

Analisis Perbandingan Algoritma Naïve Bayes, KNN dan Decision Tree untuk Hasil Prediksi Kelulusan Cpnas Auditor Ahli Pertama

Indah Dwijyanthi Nirmala¹, Dwi Budi Srisulistiwati^{2,*}

¹ Program Studi Magister Ilmu Komputer; Universitas Budi LuhurJl. Ciledug Raya, Petungkang Utara, Kota Jakarta Selatan, DKI Jakarta, 12260; e-mail:

2211602392@student.budiluhur.ac.id

²Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya, Jl.Perjuangan No.81, Marga Mulya, Kec.Bekasi Utara, Kota Bekasi, Jawa Barat 17143,

dwibudi@dsn.ubharajaya.ac.id

* Korespondensi: e-mail: dwibudi@dsn.ubharajaya.ac.id

Diterima: 31 Ags 2025; Review: 31 Ags 2025; Disetujui: 31 Ags 2025; Diterbitkan: 31 Ags 2025

Abstract

It cannot be denied that every time there is a CPNS acceptance, many are ready to compete to be able to graduate as a civil servant. To become a civil servant, one of the challenges that must be faced is taking the CAT SKD, SKB test by meeting the CAT threshold scores determined by BKN. Where the final graduation rate for CPNS participants is by getting the highest score from the specified number of formations.

Therefore, the evaluation carried out by researchers in carrying out a comparative analysis of the Decision Tree, KNN, and Naïve Bayes algorithms for the predicted results of the 2021 Ministry of Religion First Expert Auditor CPNS graduation using Rapid Minner Tools. The results of this research will help in providing more effective and efficient solutions in assessing the qualifications of CPNS candidates, as well as minimizing subjective biases that may arise in the manual selection process.

Keywords: Decision Tree, KNN, Naive Bayes, Rapid Minner

Abstrak

Tidak bisa dipungkiri bahwa setiap ada penerimaan CPNS banyak yang siap bersaing untuk bisa lulus sebagai seorang PNS. Untuk menjadi seorang PNS salah satu tantangan yang harus dihadapi adalah mengikuti tes CAT SKD, SKB dengan memenuhi nilai ambang batas CAT yang telah ditentukan oleh BKN. Dimana tingkat kelulusan akhir para peserta CPNS dengan mendapatkan nilai tertinggi dari jumlah formasi yang ditentukan. Maka dari itu evaluasi yang dilakukan peneliti dalam melakukan Analisa Perbandingan Algoritma Naïve Bayes, KNN dan Decision Tree untuk hasil prediksi kelulusan CPNS Auditor Ahli Pertama Kemenag 2021 menggunakan Tools Rapid Minner. Hasil penelitian ini akan membantu dalam memberikan solusi yang lebih efektif dan efisien dalam menilai kualifikasi calon CPNS, serta meminimalkan bias subjektif yang mungkin muncul dalam proses seleksi manual.

Kata kunci: Decision Tree, KNN, Naïve Bayes, Rapid Minner

1. Pendahuluan

Penerimaan Calon Pegawai Negeri Sipil (CPNS) adalah proses yang krusial dalam memastikan bahwa instansi pemerintah memiliki sumber daya manusia berkualitas untuk menjalankan tugas-tugas pelayanan publik (Erwansyah, 2019). Kebutuhan akan sumber daya manusia yang berkualitas dan berdaya saing di era globalisasi ini mengharuskan Pemerintah lebih serius dalam melakukan perekrutan terhadap calon abdi Negara agar perwujudan pelayanan yang baik serta kebutuhan organisasi terhadap kualifikasi jabatan yang ada dapat terpenuhi (Peraturan LAN RI No.12, 2018). Penerapan klasifikasi data mining menjadi solusi yang menarik, dalam rangka mewujudkan penyelenggaraan pemerintahan yang baik (good governance) yang bertujuan pada upaya mengembalikan kepercayaan dan kepuasan publik terhadap sistem rekrutmen CPNS (Erick Junata, 2017).

Penelitian ini mengusulkan penerapan teknik klasifikasi data mining dalam seleksi CPNS menggunakan algoritma Naïve Bayes, K-Nearest Neighbors (KNN) dan Decision Tree. Penelitian ini menggunakan pendekatan kuantitatif, dimana seluruh data disajikan dalam bentuk angka dan dianalisis menggunakan metode analisis statistik, dengan menggunakan sumber data kelulusan CPNS Auditor Ahli Pertama Kemenag. Data yang diperoleh akan diproses dan disiapkan untuk dianalisa dan dilakukan pengujian algoritma Naïve Bayes, KNN dan Decision Tree. Dengan menerapkan teknik klasifikasi dengan menggabungkan algoritma Naïve Bayes, KNN dan Decision Tree menggunakan RapidMiner, diharapkan penelitian ini dapat memberikan solusi yang lebih efektif dan efisien dalam menilai kualifikasi calon CPNS, serta meminimalkan bias subjektif yang mungkin muncul dalam proses seleksi manual.

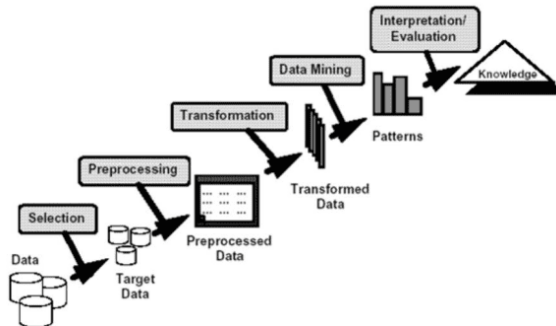
Penerapan klasifikasi data mining untuk seleksi CPNS diharapkan dapat menjadi langkah progresif dalam menghadapi tantangan modern dalam manajemen sumber daya manusia. Penelitian ini juga dapat memberikan kontribusi terhadap pengembangan metodologi seleksi yang lebih canggih dan berbasis data, yang dapat diadopsi oleh instansi pemerintah dalam meningkatkan kualitas sumber daya manusia di dalam birokrasi. Dengan adanya latar belakang ini, penelitian ini bertujuan untuk menggali lebih dalam potensi RapidMiner dalam konteks seleksi CPNS dengan mengimplementasikan teknik klasifikasi data mining dan mengevaluasi kinerja model yang dihasilkan.

2. Metode Penelitian

Dalam penelitian ini menggunakan penelitian kuantitatif, penelitian yang berfokus pada data informasi yang berguna untuk menyimpulkan hasil. Metode penelitian kuantitatif dijelaskan sebagai suatu bentuk kajian sistematis terhadap suatu fenomena yang melibatkan observasi langsung di lapangan, serta pengumpulan data yang telah didapatkan selanjutnya dianalisis memanfaatkan teknik statistik untuk menemukan korelasi antar variabel yang diteliti dan juga membandingkan dua algoritma sebagai penilaian algoritma terbaik (Fauzan et al., 2024).

Analisis Perbandingan Algoritma Naïve Bayes, KNN dan Decision Tree untuk Hasil Prediksi Kelulusan Cpns Auditor Ahli Pertama

Analisis yang diterapkan untuk hasil prediksi kelulusan CPNS Auditor Ahli Pratama dengan metode algoritma Naïve Bayes, KNN dan Decision Tree dengan mengaplikasikan tools rapidminer.



Sumber: (Silangen & Matdoan, 2022)

Gambar 1. Knowledge Discovery in Database

Pada gambar 1 menunjukkan metode Knowledge Discovery in Database (KDD). Menurut (Silangen & Matdoan, 2022), (KDD), merupakan metode untuk mencari dan mengidentifikasi pola dalam data yang dapat digunakan sebagai pengetahuan.

Menurut (Han et al., 2011), Tahapan proses KDD adalah sebagai berikut :

Tabel 1. Tahapan Knowledge Discovery in Database

1.Selection :	merupakan pengambilan data yang berkaitan dengan analisis yang dilakukan dari database.
2.Pre-processing :	merupakan proses pembersihan data dengan menghilangkan duplikasi data.
3.Transformation:	merupakan proses perubahan data menjadi bentuk data yang valid dan siap dilakukan untuk proses data mining.
4. Data Mining :	proses implementasi metode atau algoritma yang sesuai dengan informasi atau pengetahuan yang akan digali.
5.Evaluasi :	merupakan proses yang dilakukan dengan mengevaluasi hasil dari proses data mining.

Sumber: (Han et al., 2011)

Pada tabel 1 menjelaskan tahapan Tahapan Knowledge Discovery in Database. Dari tahapan selection, Pre-processing, tranformation, data mining, dan evaluasi.

Pada tahap evaluasi di penelitian ini, setelah memperoleh hasil dari proses data mining. Pada tahap ini melibatkan evaluasi dan interpretasi hasil tersebut dengan tujuan memahami makna dan relevansi dari pola atau model yang telah diidentifikasi (Triwidianti et al., 2021). Hasil evaluasi dilandasi pada Confusion Matrix. Evaluasi diperlukan untuk menampilkan hasil dengan memperlihatkan accuracy, precision dan recall (Firmansyach et al., 2023).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

Sumber: Hasil Penelitian (2025)

Gambar 2. Rumus Perhitungan *Accuracy*, *Precision*, *Recall*

Pada gambar 2 menunjukkan rumus untuk menghitung nilai memperlihatkan accuracy, precision dan recall. Terdapat pada rumus 1,2,3 (Pujiyanto & Ristanti, 2019).

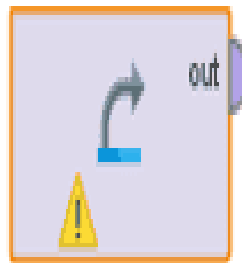
3. Hasil dan Pembahasan

Penelitian yang dilakukan mengikuti langkah-langkah dalam metode Knowledge Discovey in Database (KDD).

3.1. *Data Selection*

Dataset yang digunakan dalam penelitian ini adalah data sekunder. Dokumen digunakan dalam penelitian sebagai sumber data sekunder manakala dokumen tersebut memiliki nilai. Data sekunder diperoleh secara tidak langsung dari orang lain, kantor yang berupa laporan, profil, buku pedoman, atau pustaka (Priadana & Sunarsi, 2021).

Retrieve HASIL INTEGRASI SKD DAN SKB FORMASI AUDITOR AHLI PERTAMA KEMENAG



Sumber: Hasil Penelitian (2025)

Gambar 3. Dataset

Pada gambar 3 menunjukkan dataset hasil akhir kelulusan CPNS Auditor Ahli Pertama Kemenag yang digunakan dalam penelitian ini diperoleh dari website <https://kemenag.go.id/informasi/hasil-akhir-seleksi-calon-pegawai-negeri-sipil-cpns--kementerian-an-agama>

Analisis Perbandingan Algoritma Naïve Bayes, KNN dan Decision Tree untuk Hasil Prediksi Kelulusan Cpnas Auditor Ahli Pertama

Row No.	Kotorangan	No Peserta	Nama	Tanggal Lahir	Kode Pendid.	Pendidikan	Nilai BPK	TWK	TIU	TRP	Total	Skor SKD (L., SKB)
1	LULUS	2130122120...	KHABIBULLAH	Oct 9, 1994	5101057	TEKNIK INF...	3.910	120	155	190	465	33.618 79.325
2	LULUS	2130122110...	MUHAMMAD S.	May 16, 1998	5102959	HUKUM KEL...	3.730	115	155	192	462	33.660 79.520
3	LULUS	2130122110...	ABU YAZID B.	Aug 7, 1993	5101059	PENDIDIKAN	3.670	110	150	190	450	32.727 80.933
4	LULUS	2130122120...	WINIARTI	Mar 15, 1997	5140219	PENDIDIKAN	3.670	110	150	191	451	32.890 80.325
5	LULUS	2130122120...	DANI BETHAWI	Jan 9, 1994	5101058	PENDIDIKAN	3.590	120	150	191	461	33.527 77.983
6	LULUS	2130122110...	CHANDRA NULU	Jul 11, 1990	5140219	PENDIDIKAN	3.240	130	140	188	456	33.164 78.650
7	LULUS	2130122120...	BEBSE FATMA	Aug 10, 1989	5101058	PENDIDIKAN	3.850	125	155	181	471	34.255 76.438
8	TIDAK LULUS	2130122120...	HIRIKAH LES	Apr 13, 1994	5101058	PENDIDIKAN	3.590	115	150	188	454	33.745 77.035
9	TIDAK LULUS	2130122120...	KEN DITE AS	May 19, 1993	5101058	PENDIDIKAN	3.590	120	150	201	471	34.255 76.130
10	TIDAK LULUS	2130122110...	ADITYA WAH	Feb 15, 1994	5101058	PENDIDIKAN	3.360	115	150	188	453	32.945 78.175
11	TIDAK LULUS	2130122120...	LATIFA FITRI	May 8, 1989	5101058	PENDIDIKAN	3.290	115	150	181	456	33.164 75.617
12	TIDAK LULUS	2130122120...	EKA SUKCI PU	May 2, 1997	5101058	PENDIDIKAN	3.520	120	150	187	457	33.236 74.538
13	TIDAK LULUS	2130122110...	MUHAMMAD ...	Oct 24, 1998	5110832	PENDIDIKAN	3.700	100	155	198	453	32.945 73.825
14	TIDAK LULUS	2130122120...	EMIL NADIA	Apr 4, 1987	5101058	PENDIDIKAN	3.890	130	140	181	451	32.890 73.828
15	TIDAK LULUS	2130122110...	ZIA SHAH RE	May 5, 1995	5101059	PENDIDIKAN	3.350	105	155	200	450	33.485 72.125
16	TIDAK LULUS	2130122110...	MUHAMMAD ...	Aug 15, 1993	5140322	KOMUNIKAS...	3.290	105	150	197	452	32.873 71.383
17	TIDAK LULUS	2130122120...	ELA FARRAH	Dec 17, 1993	5101059	PENDIDIKAN	3.280	105	155	198	458	34.036 80.280
18	TIDAK LULUS	2130122110...	DANI RAHMAN	Oct 31, 1992	5101059	PENDIDIKAN	3.580	130	130	192	452	32.873 70.213
19	TIDAK LULUS	2130122120...	INDGAR GAL...	Sep 5, 1993	5110822	PENDIDIKAN	3.480	120	145	187	452	32.873 69.777
20	TIDAK LULUS	2130122120...	YULIANA	Jul 12, 1990	5101059	PENDIDIKAN	3.390	110	135	205	450	32.727 69.035
21	TIDAK LULUS	2130122120...	ANISATUL FA...	Jan 15, 1995	5101059	PENDIDIKAN	3.680	105	155	184	454	33.018 69.998

Sumber: Hasil Penelitian (2025)

Gambar 4. Hasil Read Excel

Pada gambar 4 menunjukkan Hasil Read Excel.

3.2. Pre-Processing

Pada tahapan ini data yang tidak relevan, missing value, dan redunant harus dibersihkan. Dataset yang digunakan sudah tidak ada data yang tidak lengkap/sesuai, hilang, maka data dapat digunakan ke proses selanjutnya.

3.3. Transformation

Transformasi dataset menjadi bentuk yang lebih cocok agar data dapat diproses ke dalam aplikasi Data Mining, sehingga lebih mudah untuk mendapatkan pengetahuan dari dataset. Hasil dari proses transformasi data, sebagai berikut:

Row No.	Kotorangan	No Peserta	Nama	Tanggal Lahir	Kode Pendid.	Pendidikan	Nilai BPK	TWK	TIU	TRP	Total	Skor SKD (L., SKB)
1	LULUS	2130122120...	KHABIBULLAH	Oct 9, 1994	5101057	TEKNIK INF...	3.910	120	155	190	465	33.618 79.325
2	LULUS	2130122110...	MUHAMMAD S.	May 16, 1998	5102959	HUKUM KEL...	3.730	115	155	192	462	33.660 79.520
3	LULUS	2130122110...	ABU YAZID B.	Aug 7, 1993	5101059	PENDIDIKAN	3.670	110	150	190	450	32.727 80.933
4	LULUS	2130122120...	WINIARTI	Mar 15, 1997	5140219	PENDIDIKAN	3.670	110	150	191	451	32.890 80.325
5	LULUS	2130122120...	DANI BETHAWI	Jan 9, 1994	5101058	PENDIDIKAN	3.600	120	150	191	451	33.527 77.983
6	LULUS	2130122110...	CHANDRA NULU	Jul 11, 1990	5140219	PENDIDIKAN	3.240	130	140	188	456	33.164 78.650
7	LULUS	2130122120...	BEBSE FATMA	Aug 10, 1989	5101058	PENDIDIKAN	3.850	125	155	181	471	34.255 76.438
8	TIDAK LULUS	2130122120...	HIRIKAH LES	Apr 13, 1994	5101058	PENDIDIKAN	3.590	115	150	188	454	33.745 77.035
9	TIDAK LULUS	2130122120...	KEN DITE AS	May 19, 1993	5101058	PENDIDIKAN	3.590	120	150	201	471	34.255 76.130
10	TIDAK LULUS	2130122110...	ADITYA WAH	Feb 15, 1994	5101058	PENDIDIKAN	3.360	115	150	188	453	32.945 78.175
11	TIDAK LULUS	2130122120...	LATIFA FITRI	May 8, 1989	5101058	PENDIDIKAN	3.290	115	150	181	456	33.164 75.617
12	TIDAK LULUS	2130122120...	EKA SUKCI PU	May 2, 1997	5101058	PENDIDIKAN	3.520	120	150	187	457	33.236 74.538
13	TIDAK LULUS	2130122110...	MUHAMMAD ...	Oct 24, 1998	5110832	PENDIDIKAN	3.700	100	155	198	453	32.945 73.825
14	TIDAK LULUS	2130122120...	EMIL NADIA	Apr 4, 1987	5101058	PENDIDIKAN	3.890	130	140	181	451	32.890 73.828
15	TIDAK LULUS	2130122110...	ZIA SHAH RE	May 5, 1995	5101059	PENDIDIKAN	3.350	105	155	200	450	33.485 72.125
16	TIDAK LULUS	2130122110...	MUHAMMAD ...	Aug 15, 1993	5140322	KOMUNIKAS...	3.290	105	150	197	452	32.873 71.383
17	TIDAK LULUS	2130122120...	ELA FARRAH	Dec 17, 1993	5101059	PENDIDIKAN	3.280	105	155	198	458	34.036 80.280
18	TIDAK LULUS	2130122110...	DANI RAHMAN	Oct 31, 1992	5101059	PENDIDIKAN	3.580	130	130	192	452	32.873 70.213
19	TIDAK LULUS	2130122120...	INDGAR GAL...	Sep 5, 1993	5110822	PENDIDIKAN	3.480	120	145	187	452	32.873 69.777
20	TIDAK LULUS	2130122120...	YULIANA	Jul 12, 1990	5101059	PENDIDIKAN	3.390	110	135	205	450	32.727 69.035
21	TIDAK LULUS	2130122120...	ANISATUL FA...	Jan 15, 1995	5101059	PENDIDIKAN	3.680	105	155	184	454	33.018 69.998

Sumber: Hasil Penelitian (2025)

Gambar 5. Hasil Proses Transformasi Data

Pada gambar 5 menunjukkan gambaran umum dari objek penelitian kerangka kerja Architecture Enterprise yang berbeda. Dimana setiap kerangka kerja memiliki konteks, aplikasi, dan penggunaan khususnya, dan pemilihan kerangka kerja tergantung pada tujuan dan kebutuhan spesifik organisasi.

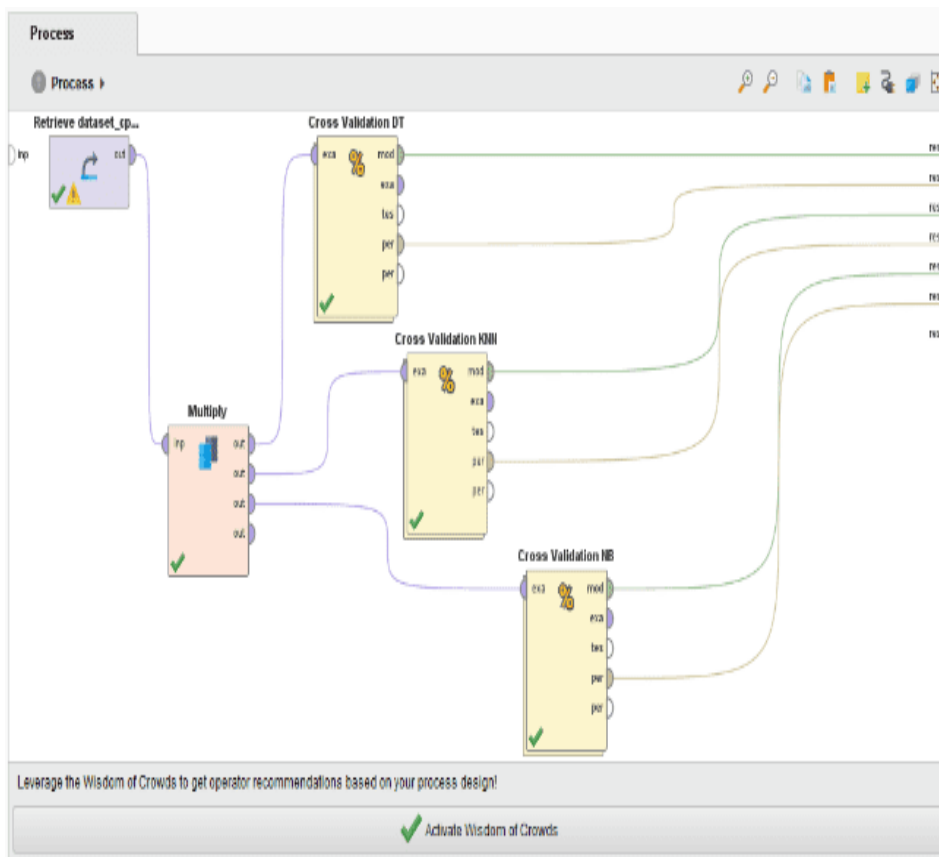
3.4. Data Mining

Transformasi dataset menjadi bentuk yang lebih cocok agar data dapat diproses ke dalam aplikasi Data Mining, sehingga lebih mudah untuk mendapatkan pengetahuan dari dataset. Hasil dari proses transformasi data, sebagai berikut:

Tahapan implementasi ini menggunakan metode klasifikasi algoritma Naïve Bayes, KNN dan Decision Tree perhitungan menggunakan tools Rapidminer Studio versi 10.3. Pengujian data sebanyak 21 record yang telah melalui proses *data selection*, *data preprocessing* dan *data transformation*.

3.4.1. Analisa Data dari Algoritma Data Mining

Implementasi dilakukan dengan menyimpan semua atribut indikator ke dalam format xlsx, kemudian *import dataset* dengan pilih *Import data* dari Rapidminer Studio.

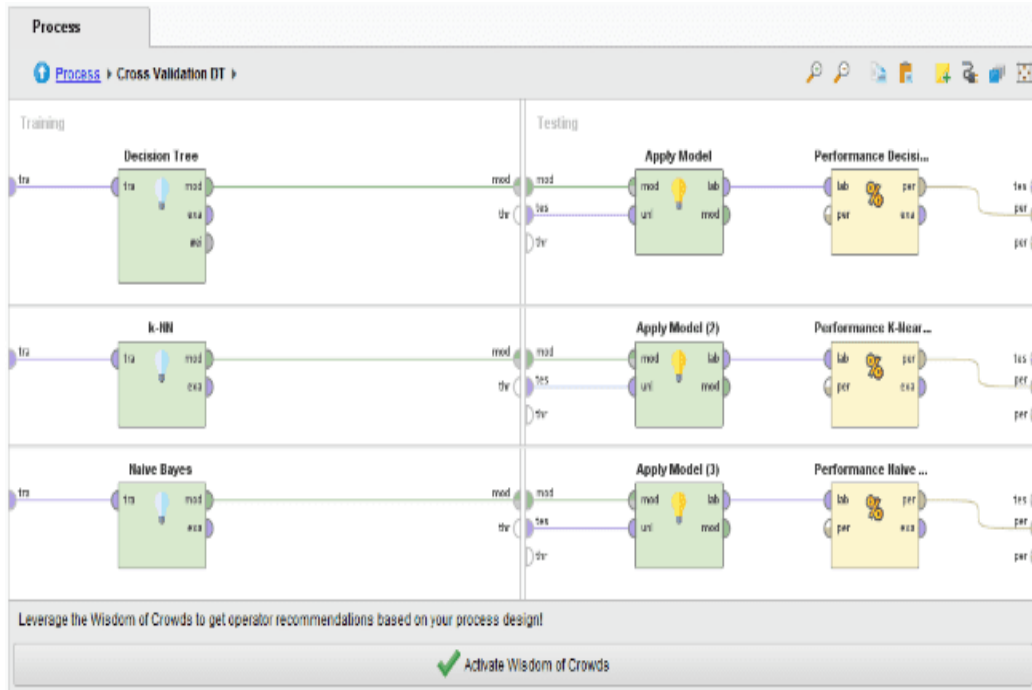


Sumber: Hasil Penelitian (2025)

Gambar 6. Model Proses Desain Import Data

Pada gambar 6 menunjukkan Model Proses Desain Import Data.

Analisis Perbandingan Algoritma Naïve Bayes, KNN dan Decision Tree untuk Hasil Prediksi Kelulusan Cpns Auditor Ahli Pertama



Sumber: Hasil Penelitian (2025)

Gambar 7. Model Validasi *Naive Bayes*, *K-NN* dan *Decision Tree*

Pada gambar 7 menunjukkan Model Validasi *Naive Bayes*, *K-NN* dan *Decision Tree*.

3.5. Evaluasi

Transformasi dataset menjadi bentuk yang lebih cocok agar data dapat diproses ke dalam aplikasi Data Mining, sehingga lebih mudah untuk mendapatkan pengetahuan dari dataset. Hasil dari proses transformasi data, sebagai berikut:

Pada tahapan evaluasi, dilakukan pengujian menggunakan cross validation dari algoritma yang digunakan untuk klasifikasi. Hasil pengujian tersebut akan dibandingkan nilai akurasi, presisi, dan recall dari algoritma Decision Tree, kNN, dan Naive Bayes. Setelah itu, pengujian juga dilanjutkan dengan metode T-test untuk melihat algoritma yang terbaik.

- Hasil Evaluasi Model *Naive Bayes*

Pengujian pada Model Naive Bayes menggunakan split data, hasil dari pengujian split data dengan 80% data training dan 20% data testing. Hasil prediksi menggunakan model Naive Bayes, yakni mendapatkan hasil akurasi prediksi lulus dan true lulus sebesar 2 data, prediksi lulus dan true tidak lulus sebesar 8 data, prediksi tidak lulus dan true lulus sebesar 5 data, prediksi tidak lulus dan true tidak lulus sebesar 6 data.

Criterion		Table View		Pod View	
Performance	accuracy	accuracy: 36.67% +/- 25.82% (micro average: 38.10%)			
Description	pred LULUS	2	8	class precision	
	pred TIDAK LULUS	5	9	54.55%	
	class recall	28.57%	42.86%		

Sumber: Hasil Penelitian (2025)

Gambar 8. Hasil akurasi *Naive Bayes*

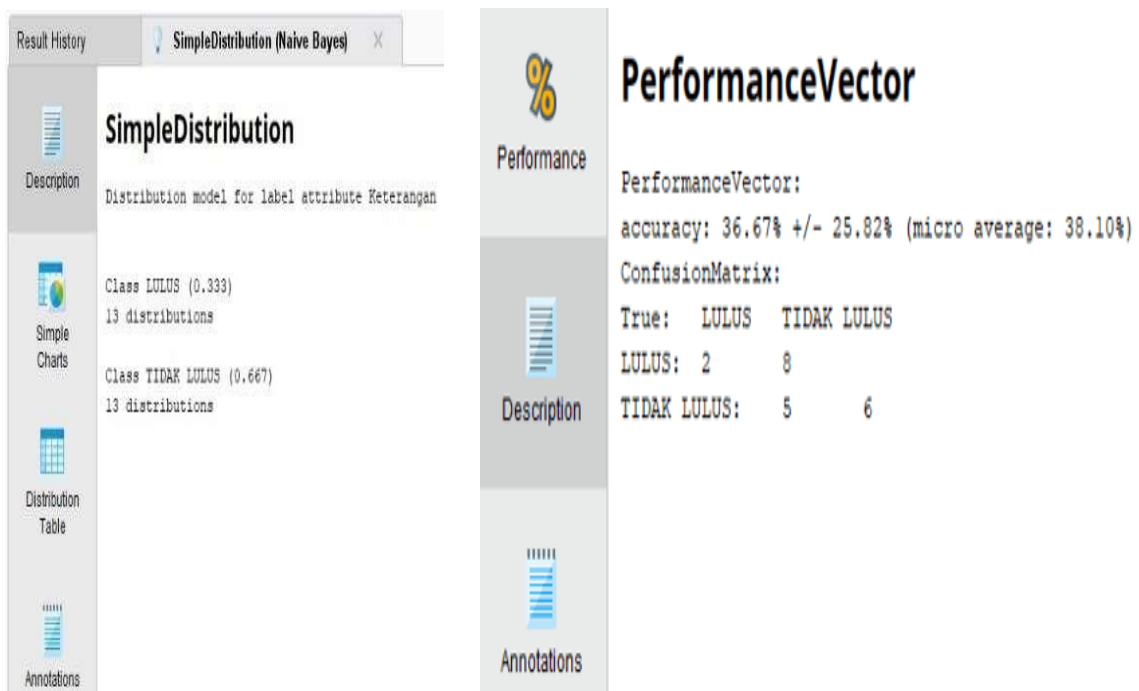
Pada gambar 8 menunjukkan Hasil akurasi *Naive Bayes*.

$$\begin{aligned} \text{Accuracy} &= \frac{2+6}{2+8+5+6} \\ &= 0,380952 \text{ (38,09\%)} \\ \text{Precision} &= \frac{6}{6+8} \\ &= 0,428571 \text{ (42,86\%)} \\ \text{Recall} &= \frac{6}{6+5} \\ &= 0,545455 \text{ (54,55\%)} \end{aligned}$$

Sumber: Hasil Penelitian (2025)

Gambar 9. Hasil perhitungan Confusion Matrix accuracy, precision, recall pada algoritma Naive Bayes

Pada gambar 9 menunjukkan hasil perhitungan Confusion Matrix accuracy, precision, recall pada algoritma Naive Bayes.



Sumber: Hasil Penelitian (2025)

Gambar 10. Hasil Performance Vector pada Naive Bayes

Pada gambar 10 menunjukkan Hasil Performance Vector pada Naive Bayes, sehingga terbentuk hasil matrix akurasi pada klasifikasi data hasil akhir kelulusan CPNS Auditor Ahli Pertama Kemenag Tahun 2021 dengan hasil sebesar 36.67% dan hasil yang lulus sebanyak 2 orang dan yang tidak lulus sebanyak 19 orang.

- Hasil Evaluasi Model KNN

	true LULUS	true TIDAK LULUS	class precision
pred LULUS	0	2	0.00%
pred TIDAK LULUS	7	12	63.16%
class recall	0.00%	85.71%	

Sumber: Hasil Penelitian (2025)

Gambar 11. Hasil akurasi KNN

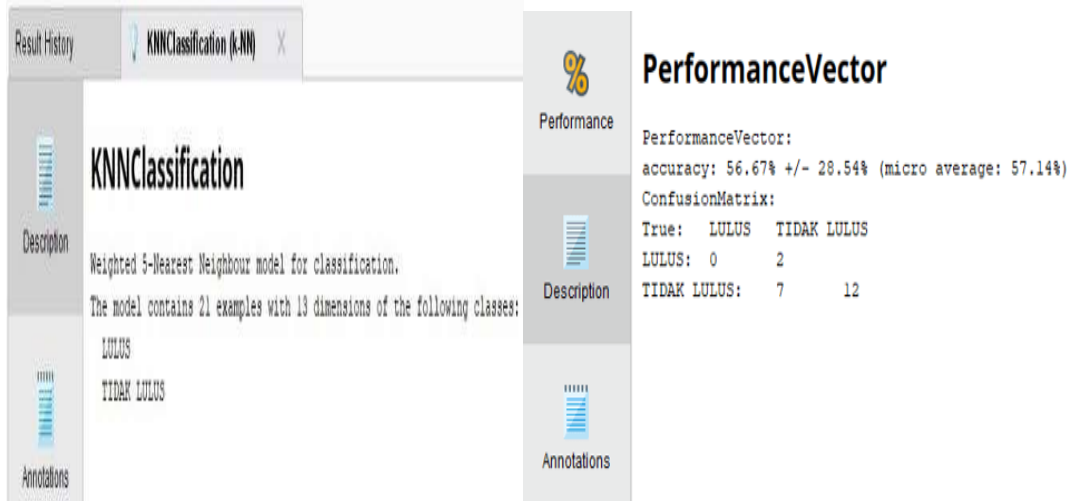
Pada gambar 11 menunjukkan Pengujian Model KNN menggunakan split data, hasil pengujian split data yakni 80% data training dan 20% data testing. Hasil prediksi menggunakan model KNN, yakni mendapatkan hasil akurasi prediksi lulus dan true lulus sebesar 0 data, prediksi lulus dan true tidak lulus sebesar 2 data, prediksi tidak lulus dan true lulus sebesar 7 data, prediksi tidak lulus dan true tidak lulus sebesar 12 data.

$$\begin{aligned}
 \text{Accuracy} &= \frac{12+0}{12+2+7+0} \\
 &= 0,571429 \text{ (57,14\%)} \\
 \text{Precision} &= \frac{12}{12+7} \\
 &= 0,631579 \text{ (63,16\%)} \\
 \text{Recall} &= \frac{12}{12+2} \\
 &= 0,857143 \text{ (85,71\%)}
 \end{aligned}$$

Sumber: Hasil Penelitian (2025)

Gambar 12. Hasil perhitungan Confusion Matrix accuracy, precision, recall pada algoritma KNN

Pada gambar 12 menunjukkan Hasil perhitungan Confusion Matrix accuracy, precision, recall pada algoritma KNN.

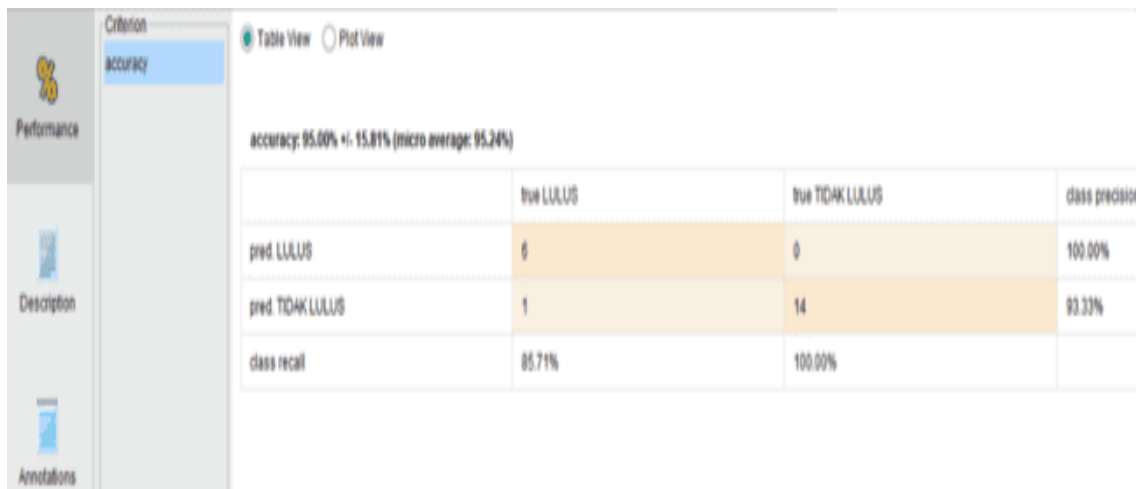


Sumber: Hasil Penelitian (2025)

Gambar 13. Hasil Performance Vector pada KNN

Pada gambar 13 menunjukkan erentuk hasil matrix akurasi pada klasifikasi data hasil akhir kelulusan CPNS Auditor Ahli Pertama Kemenag Tahun 2021 cukup dengan hasil sebesar 56.67% dengan hasil yang lulus sebanyak 0 orang dan yang tidak lulus sebanyak 21 orang.

- Hasil Evaluasi Model Decision Tree



Sumber: Hasil Penelitian (2025)

Gambar 14. Hasil akurasi Decision Tree

Pada gambar 14 menunjukkan Pengujian pada Model Decision Tree menggunakan split data, hasil dari pengujian split data dengan 80% data training dan 20% data testing. Hasil prediksi menggunakan model Decision Tree, yakni mendapatkan hasil akurasi prediksi lulus dan true lulus sebesar 6 data, prediksi lulus dan true tidak lulus sebesar 0 data, prediksi tidak lulus dan true lulus sebesar 1 data, prediksi tidak lulus dan true tidak lulus sebesar 14 data.

$$\begin{aligned} \text{Accuracy} &= \frac{14+6}{14+0+1+6} \\ &= 0,952381 \text{ (95,24\%)} \end{aligned}$$

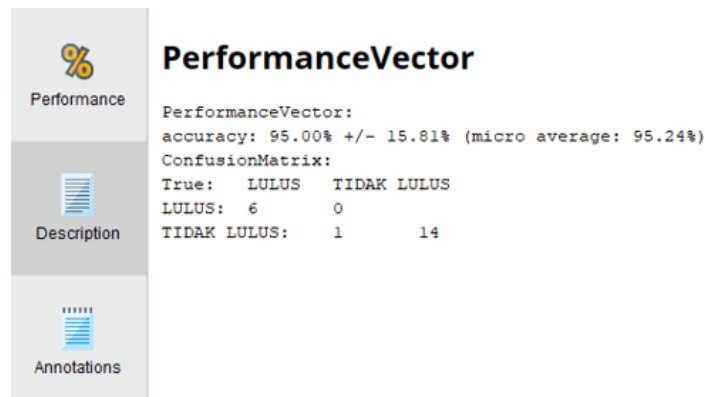
$$\begin{aligned} \text{Precision} &= \frac{14}{14+1} \\ &= 0,933333 \text{ (93,33\%)} \end{aligned}$$

$$\begin{aligned} \text{Recall} &= \frac{14}{14+0} \\ &= 1 \text{ (100,00\%)} \end{aligned}$$

Sumber: Hasil Penelitian (2025)

Gambar 15. Hasil perhitungan Confusion Matrix accuracy, precision, recall pada algoritma Decision Tree

Pada gambar 15 menunjukkan hasil perhitungan Confusion Matrix accuracy, precision, recall pada algoritma Decision Tree.



Sumber: Hasil Penelitian (2025)

Gambar 16. Hasil Performance Vector pada Decision Tree

Pada gambar 16 terbentuk hasil matrix akurasi pada klasifikasi data hasil akhir kelulusan CPNS Auditor Ahli Pertama Kemenag sangat baik dengan hasil sebesar 95.00% dengan hasil yang lulus sebanyak 6 orang dan yang tidak lulus sebanyak 15 orang.

- Hasil Komparasi 3 Algoritma

Dari hasil komparasi tersebut menunjukkan Decision tree memiliki tingkat akurasi yang paling tinggi. Hal tersebut menunjukkan bahwa kinerja Decision Tree lebih baik dibanding dengan Naive Bayes dan K-Nearest Neighbor.

4. Kesimpulan

Dari hasil proses pengujian dengan tools RapidMiner Menggunakan tiga metode yang telah dilakukan. Decision Tree (C4.5) memperoleh hasil akurasi tertinggi sebesar 90.00% pada pengujian ke 3. Metode Naïve Bayes memperoleh hasil akurasi tertinggi sebesar 61.67% pada pengujian ke 1, dan Metode K-Nearest Neighbor (K-NN) memperoleh hasil akurasi tertinggi sebesar 61.67% pada pengujian ke 2.

Daftar Pustaka

- Erick Junata, S. (2017). *Seleksi Cpnns Menurut Peraturan Menteri Pendayagunaan Aparatur Negara Dan Reformasi*. 5, 1–4.
- Erwansyah, K. (2019). Implementasi Data Mining Untuk Menganalisa Hubungan Data Penjualan Produk Bahan Kimia Terhadap Persediaan Stok Barang Menggunakan Algoritma FP (Frequent Pattern) Growth Pada PT. Grand Multi Chemicals. v, 30(2), 30–40.
- Fauzan, A. S., Irma, A., Sari, P., & Ali, I. (2024). Analisis Perbandingan Algoritma Decision Tree Dan Naïve Untuk Mengevaluasi Prestasi Belajar Siswa Studi Kasus: Smk Al-Musyawirin. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Issue 1).
- Firmansyach, W. A., Hayati, U., & Wijaya, Y. A. (2023). Analisa Terjadinya Overfitting Dan Underfitting Pada Algoritma Naive Bayes Dan Decision Tree Dengan Teknik Cross Validation. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 7, Issue 1).
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)*.
- Peraturan LAN RI No.12. (2018). Peraturan LAN RI No 12 Thn 2018 Tentang Pelatihan Dasar CPNS. *Berita Negara Republik Indonesia Tahun 2018 Nomor 1800*, 108.
- Priadana, S., & Sunarsi, D. (2021). *Metode-Penelitian-Kuantitatif*.
- Pujianto, U., & Ristanti, P. Y. (2019). *Jurusan Teknik Elektro, Universitas Negeri Malang, Indonesia | Maret 2019 U. Pujianto, Putri Yuni Ristanti | Perbandingan kinerja metode C4.5 dan Naive Bayes dalam klasifikasi ...* (Vol. 29). <http://journal2.um.ac.id/index.php/tekno>
- Silangen, R. T., & Matdoan, M. Y. (2022). Klasifikasi Hasil Seleksi Kompetensi Dasar CPNS Menggunakan Metode Decision Tree. *Inferensi*, 5(2), 69. <https://doi.org/10.12962/j27213862.v5i2.12353>
- Triwidiyanti, J., Alfian, F. Y., & Prasojo, M. (2021). *Perbandingan Metode Data Mining Untuk Prediksi Prestasi Siswa Tingkat Pendidikan Menengah Kejuruan Pada Sekolah Menengah Kejuruan Negeri (SMKN 1) Gadingrejo Pringsewu Lampung*.