



Perbandingan Model Klasifikasi C4.5, Naïve Bayes, Support Vector Machine dan K-nearest Neighbour untuk Memprediksi Kelayakan Masyarakat dalam Menerima Bantuan PBI APBD

Comparison of Classification Models C4.5, Naïve Bayes, Support Vector Machine and K-nearest Neighbour to Predict Community Eligibility to Receive PBI APBD Assistance

Tutik Ultsa Rahmatika¹, Nur Alamsyah^{2*}, Titan Parama Yoga³, Budiman⁴

¹Sistem Informasi, Fakultas Teknologi Dan Informatika, Universitas Informatika dan Bisnis Indonesia

¹tutikultsa06@gmail.com, ²nuralamsyah.bdg@gmail.com, ³titanparama@unibi.ac.id, ⁴budiman@unibi.ac.id

Abstract

This research evaluates the eligibility of the community to receive APBD Contribution Assistance (PBI) using four classification algorithms: C4.5, Naïve Bayes, K-Nearest Neighbor, and Support Vector Machine (SVM). There is a problem of inaccurate distribution of assistance, which prompted the selection of these four methods with specific considerations, C4.5 (Decision Tree) is known for its clarity and interpretability, providing an easy-to-understand understanding of the factors that influence classification decisions, Naïve Bayes was selected for its efficiency and speed in training and testing, suitable for large datasets and can be updated quickly with new data, K-Nearest Neighbor (KNN) is used for decision making based on local patterns in the data, useful if the eligibility decision is local or related to the surrounding environment while Support Vector Machine (SVM): Selected for its ability to handle complex and non-linear datasets. The results show that SVM has the highest Weighted Mean Precision, reaching 91.67%, confirming its superiority as the best choice. These findings make a significant contribution to improving the accuracy of determining the eligibility of PBI APBD beneficiaries, supporting targeting accuracy, and ensuring the effectiveness of the assistance program for people in need

Keywords: PBI APBD, C4.5, NB, KNN, SVM

Abstrak

Penelitian ini mengevaluasi kelayakan masyarakat menerima Bantuan Iuran (PBI) APBD menggunakan empat algoritma klasifikasi: C4.5, Naïve Bayes, K-Nearest Neighbor, dan Support Vector Machine (SVM). Terdapat masalah ketidaktepatan penyaluran bantuan, yang mendorong pemilihan keempat metode ini dengan pertimbangan khusus, C4.5 (Decision Tree) diketahui untuk kejelasan dan interpretabilitasnya, memberikan pemahaman yang mudah dipahami terhadap faktor-faktor yang mempengaruhi keputusan klasifikasi, Naïve Bayes dipilih karena efisiensi dalam kecepatan dalam pelatihan dan pengujian, cocok untuk dataset besar dan dapat diupdate secara cepat dengan data baru, K-Nearest Neighbor (KNN) digunakan untuk pengambilan keputusan berbasis pola lokal dalam data, berguna jika keputusan kelayakan bersifat lokal atau terkait dengan lingkungan sekitarnya sedangkan Support Vector Machine (SVM): Diseleksi karena kemampuannya menangani dataset kompleks dan non-linear. Hasil penelitian menunjukkan bahwa SVM memiliki Weighted Mean Precision tertinggi, mencapai 91.67%, menegaskan keunggulannya sebagai pilihan terbaik. Temuan ini memberikan kontribusi signifikan dalam meningkatkan akurasi penentuan kelayakan penerima bantuan PBI APBD, mendukung ketepatan sasaran, dan memastikan efektivitas program bantuan bagi masyarakat yang membutuhkan.

Kata kunci : PBI APBD, C4.5, NB, KNN, SVM

1. Pendahuluan

Kemiskinan merupakan tantangan serius di Indonesia yang membutuhkan pendekatan holistik dan solusi yang tepat [1]. Dalam amanat Undang-Undang Dasar 1945, penanggulangan kemiskinan diutamakan sebagai

langkah krusial bagi pembangunan berkelanjutan [2]. Data terbaru Badan Pusat Statistik (BPS) mencatat tingkat kemiskinan Indonesia pada September 2022 sebesar 9,57%, meningkat 0,03% dari Maret 2022, meskipun mengalami penurunan sebesar 0,14% dibandingkan September 2021. Kabupaten Bandung

juga menunjukkan perubahan, dari 7,15% pada 2021 menjadi 6,80% pada 2022.

Pemerintah merespons permasalahan ini dengan sejumlah program bantuan, termasuk Program Bantuan Iuran Anggaran Pendapatan dan Belanja Daerah (PBI APBD), yang bertujuan memberikan dukungan kesehatan kepada peserta Jaminan Kesehatan Nasional (JKN). Namun, kendala utama terletak pada ketidakpastian penyaluran bantuan dan menentukan kelayakan penerima [3]. Oleh karena itu, penelitian ini merinci penggunaan metode klasifikasi data mining, yaitu Algoritma C4.5, Naïve Bayes, K-Nearest Neighbour, dan Support Vector Machine (SVM).

Melalui telaah literatur, keempat metode tersebut telah terbukti efektif dalam berbagai konteks [4]. Contohnya, C4.5 dikenal untuk memberikan aturan keputusan yang jelas dan interpretabil, Naïve Bayes unggul dalam efisiensi pengolahan data, K-Nearest Neighbour efektif dalam pengenalan pola lokal, dan SVM mampu menangani kompleksitas data yang bersifat non-linear [5]. Keunggulan masing-masing metode ini akan menjadi landasan utama untuk mengidentifikasi pola dan informasi penting dalam data terkait kelayakan penerima bantuan PBI APBD.

Seiring dengan berkembangnya literatur yang mendukung efektivitas keempat model tersebut, penelitian ini memberikan kontribusi dengan fokus pada konteks kemiskinan di Indonesia [6]. Urgensi penelitian ini terletak pada kebutuhan mendalam untuk meningkatkan akurasi penentuan kelayakan penerima bantuan, yang pada gilirannya akan mengoptimalkan efektivitas program PBI APBD dan mendukung tujuan penanggulangan kemiskinan secara nasional. Dengan merinci dan membandingkan metode klasifikasi yang berbeda, diharapkan penelitian ini dapat menjadi landasan penting bagi perbaikan sistem penyaluran bantuan dan kontribusi nyata terhadap kesejahteraan masyarakat yang membutuhkan.

2. Metode Penelitian

Gambar 1 adalah alur penelitian yang terdiri dari dua metode yaitu metode pengumpulan data dan metode simulasi.

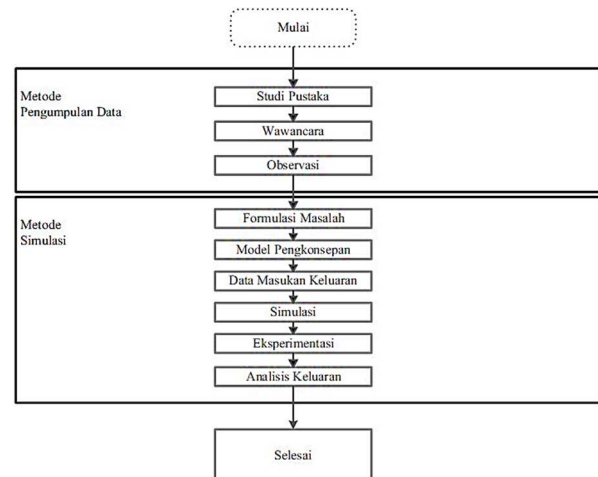
Pada metode pengumpulan data terdiri dari tiga tahapan yaitu studi pustaka, wawancara dan observasi. Sedangkan untuk metode simulasi terdiri dari enam tahapan yaitu formulasi masalah, model pengkonsepkan, data masukan/keluaran, simulasi, eksperimentasi dan analisis keluaran [7].

2.1. Pengumpulan Data

Dalam penelitian ini penulis membutuhkan data serta informasi untuk menjadi bahan pembahasan. Dengan demikian penulis melakukan penelitian terbanyak dahulu untuk mencari data serta informasi tersebut.

Adapun metode penelitian yang penulis lakukan adalah sebagai berikut :

Pada tahap pengumpulan data dengan studi pustaka ini penulis mencari referensi-referensi yang telah ada sebelumnya yang mempunyai hubungan dengan objek yang akan diteliti. Pencarian bersumber dari buku, ebook, jurnal dan studi literatur sejenis [8].



Gambar 1. Tahapan Penelitian

Pada tahap wawancara ini penulis mewawancarai Ibu Karnita selaku pegawai PUSKESOS (Pusat Kesejahteraan Sosial). Wawancara ini bertujuan untuk menggali informasi banyak dalam yang berkenaan mengenai kelayakan penerima bantuan PBI APBD seperti kriteria calon penerima serta data-data yang pernah ada [9].

Pada tahap observasi ini penulis langsung mengunjungi lokasi kantor Kecamatan Pameungpeuk yang beralamat Jl. Raya Banjaran Barat, Sukasari, Kecamatan Pameungpeuk, Kabupaten Bandung [10].

2.2. Metode Simulasi.

Pada penelitian ini, penulis menggunakan metode simulasi sebagai perbandingan tingkat Accuracy, Weighted Mean Recall, Weighted Mean Precision dan Execution Time pada algoritma C4.5, Naïve Bayes, Support Vector Machine (SVM) dan K-Nearest Neighbour (KNN) [11]. Adapun metode simulasi ini dibagi menjadi enam tahap yaitu [12]:

Tahapan Formulasi Masalah (*Problem Formulation*) merupakan tahap awal dalam metode simulasi. Setelah penulis melakukan pengumpulan data baik itu melalui studi pustaka, wawancara, dan observasi, penulis mendapatkan formulasi inti masalah yaitu perbandingan algoritma C4.5 dan Naïve Bayes untuk klasifikasi bantuan. Pada perbandingan algoritma ini parameter yang dijadikan acuan adalah tingkat Accuracy, Weighted Mean Recall, Weighted Mean Precision dan Execution Time yang dihasilkan dari tool RapidMiner Studio.

Konsep Model (*Conceptual Mode*): Setelah penulis mendapatkan masalah, penulis merancang konsep simulasi yang akan dibuat. Pada penelitian kali ini penulis menggunakan RapidMiner Studio untuk mensimulasikan studi kasus yang akan penulis teliti.

Data Masukan/Keluaran (*Input/Output Data*): Pada tahapan ini penulis menentukan input untuk proses dan output untuk mendapatkan hasilnya. Adapun input yang dibutuhkan adalah data dummy dari klasifikasi bantuan yang berisi atributnya berisi pekerjaan dan kondisi rumah, sedangkan output yang dihasilkan adalah nilai Accuracy, Weighted Mean Recall, Weighted Mean Precision dan Execution Time.

Simulasi (*Simulation*): Pada tahapan ini penulis menerapkan model yang telah ditentukan sebelumnya. Selanjutnya model tersebut akan disimulasikan dengan atribut-atribut yang telah ditentukan. Proses simulasi ini menggunakan RapidMinerStudio serta algoritma C4.5, Naïve Bayes, KNN dan SVM. Lalu hasil simulasi tersebut akan dicatat dan akan dilakukan tahap verifikasi.

Eksperimen (*Experiment*): Selanjutnya penulis akan melakukan eksperimentasi terhadap permodelan yang telah dibuat sebelumnya.

Analisis Keluaran (*Output Analysis*): Ini merupakan tahap akhir dari metode simulasi, penulis akan menganalisis hasil dari simulasi yang dilakukan pada tahapan eksperimentasi. Untuk memudahkan menganalisisnya maka hasil eksperimentasi tersebut akan disajikan dalam bentuk tabel dan grafik yang menyatakan nilai Accuracy, Weighted Mean Recall, Weighted Mean Precision dan Execution Time dari masing-masing algoritma. Pada hasil akhirnya penulis akan membandingkan algoritma C4.5, Naïve Bayes, KNN dan SVM untuk melihat keefektifan algoritma tersebut pada simulasi ini berdasarkan nilai Accuracy, Weighted Mean Recall, Weighted Mean Precision dan Execution Time.

3. Hasil dan Pembahasan

3.1. Formulasi Permasalahan.

Pada penelitian ini menggunakan *tool RapidMiner Studio*. *RapidMiner Studio* merupakan salah satu *tools* yang sangat bermanfaat sekali khususnya untuk menganalisa *data mining*, *text mining*, maupun Analisa prediksi. Selain itu *RapidMiner Studio* juga bersifat *open source* [13]. Didalam *RapidMiner Studio* memiliki banyak sekali algoritma yang telah disediakan termasuk pada penelitian ini yaitu algoritma C4.5, *Naïve Bayes*, *Support Vector Machine* (SVM) dan *K-Nearest Neighbour* (KNN) [14]. Setelah penulis mengumpulkan data studi literatur sejenis, maka penulis mengambil objek yaitu perbandingan algoritma C4.5, *Naïve Bayes*, *Support Vector Machine* (SVM) dan *K-Nearest Neighbour* (KNN) pada klasifikasi menentukan

kelayakan penerima bantuan. Untuk mengetahui seberapa efektifnya empat algoritma tersebut maka terdapat empat parameter penentu yaitu nilai *Accuracy*, *Weighted Mean Recall*, *Weighted Mean Precision* dan *Execution Time*.

Nilai *accuracy* merupakan tingkat kedekatan antara nilai prediksi dengan nilai aktual, *Weighted Mean Recall* yaitu rata-rata dari nilai *Recall* disetiap kelas yang dihitung dengan mempertimbangkan distribusi kelas, *Weighted Mean Precision* yaitu rata-rata dari nilai *Precision* di setiap kelas yang dihitung dengan mempertimbangkan distribusi kelas, sedangkan *Execution Time* yaitu banyaknya waktu yang dibutuhkan untuk memproses data. Nilai *accuracy* ini dihasilkan berdasarkan *data testing* dan *data testing* didapat dari *split data*. *Data Testing* tersebut dapat diatur banyaknya dengan *split ratio* [15]. *Split ratio* merupakan pembagian data yang menjadi banyaknya *data training* dan *data testing*. Pada penelitian kali ini sudah ditentukan besarnya *split ratio* adalah 0,5. Ini berarti 50% yang digunakan sebagai *data training* dan 50% digunakan sebagai *data testing*. Persamaan 1 adalah rumus untuk menghitung nilai *accuracy*.

$$Accuracy = \frac{\sum Data Uji Benar}{\sum Total Data Uji} \times 100\% \quad (1)$$

Dalam penelitian ini data yang digunakan adalah data yang didapatkan dari kecamatan langsung dapat dilihat pada Tabel 1.

Tabel 1. Kriteria Calon Penerima Bantuan

No	Kriteria	Deskripsi
1	Status Pekerjaan	Status pekerjaan calon bantuan iuran sesuai yang tertera pada KTP
2	Kondisi Rumah	Kondisi tempat tinggal calon bantuan iuran yang dibagi menjadi 3 poin. Biasa, Tidak Layak, dan Sangat Tidak Layak

Setelah mendapatkan kriteria-kriteria tersebut, maka peneliti sebelumnya membuat tabel aspek penelitian kriteria tersebut seperti pada Tabel 2.

Tabel 2. Aspek Penilaian Kriteria

Kriteria	Range Nilai	Label
Status Pekerjaan	Buruh Harian Lepas	Buruh Harian Lepas
	Pegawai Swasta	Pegawai Swasta
	40% - 50%	Biasa
Kondisi Rumah	50% - 60%	Tidak Layak
	60% - 100%	Sangat Tidak Layak

3.2. Tahapan Olah Data.

Berikut ini adalah tahapan metode C4.5, Naïve Bayes, KNN dan SVM dari awal pengolahan data sampai ke *RapidMiner Studio* [16].

Setelah mencari referensi-referensi yang bersumber dari internet, dilakukanlah wawancara dengan salah satu pegawai PUSKESOS yaitu Ibu Karnita yang

kemudian mendapatkan data dari PBI APBD tahun 2022-2023. Data tersebut diolah yang sebelumnya data mentah seperti terlihat pada Gambar 2. Berdasarkan wawancara dengan Ibu Karnita bahwa faktor yang paling mempengaruhi terima atau tidaknya bantuan

dilihat dari dua faktor, yaitu status pekerjaan dan kondisi rumah calon penerima bantuan. Dengan bantuan Bu Karnita data mentah sebelumnya kemudian diolah seperti Gambar 3.

No	Tahun	Bulan	Kecamatan	Desa	Aksi	Sumber Data	Status Data	KTP	KK	Foto Rumah	Geotagging	No KK	NIK	Nama Lengkap	Alamat Lengkap	Status Hubungan Keluarga	Program Bansos
1	2023	5	PAMEUNGPEUK	LANGONSARI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20414E+15	AZKAAL SYAHAN TRIAN PRATAMA	KP.LANGONSARI INDAH RT 4 / RW 4	ANAK	PBI
2	2023	5	PAMEUNGPEUK	LANGONSARI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20415E+15	NADIA RIZKI AMALIA	KP.BOJONGPEUTEUY RT 3 / RW 5	ISTRI	PBI
3	2023	5	PAMEUNGPEUK	LANGONSARI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20414E+15	RAHMAT MAULANA	KP.BOJONGPEUTEUY RT 3 / RW 5	KEPALA KELUARGA	PBI
4	2023	5	PAMEUNGPEUK	LANGONSARI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20414E+15	MUHAMMAD ALBAR NA SR	KP.CIBUK RT 3 / RW 13	ANAK	PBI
5	2023	5	PAMEUNGPEUK	LANGONSARI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20416E+15	RKA SITI RANDINI	KP.CIBUK RT 3 / RW 13	ISTRI	PBI
6	2023	5	PAMEUNGPEUK	LANGONSARI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20415E+15	RYOM MINTARSH	KP.CINAGREG RT 3 / RW 15	KEPALA KELUARGA	PBI
7	2023	4	PAMEUNGPEUK	RANCA MULYA		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20415E+15	WARISTA	KP.MENGGER RT 01 / RW 07	KEPALA KELUARGA	PBI
8	2023	4	PAMEUNGPEUK	LANGONSARI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20415E+15	SELVI ANANDA PUTRI	KP.LEGOK LOA RT 3 / RW 12	ANAK	PBI
9	2023	4	PAMEUNGPEUK	LANGONSARI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20414E+15	IRMA EVIANA	KP.LEGOK LOA RT 3 / RW 12	ISTRI	PBI
10	2023	3	PAMEUNGPEUK	BOJONGKUNCI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20414E+15	ALFAN NAHIF JH	KP.SUKAMANA RT 02 / RW 03	ANAK	PBI
11	2023	3	PAMEUNGPEUK	BOJONGKUNCI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20414E+15	FANNYA ALLIAU NAZIWA	KP.SUKAMANA RT 02 / RW 03	ISTRI	PBI
12	2023	3	PAMEUNGPEUK	BOJONGKUNCI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20416E+15	ALDI MUHAMMAD REFALDI	KP.SUKAMANA RT 02 / RW 03	KEPALA KELUARGA	PBI
13	2023	3	PAMEUNGPEUK	BOJONGKUNCI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20415E+15	SUNUNG RAT	KP.LAMPEGAN RT 02 / RW 04	ISTRI	PBI
14	2023	3	PAMEUNGPEUK	BOJONGKUNCI		LAYANAN PBI APBD	DRAFT USULAN	Ganti	Ganti	Ganti		3.20414E+15	3.20414E+15	MULYA BNI SUKATMA	KP.LAMPEGAN RT 002 / RW 04	KEPALA KELUARGA	PBI

Gambar 2. Data Mentah PBI APBD

Tabel 3. Data setelah diolah manual

No	Status Pekerjaan	Kondisi Rumah	Status
1	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
2	Buruh Harian Lepas	Tidak Layak	Diterima
3	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
4	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
5	Buruh Harian Lepas	Biasa	Diterima
6	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
7	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
8	Buruh Harian Lepas	Tidak Layak	Diterima
9	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
10	Pegawai Swasta	Tidak Layak	Ditolak
11	Pegawai Swasta	Biasa	Ditolak
12	Buruh Harian Lepas	Tidak Layak	Diterima
13	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
14	Buruh Harian Lepas	Sangat Tidak Layak	Diterima

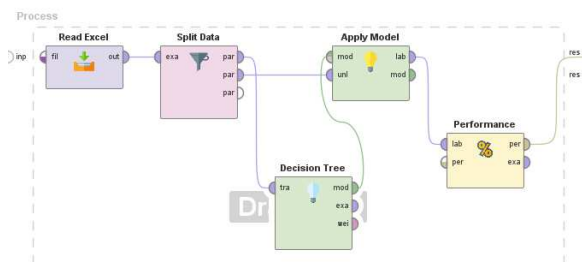
Sesudah data diolah, yang sebelumnya kondisi rumah dalam bentuk persen, pada Tabel 4 diubah menjadi 3 poin yaitu Baik, Tidak Layak dan Tidak Layak.

Kemudian data tersebut dibagi menjadi 2 bagian data *training* dan data *testing* dengan perbandingan 50:50 menggunakan *operators read excel*, *split data*, model algoritma (C4.5, Naïve Bayes, KNN), *apply model* dan *performance* dapat dilihat pada Gambar 3.

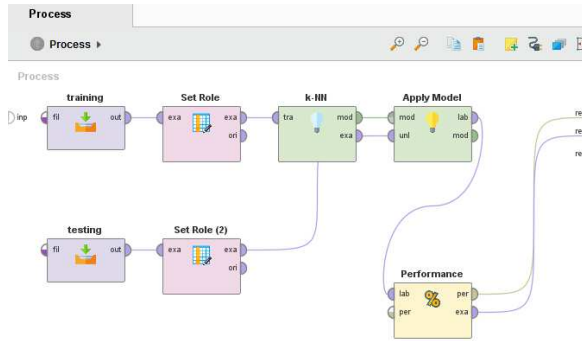
Tabel 4. Data Setelah Diolah Manual

No	Status Pekerjaan	Kondisi Rumah	Status
1	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
2	Buruh Harian Lepas	Tidak Layak	Diterima
3	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
4	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
5	Buruh Harian Lepas	Biasa	Diterima
6	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
7	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
8	Buruh Harian Lepas	Tidak Layak	Diterima
9	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
10	Pegawai Swasta	Tidak Layak	Ditolak
11	Pegawai Swasta	Biasa	Ditolak
12	Buruh Harian Lepas	Tidak Layak	Diterima
13	Buruh Harian Lepas	Sangat Tidak Layak	Diterima
14	Buruh Harian Lepas	Sangat Tidak Layak	Diterima

Setelah pembagian data menggunakan *split ratio* 0.5, kemudian data tersebut dibagi menjadi *data training* dan *data testing*. Lalu masukkan *data training* dan *data testing* ke dalam *RapidMiner*, *set role*, model algoritma (C4.5, Naïve Bayes dan KNN), *apply model* dan *performance* seperti terlihat pada Gambar 4.

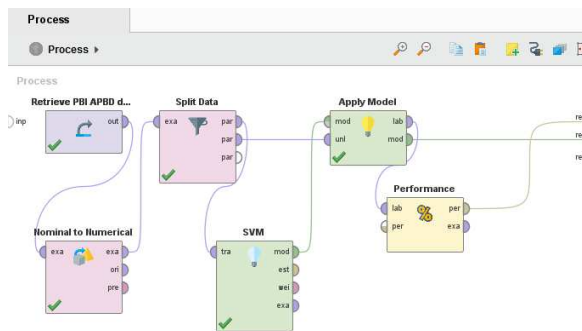


Gambar 1. Split data



Gambar 4. Operators untuk C4.5, Naive Bayes dan KNN

Untuk model algoritma SVM beberapa operator sedikit berbeda dari ketiga algoritma sebelumnya. Pada Gambar 5 dapat dilihat operator dari SVM terdiri dari *Retrieve*, *Nominal to Numerical*, *Split Data*, *SVM*, *Apply Model* dan *Performance*



Gambar 2. Operators untuk SVM

Pada bagian *result* pada Gambar 6 dapat dilihat terdiri dari *Example Set (Apply Model)* dan *PerformanceVector*.

Gambar 3. Proses Akhir Implementasi Menggunakan RapidMinerStudio

3.3. Analisis Keluaran.

Pada analisis keluaran ini menghasilkan 4 metode algoritma yang menampilkan *accuracy*, *weighted mean recall*, *weighted mean precision* dan *execution time*.

	true Ditolak	true Diterima	class precision
pred. Ditolak	6	1	85.71%
pred. Diterima	0	58	100.00%
class recall	100.00%	98.31%	

Gambar 4 Nilai Accuracy Algoritma C4.5

Dari Gambar 7 dihasilkan nilai *Accuracy* sebesar 98.46% dengan menggunakan *split ratio* 0.5.

	true Ditolak	true Diterima	class precision
pred. Ditolak	6	1	85.71%
pred. Diterima	0	58	100.00%
class recall	100.00%	98.31%	

Gambar 5. Nilai Weighted Mean Recall Algoritma C4.5

Dari Gambar 8 dihasilkan nilai *Weighted Mean Recall* sebesar 99.15% dengan menggunakan *split ratio* 0.5.

	true Ditolak	true Diterima	class precision
pred. Ditolak	6	1	85.71%
pred. Diterima	0	58	100.00%
class recall	100.00%	98.31%	

Gambar 6. Nilai Weighted Mean Precision Algoritma C4.5

Dari Gambar 9 dihasilkan nilai *Weighted Mean Precision* sebesar 92.86% dengan menggunakan *split ratio* 0.5.

	true Ditolak	true Diterima	class precision
pred. Ditolak	6	1	85.71%
pred. Diterima	0	58	100.00%
class recall	100.00%	98.31%	

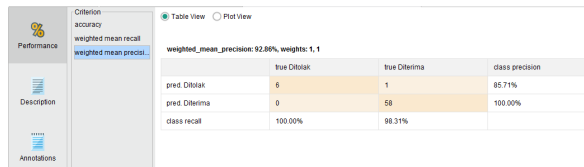
Gambar 7. Nilai Accuracy Naive Bayes

Dari Gambar 10 dihasilkan nilai *Accuracy* sebesar 98.46% dengan menggunakan *split ratio* 0.5.

	true Ditolak	true Diterima	class precision
pred. Ditolak	6	1	85.71%
pred. Diterima	0	58	100.00%
class recall	100.00%	98.31%	

Gambar 8. Nilai Weighted Mean Recall Naive Bayes

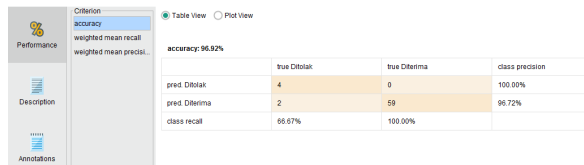
Dari Gambar 11 dihasilkan nilai *Weighted Mean Recall* sebesar 99.15% dengan menggunakan *split ratio* 0.5.



	true Ditolak	true Diterima	class precision
pred Ditolak	6	1	85.71%
pred Diterima	0	59	100.00%
class recall	100.00%	98.33%	

Gambar 9. Nilai Weighted Mean Precision Naive Bayes

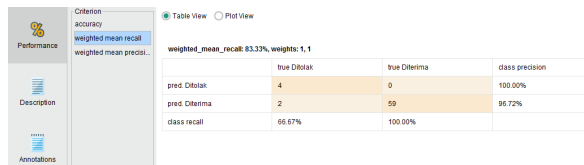
Dari Gambar 12 dihasilkan nilai *Weighted Mean Precision* sebesar 92.86% dengan menggunakan *split ratio* 0.5.



	true Ditolak	true Diterima	class precision
pred Ditolak	4	0	100.00%
pred Diterima	2	59	96.72%
class recall	66.67%	100.00%	

Gambar 10. Nilai Accuracy K-Nearest Neighbor

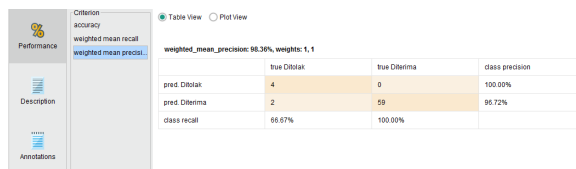
Dari Gambar 13 dihasilkan nilai *Accuracy* sebesar 96.92% dengan menggunakan *split ratio* 0.5.



	true Ditolak	true Diterima	class precision
pred Ditolak	4	0	100.00%
pred Diterima	2	59	96.72%
class recall	66.67%	100.00%	

Gambar 11. Nilai Weighted Mean Recall K-Nearest Neighbor

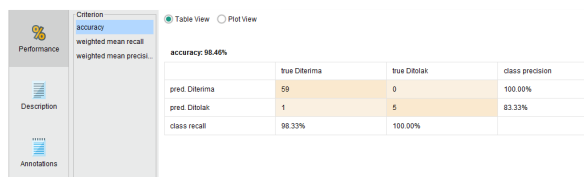
Dari Gambar 14 dihasilkan nilai *Weighted Mean Recall* sebesar 83.33% dengan menggunakan *split ratio* 0.5.



	true Ditolak	true Diterima	class precision
pred Ditolak	4	0	100.00%
pred Diterima	2	59	96.72%
class recall	66.67%	100.00%	

Gambar 12. Nilai Weighted Mean Precision K-Nearest Neighbor

Dari Gambar 15 dihasilkan nilai *Weighted Mean Precision* sebesar 98.36% dengan menggunakan *split ratio* 0.5.



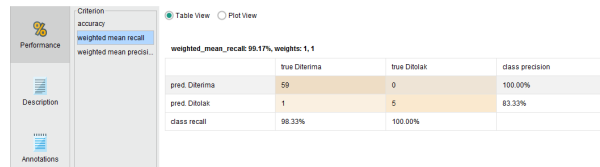
	true Ditolak	true Diterima	class precision
pred Ditolak	59	0	100.00%
pred Diterima	1	5	83.33%
class recall	98.33%	100.00%	

Gambar 13. Nilai Accuracy SVM

Dari Gambar 16 dihasilkan nilai *Accuracy* sebesar 98.46% dengan menggunakan *split ratio* 0.5.

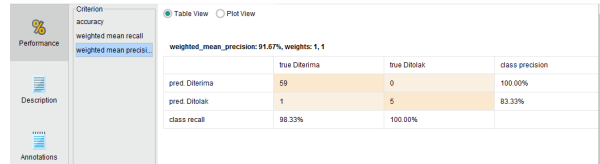
Dari Gambar 17 dihasilkan nilai *Weighted Mean Recall* sebesar 99.17% dengan menggunakan *split ratio* 0.5.

Dari gambar 18 dihasilkan nilai *Weighted Mean Precision* sebesar 91.67% dengan menggunakan *split ratio* 0.5.



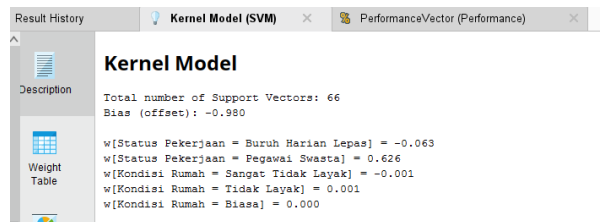
	true Ditolak	true Diterima	class precision
pred Ditolak	59	0	100.00%
pred Diterima	1	5	83.33%
class recall	98.33%	100.00%	

Gambar 14. Nilai Weighted Mean Recall SVM



	true Ditolak	true Diterima	class precision
pred Ditolak	59	0	100.00%
pred Diterima	1	5	83.33%
class recall	98.33%	100.00%	

Gambar 15. Nilai Weighted Mean Precision SVM



	Weight
w[Status Pekerjaan = Buruh Harian Lepas]	-0.063
w[Status Pekerjaan = Pegawai Swasta]	0.626
w[Kondisi Rumah = Sangat Tidak Layak]	-0.001
w[Kondisi Rumah = Tidak Layak]	0.001
w[Kondisi Rumah = Biasa]	0.000

Gambar 19. Kernel Model SVM

Dari Gambar 19 dihasilkan *kernel model* SVM dimana *Total number of support Vectors* yaitu 66 dan *Bias (offset)* : -0.980

3.4 Hasil Perbandingan.

Setelah melakukan semua tahapan pada metode simulasi, maka penulis akan membandingkan hasil dari output analysis dalam bentuk tabel dan grafik. Adapun yang akan dibandingkan adalah nilai accuracy, Weighted Mean Recall, Weighted Mean Precision dan execution time dengan split ratio 0.5. Perbandingan dapat dilihat pada Tabel 5.

Tabel 5. Hasil Penelitian Akhir

Methode	Accuracy	Weighted Mean Recall	Weighted Mean Precision	Execution Time
Algoritma C4.5	98.46%	99.15%	92.86%	0 s
Algoritma Naïve Bayes	98.46%	99.15%	92.86%	0 s
K-Nearest Neighbor	96.92%	83.33%	98.36%	0 s
Support Vector Machine	98.46%	99.17%	91.67%	0 s

Berdasarkan Tabel 5 dihasilkan nilai *Accuracy*, *Weighted Mean Recall*, *Weighted Mean Precision* dan *Execution Time* pada data yang telah diteliti. *Accuracy* merupakan tingkat kedekatan antara nilai prediksi dengan nilai aktual, *Weighted Mean Recall* merupakan rata-rata dari *recall* di setiap atribut, *Weighted Mean Precision* merupakan rata-rata nilai *precision* di setiap kelas sedangkan *Execution Time* yaitu banyaknya waktu yang dibutuhkan untuk memproses data. Pada Algoritma C4.5 dihasilkan nilai *accuracy* sebesar

98,46%, pada Algoritma *Naïve Bayes* dihasilkan nilai *accuracy* sebesar 98,46%, KNN menghasilkan nilai *accuracy* sebesar 95,38%, sedangkan SVM menghasilkan nilai *accuracy* sebesar 98,46% kemudian pada Algoritma C4.5 dihasilkan nilai *Weighted Mean Recall* sebesar 99.15%, *Naïve Bayes* dihasilkan nilai *Weighted Mean Recall* sebesar 99.15%, KNN dihasilkan nilai *Weighted Mean Recall* sebesar 83.33%, SVM dihasilkan nilai *Weighted Mean Recall* sebesar 99.17%, kemudian pada Algoritma C4.5 dihasilkan nilai *Weighted Mean Precision* sebesar 92.86%, *Naïve Bayes* dihasilkan nilai *Weighted Mean Precision* sebesar 92.86%, KNN dihasilkan nilai *Weighted Mean Precision* sebesar 98.36%, SVM dihasilkan nilai *Weighted Mean Precision* sebesar 91.67%, dan untuk *execution time* keempat algoritma tersebut sama-sama mempunyai nilai 0 s.

4. Kesimpulan

Hasil perbandingan Algoritma C4.5, Naive Bayes, KNN dan SVM pada prediksi bantuan dengan menggunakan metode simulasi yang terdiri dari enam tahapan yaitu problem formulation, conceptual model, input and output data, simulation, experimentation, dan output analysis menunjukkan bahwa Algoritma C4.5, Naïve Bayes dan SVM mempunyai nilai *accuracy* yang tinggi dibandingkan dengan KNN. Hal ini dibuktikan dengan ketiga algoritma ini memiliki nilai *accuracy* sebesar 98,46% sedangkan algoritma KNN mempunyai nilai *accuracy* 95,38%. Sedangkan untuk nilai *Weighted Mean Recall* SVM memiliki nilai paling tinggi. Hal ini dibuktikan bahwa SVM memiliki nilai *Weighted Mean Recall* sebesar 99.17%, Algoritma C4.5 dan Naïve Bayes memiliki nilai *Weighted Mean Recall* sebesar 99.15%, dan KNN memiliki nilai *Weighted Mean Recall* paling rendah yaitu sebesar 83.33%. Dan untuk *Weighted Mean Precision* KNN memiliki nilai paling tinggi diantara keempat algoritma tersebut. Hal ini dibuktikan bahwa KNN memiliki nilai *Weighted Mean Precision* sebesar 98.36%, Algoritma C4.5 dan Naïve Bayes memiliki nilai *Weighted Mean Precision* sebesar 92.86%, dan SVM memiliki nilai *Weighted Mean Precision* sebesar 91.67%. Jika memprioritaskan akurasi, maka C4.5, Naïve Bayes dan SVM yang terbaik. Jika untuk memastikan bahwa sebanyak mungkin orang yang berhak mendapatkan bantuan benar-benar menerima bantuan, maka KNN yang lebih baik karena memiliki nilai *Weighted Mean Recall* tertinggi. Dan untuk memprediksi seseorang layak atau tidaknya mendapatkan bantuan, maka SVM pilihan yang terbaik karena memiliki nilai *Weighted Mean Precision* tertinggi. Kedepannya, penelitian ini dapat disempurnakan dengan menggunakan model *stacing ensemble*, Dimana menggabungkan beberapa model machine learning untuk satu luaran.

Daftar Rujukan

- [1] R. Ramadani and E. Revida, "Pemberdayaan Masyarakat Melalui Program Kelompok Usaha Bersama (KUBE) Untuk Menanggulangi Kemiskinan Di Kelurahan Bandar Utama Kota Tebing Tinggi".
- [2] A. Yuditia, Y. Hidayat, and S. Achmad, "PELAKSANAAN JAMINAN KESEHATAN NASIONAL OLEH BPJS BERDASARKAN UNDANG-UNDANG NO.40 TAHUN 2004 TENTANG SISTEM JAMINAN SOSIAL NASIONAL," *J. Magister Ilmu Huk.*, vol. 6, no. 1, p. 43, Aug. 2021, doi: 10.36722/jmih.v6i1.796.
- [3] M. R. Iswardhana and A. M. S. Attamimi, "EFEKTIVITAS PENERAPAN KEBIJAKAN OTONOMI DAERAH DALAM PENURUNAN TINGKAT KEMISKINAN DI BANDUNG BARAT TAHUN 2019," no. 2, 2023.
- [4] U. E. Muhamad Khandava Mulyadien, "Algoritma K-Means Untuk Pengelompokan Bantuan Langsung Tunai (BLT)," Jul. 2022, doi: 10.5281/ZENODO.6944517.
- [5] A. Yonathan, H. Sujaini, and E. E. Pratama, "Perbandingan Algoritma Klasifikasi dalam Pendeteksian Hoax pada Media Sosial," vol. 01, no. 1, 2022.
- [6] L. S. Hasibuan, "ANALISIS PENGARUH IPM, INFLASI, PERTUMBUHAN EKONOMI TERHADAP PENGANGGURAN DAN KEMISKINAN DI INDONESIA," vol. 8, 2023.
- [7] N. Alamsyah, Budiman, and Titan Parama, "Analysis of E-learning user Acceptance using the Technology Acceptance Model (TAM) and end-User Computing Satisfaction (EUCS)," *Formosa J. Appl. Sci.*, vol. 2, no. 8, pp. 1873–1892, Aug. 2023, doi: 10.55927/fjas.v2i8.5405.
- [8] N. Alamsyah and R. A. Krisdiawan, "PEMBANGUNAN APLIKASI SEBAGAI MEDIA PEMBELAJARAN BANGUN RUANG TINGKAT SD/SMP DENGAN MENGGUNAKAN METODE MARKER AUGMENTED REALITY," *NUANSA Inform.*, vol. 15, no. 1, p. 23, Jan. 2021, doi: 10.25134/nuansa.v15i1.3847.
- [9] R. W. Abdullah, D. Hartanti, H. Permatasari, A. W. Septyanto, and Y. A. Bagaskara, "Penerapan Data Mining untuk Memprediksi Jumlah Produk Terlaris Menggunakan Algoritma Naive Bayes Studi Kasus (Toko Prapti)," *J. Ilm. Inform. Glob.*, vol. 13, no. 1, Mar. 2022, doi: 10.36982/jiig.v13i1.2060.
- [10] N. Alamsyah and N. Safitri, "Sistem Pendeteksi Kebocoran Tabung Gas Elpiji (Lpg) Berbasis Nodemcu Dan Telegram," vol. 17, 2023.
- [11] A. C. Khotimah and E. Utami, "COMPARISON NAÏVE BAYES CLASSIFIER, K-NEAREST NEIGHBOR AND SUPPORT VECTOR MACHINE IN THE CLASSIFICATION OF INDIVIDUAL ON TWITTER ACCOUNT".
- [12] A. Simangunsong, "Penerapan Metode Monte Carlo Dalam Simulasi Pengelolaan Persediaan Alat Tulis Kantor," *J. SAINTIKOM J. Sains Manaj. Inform. Dan Komput.*, vol. 22, no. 2, p. 280, Aug. 2023, doi: 10.53513/jis.v22i2.8718.
- [13] E. Junaedi, A. M. Siregar, and E. Nurlaelasari, "Implementasi C4.5 Dan Algoritma K Nearest Neighbor Untuk Prediksi Kelayakan Pemberian Kredit Menggunakan RapidMiner Studio," no. 1, 2022.
- [14] A. G. Putrada, N. Alamsyah, S. F. Pane, and M. Nurkamal Fauzan, "Feature Importance on Text Analysis for a Novel Indonesian Movie Recommender System," in *2023 11th International Conference on Information and Communication Technology (ICoICT)*, 2023, pp. 34–39. doi: 10.1109/ICoICT58202.2023.10262504.
- [15] N. Alamsyah, Saparudin, and A. P. Kurniati, "A Novel Airfare Dataset To Predict Travel Agent Profits Based On Dynamic Pricing," in *2023 11th International Conference on Information and Communication Technology (ICoICT)*, 2023, pp. 575–581. doi: 10.1109/ICoICT58202.2023.10262694.
- [16] I. K. Sutarga, "Analisis Pola Spasial Sebaran COVID-19 Kota Bogor Berdasarkan Indek Moran," *Media Komun. Geogr.*, vol. 23, no. 2, pp. 265–276, Dec. 2022, doi: 10.23887/mkg.v23i2.55183.