ISSN:2302-8149
e-ISSN:2540-9719

Implementasi *Random Forest* untuk Prediksi Prioritas Penanganan Aduan di Dinas Pendidikan Minahasa

Implementation of Random Forest for Predicting Complaint Handling Priorities at the Minahasa Education Office

[1]Sinta Tumbo, [2]Medi Hermanto Tinambunan*, [3]Kristofel Santa
[1,2,3]Program Studi Teknik Informatika, Fakultas Teknik, Universitas Negeri Manado
[1,2,3]Kec. Tondano Selatan, Kab. Minahasa, Sulawesi Utara, 95618
*e-mail: meditinambunan@unima.ac.id

(***received*:** 9 January 2026, ***revised*:** 19 February 2026, ***accepted*:** 20 February 2026)

Abstrak

Pelayanan publik di sektor pendidikan menuntut kemampuan instansi pemerintah dalam menangani aduan masyarakat secara cepat dan tepat, namun Dinas Pendidikan Kabupaten Minahasa masih menghadapi permasalahan dalam penentuan prioritas penanganan aduan yang cenderung bersifat subjektif dan manual. Kondisi tersebut berpotensi menyebabkan keterlambatan penanganan aduan yang bersifat mendesak. Penelitian ini bertujuan untuk mengimplementasikan algoritma *Random Forest* dalam memprediksi prioritas penanganan aduan masyarakat secara objektif dan berbasis data. Metode penelitian meliputi pengumpulan *dataset* aduan masyarakat sebanyak 500 data dengan 6 atribut, yaitu jenis aduan, tingkat pelanggaran, unit kerja, tindak lanjut, hasil, dan prioritas, tahapan *preprocessing* data, pembagian data *training* dan *testing*, pembangunan model *Random Forest*, serta evaluasi model. Proses pengolahan dan pemodelan data dilakukan menggunakan Jupyter Notebook melalui Anaconda. Evaluasi kinerja model dilakukan dengan menggunakan metrik *accuracy, confusion matrix, precision, recall,* dan *F1-score*. Hasil penelitian menunjukkan bahwa algoritma *Random Forest* mampu memprediksi prioritas penanganan aduan dengan tingkat *accuracy* sebesar 100%, serta menghasilkan nilai *precision, recall,* dan *F1-score* sebesar 1.00 pada seluruh kelas prioritas. Hasil ini menunjukkan bahwa model memiliki performa klasifikasi yang sangat baik dan stabil, sehingga dapat dimanfaatkan sebagai dasar pengembangan sistem pendukung keputusan untuk meningkatkan efektivitas dan kualitas pelayanan publik pada Dinas Pendidikan Kabupaten Minahasa.

Kata kunci: *data mining*, aduan masyarakat, pelayanan publik, prediksi prioritas, *random forest*

Abstract

Public service delivery in the education sector requires government institutions to handle public complaints promptly and accurately. However, the Minahasa Regency Education Office still faces challenges in determining complaint handling priorities, which are often subjective and manually assigned. This condition may lead to delays in addressing urgent complaints. This study aims to implement the Random Forest algorithm to objectively predict complaint handling priorities based on data. The research methodology includes collecting a dataset of 500 public complaints with six attributes: complaint type, violation level, work unit, follow-up action, outcome, and priority. The process involves data preprocessing, splitting the dataset into training and testing sets, building the Random Forest model, and evaluating its performance. Data processing and modeling were conducted using Jupyter Notebook within the Anaconda environment. Model performance was evaluated using accuracy, confusion matrix, precision, recall, and F1-score metrics. The results show that the Random Forest algorithm achieved an accuracy of 100%, with precision, recall, and F1-score values of 1.00 across all priority classes. These findings indicate that the model demonstrates excellent and stable classification performance. Therefore, it can serve as a foundation for developing a decision support system to improve the effectiveness and quality of public service delivery at the Minahasa Regency Education Office.

Keywords: *data mining, public complaints, public services, priority prediction, random forest*

 ISSN:2302-8149
e-ISSN:2540-9719

1 Pendahuluan

Pelayanan publik adalah bentuk usaha sadar dari penyelenggara negara kepada masyarakat berupa barang dan/ atau jasa guna pemenuhan kebutuhan masyarakat, karena itu merupakan hak dari setiap warga negara karena dijamin oleh undang undang dan kepada pelayan publik wajib untuk melakukannya [1]. Pelayanan publik ini sendiri sudah diatur dalam Undang-undang yaitu dalam Undang-undang yaitu pada Nomor 25 tahun 2009 [2]. Pelayanan publik yang berkualitas menjadi tuntutan utama bagi instansi pemerintah [3], termasuk di sektor pendidikan. Salah satu indikator penting dalam menilai kualitas pelayanan adalah kemampuan instansi dalam menerima dan menangani aduan masyarakat secara cepat dan tepat [4] dan untuk meningkatkan efisiensi dan efektivitas layanan publik [5]. Dinas Pendidikan sebagai penyelenggara layanan pendidikan sering menerima berbagai aduan terkait administrasi, tenaga pendidik, sarana dan prasarana, maupun kebijakan pendidikan [6]. Banyaknya aduan yang masuk dengan karakteristik yang beragam menyebabkan proses penanganan aduan menjadi kurang optimal apabila masih dilakukan secara manual [7].

Permasalahan utama yang dihadapi adalah kesulitan dalam menentukan prioritas penanganan aduan. Selama ini, penentuan prioritas cenderung bersifat subjektif dan bergantung pada penilaian petugas, sehingga berpotensi menimbulkan keterlambatan penanganan aduan yang bersifat mendesak [8]. Kondisi ini menunjukkan perlunya pendekatan berbasis data yang mampu membantu instansi dalam mengklasifikasikan dan memprediksi tingkat prioritas aduan secara objektif dan konsisten.

Pemanfaatan teknik *data mining* dan *machine learning* menjadi salah satu solusi yang dapat diterapkan untuk mengatasi permasalahan dalam penelitian ini. Algoritma *Random Forest* merupakan metode *ensemble learning* yang memiliki keunggulan dalam menangani data dengan banyak atribut, mampu mengurangi *overfitting*, serta menghasilkan tingkat akurasi yang tinggi [9]. Algoritma *Random Forest* merupakan salah satu dari algoritma komputasi yang memiliki manfaat untuk membantu proses pengelolaan dan penanganan aduan secara terstruktur [10]. Oleh karena itu, algoritma ini dinilai sesuai untuk diterapkan dalam prediksi prioritas penanganan aduan berdasarkan data historis aduan [11].

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengimplementasikan algoritma *Random Forest* dalam memprediksi prioritas penanganan aduan [12] pada Dinas Pendidikan Kabupaten Minahasa. Penelitian ini diharapkan dapat memberikan manfaat dalam mendukung pengambilan keputusan penentuan prioritas aduan secara lebih cepat, objektif, dan berbasis data, serta menjadi referensi pengembangan sistem pendukung keputusan dalam peningkatan kualitas pelayanan publik di bidang pendidikan [13].

2 Tinjauan Literatur

Penggunaan algoritma *Random Forest* telah banyak diterapkan dalam berbagai konteks layanan publik dan menunjukkan kinerja yang sangat baik. Pada domain layanan kesehatan, *Random Forest* mampu menghasilkan nilai AUC di atas 98% serta *accuracy* lebih dari 97% dalam prediksi waktu tunggu pelayanan pasien rawat jalan, yang menunjukkan efektivitasnya dalam menangani permasalahan prediktif multikategori dengan tingkat *noise* data yang relatif tinggi [14]. Dalam konteks pendidikan, algoritma ini juga terbukti mampu memberikan prediksi yang andal terhadap prestasi akademik siswa dengan evaluasi metrik *accuracy, precision,* dan *recall* yang tinggi, sehingga secara metodologis relevan untuk diterapkan pada permasalahan prediksi berbasis data pendidikan [15]. Selain itu, *Random Forest* juga menunjukkan performa yang baik pada domain lingkungan, di mana algoritma ini mampu mencapai tingkat *accuracy* sebesar 88,33% dalam prediksi kelayakan air minum, meskipun *dataset* yang digunakan bersifat multikategori dan kompleks, selama tahapan *preprocessing* dilakukan dengan tepat [16].

Pada konteks pengelolaan dan klasifikasi aduan pelanggan, *Random Forest* dilaporkan memiliki tingkat stabilitas dan *accuracy* yang lebih konsisten dibandingkan algoritma pembanding ketika dihadapkan pada *dataset* aduan yang heterogen dan tidak seimbang [17]. Hal ini mengindikasikan bahwa *Random Forest* efektif dalam menangani variasi tingkat urgensi dan kompleksitas permasalahan yang umumnya terdapat pada data aduan masyarakat. Keunggulan *Random Forest* dalam mengelola data berdimensi tinggi dan menangkap hubungan *non-linear* antar atribut juga terlihat pada pengolahan data publik pemerintahan, di mana algoritma ini mampu menghasilkan

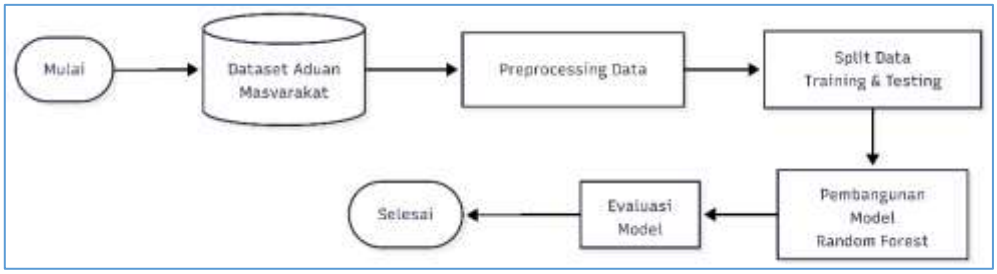
 ISSN:2302-8149
e-ISSN:2540-9719

prediksi yang lebih akurat dibandingkan metode regresi konvensional dalam memprediksi indikator sosial seperti persentase kemiskinan daerah [18].

Dibandingkan dengan algoritma lain yang sering digunakan pada kasus klasifikasi teks dan aduan, seperti *Naive Bayes* dan *Support Vector Machine (SVM)*, *Random Forest* memiliki beberapa keunggulan utama. *Naive Bayes* dikenal sederhana dan efisien, namun cenderung memiliki keterbatasan dalam menangkap hubungan kompleks antar fitur. Sementara itu, *SVM* mampu memberikan performa yang baik pada data berdimensi tinggi, tetapi memerlukan proses penentuan parameter yang lebih kompleks dan sensitif terhadap pemilihan *kernel*. Oleh karena itu, berdasarkan hasil-hasil penelitian terdahulu dan karakteristik data aduan yang bersifat tekstual, multikategori, dan heterogen, *Random Forest* dipilih dalam penelitian ini karena dinilai lebih *robust*, stabil, serta mampu memberikan performa prediksi yang konsisten pada data layanan publik.

3 Metode Penelitian

Penelitian ini terdiri atas beberapa tahapan alur yang disajikan pada Gambar 1 di bawah ini dalam bentuk digram alir, gambar tersebut menunjukkan alur penelitian yang dilakukan.



Gambar 1. Alur metode penelitian

Dari kerangka alur metode penelitian pada Gambar 1. Alur metode penelitian dapat dijelaskan sebagai berikut:

a) *Dataset* aduan masyarakat adalah *dataset* yang digunakan dalam penelitian ini, yaitu dataset pengaduan masyarakat Minahasa, dengan jumlah data sebanyak 500 data, 6 atribut, 3 kelas
b) *Preprocessing* data, bertujuan untuk menyiapkan data agar layak digunakan dalam proses *machine learning*. Pada tahap ini dilakukan pembersihan data, penghapusan data duplikat, penanganan data kosong, transformasi data kategorikal menjadi data numerik, serta normalisasi atau standarisasi data jika diperlukan. *Preprocessing* dilakukan untuk meningkatkan kualitas data dan mengurangi potensi kesalahan pada proses pemodelan.
c) *Split data training* & *testing*, data latih digunakan untuk membangun model *Random Forest*, sedangkan data uji *(testing data)* digunakan untuk mengukur kinerja model yang telah dibangun. Pembagian data dilakukan untuk memastikan bahwa model dapat diuji secara objektif terhadap data yang belum pernah dilihat sebelumnya.
d) Pembangunan *model random forest*, algoritma *Random Forest* diterapkan pada data latih untuk mempelajari pola hubungan antara atribut aduan dan kelas prioritas penanganan aduan. Model ini bekerja dengan membangun sejumlah pohon keputusan *(decision tree)* dan menggabungkan hasil prediksi dari setiap pohon untuk menghasilkan keputusan akhir yang lebih akurat dan stabil.
e) Evaluasi model, yang dilakukan dengan menggunakan data uji. Evaluasi model bertujuan untuk menilai performa model dalam memprediksi prioritas penanganan aduan. Metrik evaluasi yang digunakan meliputi *accuracy, confusion matrix, precision, recall,* dan *F1-score*. Hasil evaluasi ini digunakan untuk menentukan sejauh mana model *Random Forest* mampu memenuhi tujuan penelitian.

3.1. Deskripsi Singkat Algoritma yang Digunakan

Random Forest

Pendekatan *random forest* yang diusulkan oleh Breiman adalah algoritma pembelajaran mesin dengan banyak pohon keputusan[19]. *Random Forest* adalah model pembelajaran mesin yang digunakan dalam klasifikasi dan peramalan [20]. *Random Forest* dapat menangani data besar dan data berdimensi tinggi [21]. *Random Forest* dibangun dengan mengambil subset acak dari sampel dalam

 ISSN:2302-8149
e-ISSN:2540-9719

dataset pelatihan [22]. Teknik ini memprediksi variabel dependen dengan mengumpulkan prediksi dari pohon keputusan [23]. Untuk melatih algoritma pembelajaran mesin dan model kecerdasan buatan, sangat penting untuk memiliki jumlah data yang memadai guna pengumpulan yang efektif [20]. Data kinerja sistem sangat penting untuk menyempurnakan algoritma, meningkatkan efisiensi perangkat lunak dan perangkat keras, mengevaluasi perilaku pengguna, memfasilitasi identifikasi pola, pengambilan keputusan, pemodelan prediktif, dan pemecahan masalah, yang pada akhirnya menghasilkan peningkatan efektivitas dan *accuracy* [20]. *Random Forest* menentukan kelas akhir berdasarkan suara terbanyak *(majority voting)* dari seluruh pohon keputusan. Secara matematis, proses penentuan kelas akhir dapat dinyatakan pada Persamaan (1).

Pada persamaan (1), prediksi akhir $y^$ diperoleh dengan memilih kelas c yang memaksimalkan jumlah suara dari N pohon keputusan. Setiap pohon ke-i menghasilkan prediksi $hi(x)$ terhadap data masukan x. Fungsi indikator $I(hi(x)=c)$ bernilai 1 apabila pohon ke-i memprediksi kelas c, dan bernilai 0 apabila tidak. Penjumlahan nilai indikator untuk seluruh pohon menghasilkan total suara pada masing-masing kelas, sehingga kelas dengan akumulasi suara tertinggi ditetapkan sebagai hasil prediksi akhir. Mekanisme agregasi ini menjadikan *Random Forest* lebih *robust*, stabil, serta mampu meminimalkan risiko *overfitting* dibandingkan penggunaan satu pohon keputusan tunggal.

$$\hat{y} = \arg\max_{c} \sum_{i=1}^{N} I(h_i(x) = c) \quad (1)$$

Keterangan:
y^= hasil prediksi akhir
c= kelas (prioritas tinggi, sedang, rendah)
N= jumlah pohon keputusan
h_i (x)= hasil prediksi pohon ke-i
I= fungsi indikator (bernilai 1 jika benar, 0 jika salah)

4 Hasil dan Pembahasan

Penelitian ini menggunakan lingkungan pengembangan Jupyter Notebook yang dijalankan melalui Anaconda sebagai alat bantu dalam pengolahan data, *preprocessing*, pembangunan *model Random Forest*, serta evaluasi model.

4.1. Data Set

Data yang digunakan dalam penelitian ini merupakan data aduan masyarakat pada Dinas Pendidikan Kabupaten Minahasa. Tabel 1 menunjukkan deskripsi singkat dari *dataset*.

Tabel 1. Dataset

Dataset	Jumlah Atribut	Jumlah Data
Aduan Masyarakat	6	500

Tabel 2. deskripsi atribut di bawah ini memberikan gambaran rinci mengenai atribut yang digunakan dalam penelitian ini. Kombinasi atribut tersebut merepresentasikan karakteristik aduan masyarakat secara komprehensif dan digunakan sebagai dasar dalam proses pemodelan menggunakan algoritma *Random Forest.*

Tabel 2. Deskripsi atribut

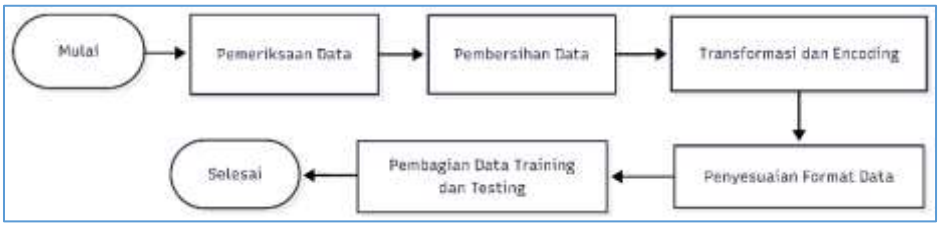
No	Nama Atribut	Definisi Operasional	Jenis Data	Peran dalam Model
1	Jenis Aduan	Kategori atau jenis permasalahan yang dilaporkan oleh masyarakat, seperti administrasi, tenaga pendidik, sarana dan prasarana, atau layanan pendidikan lainnya.	Kategorikal	Fitur
2	Tingkat Pelanggaran	Tingkat pelanggaran atau tingkat urgensi aduan berdasarkan dampak permasalahan dan kebutuhan penanganannya.	Kategorikal	Fitur

 ISSN:2302-8149
e-ISSN:2540-9719

3	Unit Kerja	Unit kerja atau instansi pendidikan yang menjadi objek aduan dan bertanggung jawab atas permasalahan yang dilaporkan.	Kategorikal	Fitur
4	Tindak Lanjut	Bentuk tindak lanjut atau respons awal yang diberikan terhadap aduan yang masuk	Kategorikal	Fitur
5	Hasil	Status atau hasil penanganan awal aduan yang digunakan sebagai informasi pendukung dalam proses klasifikasi.	Kategorikal	Fitur
6	Prioritas	Tingkat prioritas penanganan aduan yang ditetapkan oleh Dinas Pendidikan Kabupaten Minahasa dan digunakan sebagai label target dalam proses klasifikasi.	Kategorikal	Label *(Target)*

4.2 Preprocessing Data

Data *preprocessing* adalah proses mengubah data mentah kedalam bentuk yang lebih mudah dipahami. Proses ini dilakukan untuk memperbaiki kesalahan pada data mentah yang tidak lengkap dan memiliki format yang tidak teratur. Dalam tahapan *preprocessing* dalam penelitian ini adalah dilakukan pengecekan pada data yang hilang atau biasa disebut *missing values* yang terdapat dalam dataset [24]. Tujuan dari *preprocessing* ini adalah untuk meningkatkan kualitas *output* model dengan melakukan *cleaning* pada data, seperti memperbaiki atau menghapus data yang rusak atau tidak relevan [25].Tahap *preprocessing* data dilakukan untuk memastikan data aduan berada dalam kondisi siap digunakan pada proses pemodelan menggunakan algoritma *Random Forest*. Pada tahap ini, data diolah secara bertahap agar bersih, konsisten, dan sesuai dengan kebutuhan analisis.



Gambar 2 Tahapan preprocessing data

Dari Gambar 2. Tahapan *Preprocessing* Data dapat dijelaskan sebagai berikut:

a) Proses dimulai dengan pemeriksaan data untuk memastikan kelengkapan dan konsistensi *dataset* aduan masyarakat.
b) Tahap selanjutnya adalah pembersihan data, yaitu menghilangkan data duplikat serta memastikan tidak terdapat nilai kosong yang dapat memengaruhi kinerja model.
c) Setelah data bersih, dilakukan transformasi dan encoding terhadap atribut kategorikal agar dapat diproses oleh algoritma *Random Forest*.
d) Tahap penyesuaian format data dilakukan untuk memastikan seluruh atribut berada dalam bentuk numerik dan memiliki struktur yang sesuai.
e) Proses *preprocessing* diakhiri dengan pembagian data ke dalam data latih *(training data)* dan data uji *(testing data)* yang digunakan pada tahap pembangunan dan evaluasi model.

Hasil *preprocessing* disajikan dalam tabel 3. *Preprocessing* Data setelah dilakukan proses pembersihan dan *encoding.*

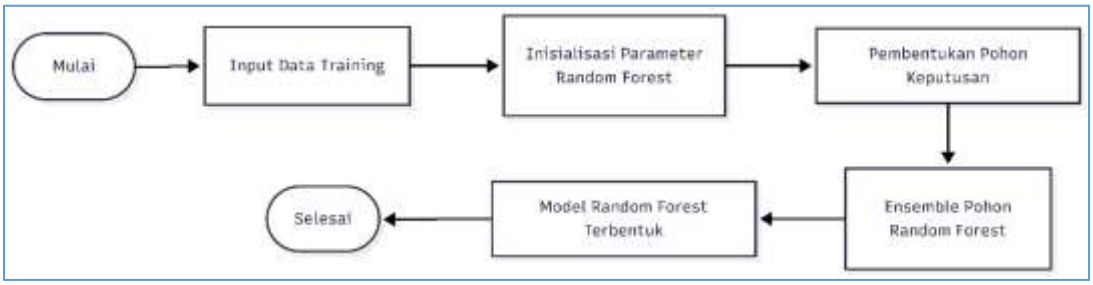
Tabel 3. Preprocessing data

Id Aduan	Jenis Aduan	Tingkat Pelanggaran	Unit Kerja	Tindak Lanjut	Hasil	Prioritas
0	3	2	1	2	2	1
1	2	0	4	1	1	2
2	0	1	0	0	1	0
3	2	0	1	1	0	2
4	1	0	5	0	1	2

 ISSN:2302-8149
e-ISSN:2540-9719

4.3. Pembangunan Model Random Forest

Pembangunan model pada penelitian ini dilakukan menggunakan algoritma *Random Forest* untuk memprediksi prioritas penanganan aduan pada Dinas Pendidikan Kabupaten Minahasa. *Random Forest* dipilih karena kemampuannya dalam menangani data kategorikal, mengurangi risiko *overfitting*, serta menghasilkan performa klasifikasi yang stabil pada data multikelas.



Gambar 3. Pembangunan model random forest

Dari kerangka pembangunan model *Random Forest* pada Gambar 3. Pembangunan Model *Random Forest* dapat dijelaskan sebagai berikut:

a) *Input Data Training*: Data hasil *preprocessing* yang digunakan untuk melatih model
b) Inisialisasi Parameter: Penentuan parameter seperti jumlah pohon *(n_estimators)*
c) Pembentukan Pohon Keputusan: Model membangun beberapa *decision tree* secara acak
d) Ensemble Pohon: Penggabungan seluruh pohon untuk membentuk *Random Forest*
e) Model Terbentuk: Model siap digunakan untuk prediksi dan evaluasi

Pada tahap inisialisasi parameter, algoritma *Random Forest* dikonfigurasi dengan parameter utama yaitu *n_estimators* sebesar 100, yang menunjukkan jumlah pohon keputusan yang dibangun dalam model. Pemilihan jumlah pohon ini bertujuan untuk meningkatkan stabilitas dan *accuracy* prediksi melalui mekanisme *ensemble*. Selain itu, parameter *random_state* ditetapkan bernilai 42 untuk memastikan hasil pelatihan model bersifat konsisten dan dapat direproduksi. Parameter lainnya menggunakan nilai bawaan *(default)* dari pustaka *scikit-learn*. Konfigurasi parameter ini dinilai cukup efektif dalam menangani data aduan yang bersifat multikelas dan heterogen tanpa meningkatkan kompleksitas model secara berlebihan. Berikut tabel *parameter model random forest* disajikan dalam Tabel 4

Tabel 4. Parameter model random forest

Parameter	Nilai	Keterangan
n_estimators	100	Jumlah pohon keputusan
Random_state	42	Menjaga reprodusibilitas hasil
Parameter lain	Default	Menggunakan pengaturan bawaan *scikit-learn*

4.4. Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja algoritma *Random Forest* dalam memprediksi prioritas penanganan aduan pada Dinas Pendidikan Kabupaten Minahasa. Evaluasi dilakukan menggunakan data uji *(testing data)* yang tidak digunakan pada proses pelatihan model. Beberapa metrik evaluasi yang digunakan meliputi *accuracy*, *confusion matrix*, serta *precision*, *recall*, dan *F1-score*. Hasil *accuracy* prediksi yang nantinya akan diukur dengan *confusion matrix* [26]. Penggunaan beberapa metrik ini bertujuan untuk memberikan gambaran kinerja model secara menyeluruh pada setiap kelas prioritas.

a. Accuracy

Hasil pengujian *accuracy* model *Random Forest* disajikan pada Tabel 5. Berdasarkan hasil pengujian tersebut, dari total 100 data uji yang digunakan, seluruh data berhasil diprediksi dengan benar oleh model, sehingga diperoleh nilai *accuracy* sebesar 100%. Nilai *accuracy* ini menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik terhadap data aduan yang digunakan dalam penelitian ini.

 ISSN:2302-8149
e-ISSN:2540-9719

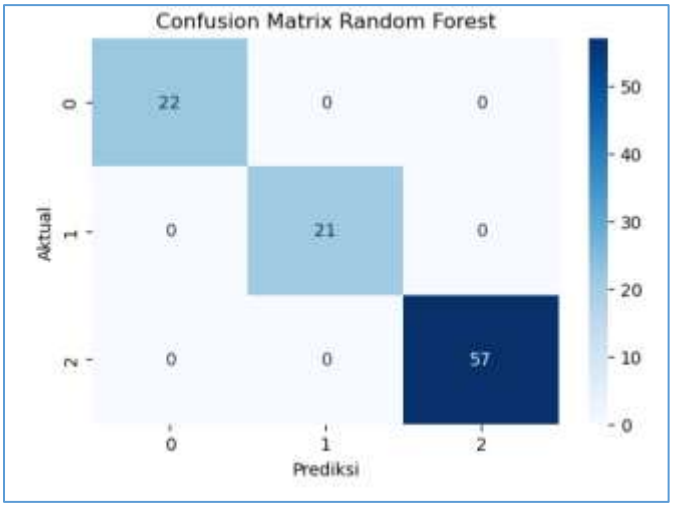
Tabel 5. Accuracy

Jumlah Data Uji	Prediksi Benar	Prediksi Salah	Accuracy
100	100	0	100%

Meskipun *accuracy* yang diperoleh sangat tinggi, evaluasi lanjutan tetap dilakukan untuk memastikan bahwa performa model tidak disebabkan oleh *overfitting* atau kebocoran data. Oleh karena itu, dilakukan validasi tambahan menggunakan metode *K-Fold Cross Validation* untuk menilai konsistensi kinerja model pada berbagai pembagian data. Pengujian ini dilakukan menggunakan 5 fold, di mana *dataset* dibagi menjadi lima bagian dan proses pelatihan serta pengujian dilakukan secara bergantian. Hasil pengujian menunjukkan bahwa model memperoleh nilai *accuracy* sebesar 100% pada seluruh fold, dengan rata-rata *accuracy* sebesar 1,0. Hasil ini mengindikasikan bahwa performa model bersifat stabil dan konsisten pada berbagai pembagian data, sehingga *accuracy* 100% yang diperoleh tidak disebabkan oleh *overfitting* atau kebocoran data.

b. Confusion Matrix

Visualisasi hasil klasifikasi model *Random Forest* ditunjukkan pada Gambar 4, yang menggambarkan distribusi prediksi model terhadap kelas prioritas rendah, sedang, dan tinggi.



Gambar 4. Confusion matrix random forest

Rincian hasil *confusion matrix* disajikan pada Tabel 6, yang menunjukkan bahwa kelas prioritas rendah memiliki 22 data yang seluruhnya diprediksi dengan benar, kelas sedang sebanyak 21 data juga terklasifikasi dengan tepat, dan kelas tinggi memiliki 57 data yang seluruhnya diprediksi sesuai dengan kelas aktualnya. Tidak adanya kesalahan klasifikasi pada ketiga kelas tersebut menegaskan bahwa model *Random Forest* mampu mengenali pola data aduan dengan sangat baik.

Tabel 6. Confusion matrix

Aktual / Prediksi	Rendah	Sedang	Tinggi
Rendah	22	0	0
Sedang	0	21	0
Tinggi	0	0	57

c. Hasil Precision, Recall, dan F1-Score

Evaluasi kinerja model selanjutnya dilakukan menggunakan metrik *precision*, *recall*, dan *F1-score* yang disajikan pada Tabel 7

Tabel 7. Hasil Precision, Recall, dan F1-Score

Kelas Prioritas	Precision	Recall	F1-Score
Rendah	1.00	1.00	1.00
Sedang	1.00	1.00	1.00
Tinggi	1.00	1.00	1.00

 ISSN:2302-8149
e-ISSN:2540-9719

Berdasarkan tabel 7. Hasil *Precision*, *Recall*, dan *F1-Score*, *Precision* menunjukkan ketepatan model dalam memprediksi suatu kelas, *recall* menunjukkan kemampuan model dalam mengenali seluruh data aktual pada kelas tersebut, sedangkan *F1-score* merupakan nilai keseimbangan antara *precision* dan *recall*. Hasil evaluasi menunjukkan bahwa seluruh kelas prioritas memperoleh nilai sempurna, yang menandakan performa model sangat stabil dan konsisten.

5 Kesimpulan

Penelitian ini berhasil mengimplementasikan algoritma *Random Forest* untuk memprediksi prioritas penanganan aduan masyarakat pada Dinas Pendidikan Kabupaten Minahasa secara objektif dan berbasis data. Proses penelitian dilakukan secara sistematis, dimulai dari tahapan *preprocessing* data yang mencakup pembersihan data, transformasi dan *encoding* atribut kategorikal, hingga pembagian data menjadi data latih *(training data)* dan data uji *(testing data)*, sehingga data berada dalam kondisi siap digunakan untuk pemodelan. Hasil evaluasi menunjukkan bahwa *model Random Forest* menghasilkan nilai *accuracy, precision, recall,* dan *F1-score* yang sangat tinggi pada seluruh kelas prioritas. Capaian performa yang optimal ini mengindikasikan bahwa model mampu mempelajari pola hubungan antar atribut aduan dan kelas prioritas secara efektif. Tingginya nilai evaluasi tersebut dapat dijelaskan oleh karakteristik dataset yang relatif terstruktur dan konsisten, di mana atribut-atribut yang digunakan memiliki relevansi yang kuat terhadap penentuan prioritas aduan. Selain itu, proses *preprocessing* dilakukan sebelum pemisahan data latih dan data uji, sehingga meminimalkan potensi terjadinya kebocoran data *(data leakage)*. Untuk memastikan bahwa performa model tidak disebabkan oleh overfitting, penelitian ini juga menerapkan metode *K-Fold Cross Validation* yang menunjukkan hasil kinerja yang konsisten pada setiap fold. Dengan demikian, *accuracy* yang diperoleh dapat dianggap valid dan representatif terhadap data yang digunakan. Meskipun demikian, penelitian lanjutan dengan jumlah data yang lebih besar, variasi atribut yang lebih beragam, serta pengujian pada data aduan dari periode atau instansi yang berbeda masih diperlukan untuk menguji kemampuan generalisasi model. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa algoritma *Random Forest* memiliki potensi yang kuat untuk diterapkan sebagai dasar pengembangan sistem pendukung keputusan dalam meningkatkan efektivitas dan kualitas pelayanan publik di sektor pendidikan.

Referensi

[1] N. K. Riani, “Strategi Peningkatan Pelayanan Publik,” Vol. 1, No. 11, pp. 2443–2452, 2021.
[2] K. Santa *et al.*, “Implementasi Metode *Extreme Programming* pada Aplikasi Pelayanan Publik Kewaspadaan Nasional dan Penanganan Konflik Kesbangpol Kota Manado,” Vol. 7, No. 1, pp. 1–12, 2025.
[3] T. Cahyati, “*Improving Public Service Quality through the Development of Online Service Innovations in Public Sector Organizations in Indonesia*,” Vol. 13, No. 1, pp. 145–154, 2023.
[4] Y. Maulana, R. Hurriyati, and F. Hidayat, “*Public Complaint Services : How Designing Integrated Services for Websites , Email and Social Media can increase Public Trust*,” Vol. 03, No. 09, pp. 2287–2295, 2023.
[5] R. R. Makalalag, G. David, and P. Maramis, “Implementasi Sistem Manajemen Dokumen *Cloud* di Dinas Kebudayaan dan Pariwisata Kabupaten Minahasa,” Vol. 9, No. 3, 2025.
[6] S. N. Lia, “*Quality of Complaint Services through the Lapor Website at Departement of Education the East Java Provincial*,” Vol. 05, No. 05, pp. 1380–1389, 2024, DOI: 10.37899/journal-la-sociale.v5i5.1275.
[7] E. M. Bertho *et al.*, “Analisis Pengelolaan Pengaduan Pelayanan Masyarakat pada Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu Kota Palangka Raya,” Vol. 9, No. November, pp. 2203–2211, 2025.
[8] M. Laliberté, L. Casgrain, and K. D. Volesky, “*Revue Canadienne de Bioéthique Handling Complaints : Considerations for Prioritizing Complaints Handling Complaints : Considerations for Prioritizing Complaints*,” 2022.
[9] M. W. Nugroho, “Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi) Analisis Performa Algoritma *Random Forest* dalam mengatasi *Overfitting* pada Model Prediksi,” Vol. 9, No. December, pp. 1562–1571, 2025.

 ISSN:2302-8149
e-ISSN:2540-9719

[10] A. A. K. Efraim Ronald Stefanus Moningkey, Sifra Ropa, "*Web-based Public Complaint System using First in First Out Algorithm in Minahasa Regency: Sistem Pengaduan Masyarakat berbasis Web dengan Algoritma First in First Out di Kabupaten Minahasa*," Vol. 26, No. 4, 2025, DOI: 10.21070/ijins.v26i4.1746.
[11] O. I. Catherine and I. J. Sunday, "*Machine Learning based Student Complaint and Priority Determination Support System*," No. 11, pp. 380–399, 2025.
[12] C. S. K. Selvi, S. J. Shri, R. Prasshanthini, and R. Sanjay, "*Customer Support Ticket Categorization and Prioritization using Natural Language Processing*," Vol. 3, No. Incoft, pp. 690–698, 2025, DOI: 10.5220/0013640400004664.
[13] H. Srivastava, M. Jha, and T. Karthick, "*Decision Support Complaint Prioritization System using a Statistical Multi-Method Algorithmic approach*".
[14] J. Homepage *et al.*, "*MALCOM: Indonesian Journal of Machine Learning and Computer Science Predicting Outpatient Service Waiting Times with Random Forest Algorithm*," Vol. 5, No. 1, pp. 35–40, 2025.
[15] R. A. Saputri and L. Rosnita, "*A Random Forest-based Predictive Model for Student Academic Performance : A Case Study in Indonesian Public High Schools*," Vol. 9, No. 3, pp. 1042–1049, 2025.
[16] K. Abdi, A. Warjaya, I. Muthmainnah, and P. H. Pahutar, "Penerapan Algoritma *Random Forest* dalam Prediksi Kelayakan Air Minum," Vol. 3, No. 2, pp. 81–88, 2023.
[17] P. Reksi, S. Susandri, and M. Syawalludin, "*Utilization of Machine Learning Algorithms for Customer Complaint Classification at Perumdam*," Vol. 1, pp. 9–17, 2025.
[18] D. N. Handayani and S. Qutub, "Penerapan *Random Forest* untuk Prediksi dan Analisis Kemiskinan," Vol. 4, No. 2, pp. 406–412, 2025.
[19] W. Apriliah *et al.*, "Prediksi Kemungkinan Diabetes pada Tahap Awal menggunakan Algoritma Klasifikasi *Random Forest*," Vol. 10, pp. 163–171, 2021.
[20] H. A. Salman, A. Kalakech, and A. Steiti, "*Random Forest Algorithm Overview*," Vol. 2024, pp. 69–79, 2024.
[21] Y. Y. Jihoon Kim, "*Prediction of IMEP Fluctuation Trends in a Gasoline Engine using Random Forest*," 2024, DOI: 10.1016/j.ifacol.2024.11.136.
[22] A. Raza *et al.*, "*International Journal of Applied Earth Observation and Geoinformation Predicting Regional-Scale Groundwater Levels at High Spatial Resolution using Spatial Random Forest Models*," *Int. J. Appl. Earth Obs. Geoinf.*, Vol. 144, No. October, p. 104918, 2025, DOI: 10.1016/j.jag.2025.104918.
[23] E. Asamoah, G. B. M. Heuvelink, and I. Chairi, "*Heliyon Random Forest Machine Learning for Maize Yield and Agronomic Efficiency Prediction in Ghana*," *Heliyon*, Vol. 10, No. 17, p. e37065, 2024, DOI: 10.1016/j.heliyon.2024.e37065.
[24] M. Kesuma, E. Data, P. Algoritma, P. Algoritma, R. Forest, and D. Mining, "Prediksi Penyakit Liver menggunakan Algoritma *Random Forest*," No. 2, pp. 184–189, 2023.
[25] R. Hidayat *et al.*, "Implementasi Algoritma *Random Forest Regression* untuk memprediksi Penjualan Produk di Supermarket," Vol. 10, No. 1, pp. 101–109, 2025.
[26] J. Bisnis and N. Volume, "Analisis Perbandingan Hasil Prediksi Sistem Pendukung Keputusan Metode *Simple Additive Weighting* dengan *Preference Selection Index* dalam menentukan Mahasiswa Penerima Beasiswa Medi Hermanto Tinambunan , Sri Wahyuni 1 . Fakultas Teknik , Universitas Negeri," No. 2, pp. 765–772, 2023.