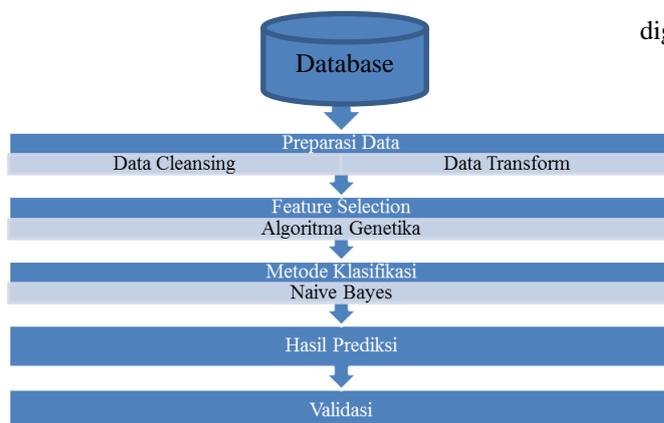
Wahyuni dkk dengan judul penelitiannya “Prediksi hasil pemilu legislatif DKI Jakarta menggunakan naïve bayes dengan algoritma genetika sebagai fitur seleksi”. Sedangkan prediksi menggunakan naïve bayes dan AG sebagai fitur seleksi memiliki akurasi 97,84% dan nilai AUC 0,994. jadi penggunaan AG sebagai fitur seleksi dapat meningkatkan akurasi prediksi hasil pemilu legislatif DKI Jakarta [7].

Oman Somantri dan M. Khambali (2017) dalam penelitiannya yang berjudul “Feature Selection Klasifikasi Kategori Cerita Pendek Menggunakan Naïve Bayes dan Algoritme Genetika” hasilnya menunjukkan bahwa algoritma Naive Bayes ketikadilakukan *feature selection* menggunakan Algoritma Genetika mengalami kenaikan akurasi yaitu dari 78,59% menjadi 84,29% [16].

III. METODE PENELITIAN

A. Model yang diusulkan

Model yang diusulkan disajikan pada Gbr 1, permodelan yang dimulai dari pengumpulan dataset sampai validasi dataset. pada tahapan ini *dataset* yang di dapat dilakukan preparasi data, preparasi data dilakukan 2 kegiatan yaitu data *cleansing* dan data *transform*, berikutnya dilakukan seleksi *feature* sebelum dilakukan klasifikasi, setelah proses klasifikasi hasil prediksi dilakukan validasi.



Gbr 1. Model yang diusulkan

a. Dataset :

Dataset yang dikumpulkan sebanyak 3977 dari hasil penyebaran kuesioner di kecamatan batujajar kabupaten Bandung barat. Kuesioner yang disebar ke responden masyarakat di Kecamatan Batujajar dengan indikator yang dipakai oleh Direktorat Jenderal Sumber Daya dan Perangkat Pos dan Informatika (Ditjen

SDPPI).

b. Preparasi Data :

Pada tahap *data Cleansing* ini *dataset* akan di reparasi yaitu menghilangkan *field* yang tidak perlu dipakai ketika proses *data mining* di *rapidminer*. Dataset yang digunakan terdapat 21 *field* bisa dilihat pada tabel 4.1. pada proses ini *field* yang di hilangkan adalah *field* nama dan jumlah. Sehingga *field* yang digunakan pada *dataset* yang digunakan untuk proses datamining menjadi 19 *field*.

Transformasi Data menyiapkan data kedalam data yang siap digunakan uji coba.

c. Feature Selection :

Algoritma yang di gunakan untuk seleksi atribut adalah algoritma genetika.

d. Metoda klasifikasi :

Metode yang digunakan adalah *Naïve Bayes* untuk klasifikasi.

e. Hasil prediksi : output yang dihasilkan oleh Naïve Bayes

f. Validasi : Pengujian dilakukan menggunakan k-flip lintas validasi teknik. Cross-validasi metode digunakan untuk memprediksi keakuratan data pengujian.

IV. HASIL DAN PEMBAHASAN

Berikut adalah hasil dan pembahasan yang di dapat dari penelitian yang dilakukan :

A. Data yang Digunakan

Pada Tabel I merupakan rincian *data set* yang digunakan dalam proses *data mining*

TABLE I. DATA YANG DIGUNAKAN

No	Atribut	Jenis Data	Type Data
1	NO	Integer	id
2	A1	Integer	atribut
3	A2	Integer	atribut
4	A3	Integer	atribut
5	A4	Integer	atribut
6	A5	Integer	atribut
7	A6	Integer	atribut
8	A7	Integer	atribut
9	A8	Integer	atribut
10	A9	Integer	atribut
11	A10	Integer	atribut

12	A11	Integer	atribut
13	A12	Integer	atribut
14	A13	Integer	atribut
15	A14	Integer	atribut
16	A15	Integer	atribut
17	A16	Integer	atribut
18	A17	Integer	atribut
19	KETERANGAN	bynominal	label

B. Hasil Pengujian Naïve Bayes +GA

Pada Tabel II adalah hasil pengujian yang dilakukan oleh penulis pada algoritma *Naïve Bayes* + *Algoritma Genetika* tanpa dilakukanya perubahan pada algoritma yang digunakan.

TABLE II. PENGUJIAN ALGORITMA NAÏVE BAYES + ALGORITMA GENETIKA

	<i>True Puas</i>	<i>True Cukup Puas</i>	<i>Class Precision</i>
Pred. Puas	2365	438	84.37%
Pred. Cukup Puas	237	937	79.81%
Class Recall	90.89%	68.15%	

Berikut merupakan hasil pengujian menggunakan *algoritma naïve bayes* dengan validasi *cross validation* mendapatkan hasil seperti yang terlihat pada tabel II mendapat Akurasi 83,03%, *Precision* 79,58%, *Recall* 68,01%, *AUC* 0,903 dan *execution time* 14 second

C. Hasil Pengujian Naïve Bayes(Kernel) +GA

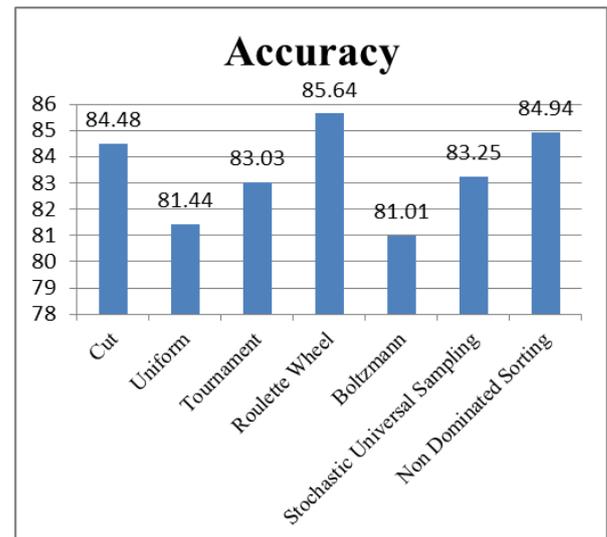
Pada Tabel III adalah hasil pengujian yang dilakukan oleh penulis pada algoritma *Naïve Bayes (Kernel)* + *Algoritma Genetika* tanpa dilakukanya perubahan pada algoritma yang digunakan.

TABLE III. PENGUJIAN ALGORITMA NAÏVE BAYES (KERNEL) + ALGORITMA GENETIKA

	<i>True Puas</i>	<i>True Cukup Puas</i>	<i>Class Precision</i>
Pred. Puas	2117	700	75.15%
Pred.	485	675	58.19%

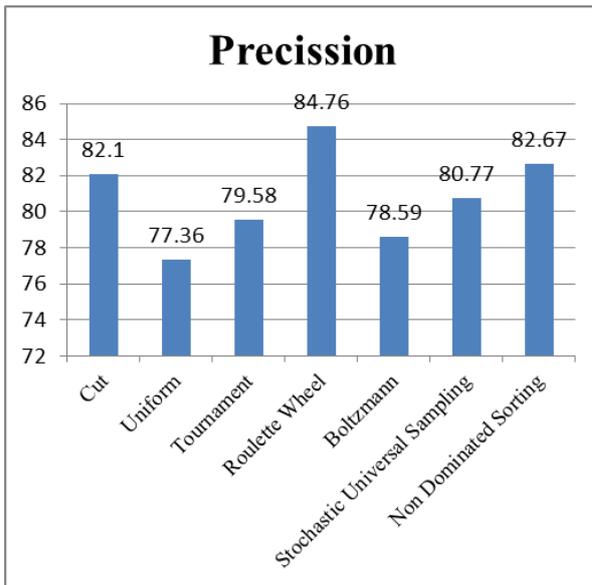
Cukup Puas			
Class Recall	81.36%	49.09%	

Berikut merupakan hasil pengujian menggunakan *algoritma naïve bayes* dengan validasi *cross validation* mendapatkan hasil seperti yang terlihat pada tabel III mendapat Akurasi 70,20%, *Precision* 58,21%, *Recall* 48,60%, *AUC* 0,715 dan *execution time* 26 second



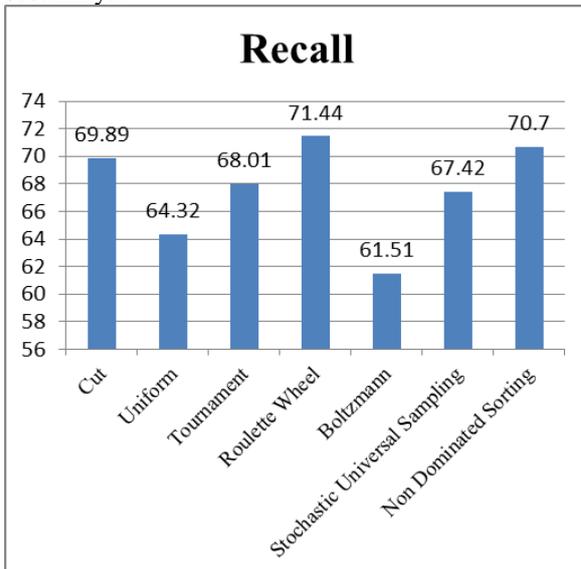
Gbr 2. Perbandingan Nilai Accuracy

Pada gambar 2 - 5 dapat dilihat optimasi menggunakan Roulette Whell menghasilkan akurasi yang terbaik, sedang untuk akurasi yang paling tinggi adalah menggunakan Roulette Wheel karena proses dari optimasi ini akan menjalankan algoritma genetika, setiap individu akan disegmentasi dalam garis yang berdekatan, proses ini akan membuat ukuran segmen menjadi sama. Sehingga proses ini akan terus diulang sampai mendapatkan nilai optimal global. Berbeda dengan Stochastic Universal Sampling akan memperhitungkan jumlah individu yang akan dipilih, ini berarti akan berpengaruh jumlah iterasinya. Sedangkan pada Boltzmann menjadi nilai akurasi yang paling rendah. Seperti yang diungkapkan pada nilai precision dan AUC berbanding lurus dengan nilai Akurasi. Dapat dilihat pada gambar 3. Precision , gambar 4. Recall dan gambar 5. AUC



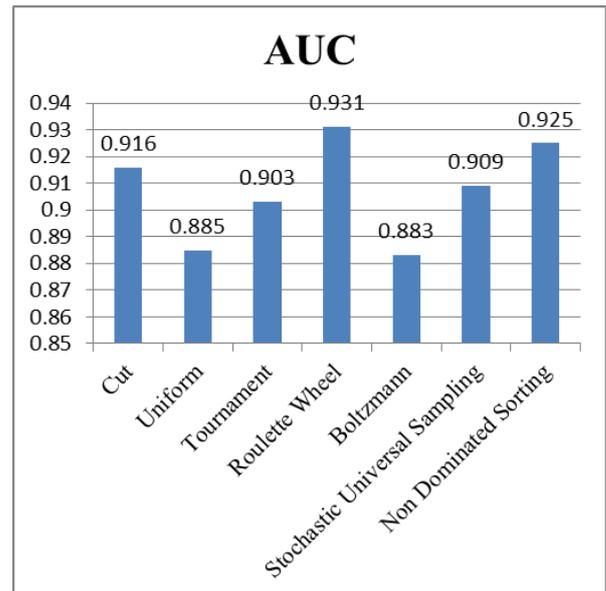
Gbr 3. Perbandingan Nilai Precision

Pada gambar 3. Perbandingan Nilai Precision nilai tertinggi didapat pada optimasi menggunakan Roulette Wheel sedangkan Uniform menunjukkan sebaliknya.



Gbr 4. Perbandingan Nilai Recall

Pada gambar 4 menunjukkan ke konsistenan pada skema Roulette Wheel dengan menunjukkan nilai recall 71.44% dan yang terendah adalah Boltzmann.



Gbr 5. Perbandingan Nilai AUC

Pada gambar 5 menunjukkan ke konsistenan pada skema Roulette Wheel dengan menunjukkan nilai AUC 0.931 dan yang terendah adalah Boltzmann.

D. Pengujian T-Test

Pada Tabel IV merupakan hasil *T-Test* yang dilakukan pada Penelitian ini, Adapun hasil pengujian T-test yang didapat dari pengujian Naïve Bayes + GA dan Naïve Bayes (Kernel) + GA didapatkan hasil sebesar 0,000 hal ini menunjukkan bahwa signifikanya perbedaan antara hasil yang didapat karena nilai yang didapat dibawah alpha T-test yaitu 0,050.

TABLE IV. HASIL PENGUJIAN T-TEST

	<i>Naïve Bayes+GA</i>	<i>Naïve Bayes(Kernel)+GA</i>
<i>Naïve Bayes+GA</i>		0,000
<i>Naïve Bayes(Kernel)+GA</i>		

V. KESIMPULAN

Berdasarkan keseluruhan penelitian yang dilakukan oleh penulis maka bisa ditarik beberapa kesimpulan yaitu sebagai berikut :

- 1) Berdasarkan hasil uji yang dilakukan dengan menggunakan algoritma naïve bayes + optimasi menggunakan Algoritma Genetika dibandingkan dengan Naïve Bayes (Kernel) + Optimasi menggunakan algoritma genetika. Hasilnya menunjukkan bahwa Naïve Bayes + algoritma Genetika dengan

pemilihan skema Roulette Wheel memiliki nilai akurasi yang lebih baik dibandingkan pemilihan skema *default*.

- 2) Untuk selection Shceme menunjukan Roulette Wheel memiliki nilai presisi yang terbaik sebesar 84,76%.
- 3) Berdasarkan hasil pengujian T-Test antara algoritma Naïve Bayes + GA dan Naïve Bayes (Kernel) + GA bahwa menunjukan perbedaan yang signifikan karena nilai sebesar 0,000.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Ditjen Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi, dan Pendidikan Tinggi, Sebagai pemberi hibah kompetitif nasional yang telah memberi dukungan financial terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] Menteri Pendayagunaan Aparatur Negara, "Keputusan Menteri Pendayagunaan Aparatur Negara Nomor: Kep/25/M.Pan/2/2004 Tentang Pedoman Umum Penyusunan Indeks Kepuasan Masyarakat Unit Pelayanan Instansi Pemerintah," p. 23, 2004.
- [2] Pojokjabar.com, "Disdukcapil Kabupaten Bandung Barat Kebut Pencetakan E-KTP," 2016. .
- [3] Fitriyani and R. S. Wahono, "Integrasi Bagging dan Greedy Forward Selection pada Prediksi Cacat Software dengan Menggunakan Naïve Bayes," *J. Softw. Eng.*, vol. 1, no. 2, pp. 101–108, 2015.
- [4] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining Practical Machine Learning Tools and Techniques*. Elsevier, 2011.
- [5] C. A. Sugianto, "Analisis Komparasi Algoritma Klasifikasi Untuk Menangani Data Tidak Seimbang Pada Data Kebakaran Hutan," *Techno.com*, vol. 14, no. 4, pp. 336–342, 2015.
- [6] T. H. Apandi and C. A. Sugianto, "Analisis Komparasi Machine Learning Pada Data Spam Sms," *J. TEDC*, vol. 12, no. 1, pp. 58–62, 2018.
- [7] D. T. Wahyuni, T. Sutojo, and A. Luthfiarta, "PREDIKSI HASIL PEMILU LEGISLATIF DKI JAKARTA MENGGUNAKAN NAÏVE BAYES DENGAN ALGORITMA GENETIKA SEBAGAI FITUR SELEKSI," pp. 1–14, 2014.
- [8] N. Bidi and Z. Elberrichi, "Feature Selection For Text Classification Using Genetic Algorithms," pp. 806–810, 2016.
- [9] L. Kumar and S. K. Rath, "Application of Genetic Algorithm as Feature Selection Technique in Development of Effective Fault Prediction Model," pp. 432–437, 2016.
- [10] A. Zagorecki, "Feature Selection for Naive Bayesian Network Ensemble using Evolutionary Algorithms," vol. 2, pp. 381–385, 2014.
- [11] W. Punlumjeak and N. Rachburee, "A Comparative Study of Feature Selection Techniques for Classify Student Performance," pp. 425–429, 2015.
- [12] F. Rezha, S. Rochmah, and Siswidiyanto, "Analisis Pengaruh Kualitas Pelayanan Publik terhadap Kepuasan Masyarakat (Study Tentang Pelayanan Perekaman Kartu Tanda Penduduk Elektronik (e-KTP) di Kota Depok)," p. 10, 2016.
- [13] F. Hamta, "Analisis Penerapan Data Mining Dalam Mengukur Tingkat Kepuasan Masyarakat Pada Pelayanan Samsat Batam," pp. 1–17, 2017.
- [14] R. E. Maris, "Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5," pp. 1–14, 2016.
- [15] D. C. P. Buani, "Optimasi Algoritma Naïve Bayes dengan Menggunakan Algoritma Genetika untuk Prediksi Kesuburan (Fertility)," vol. 4, no. 1, pp. 54–63, 2016.
- [16] O. Somantri and M. Khambali, "Feature Selection Klasifikasi Kategori Cerita Pendek Menggunakan Naïve Bayes dan Algoritme Genetika," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 6, no. 3, pp. 301–306, 2017.