

Klasifikasi Persediaan Barang Menggunakan Naïve Bayes

Ningrum Pramesti¹

¹Program Studi Komputerisasi Akuntansi, STMIK IKMI Cirebon

ningrumpraesti9@gmail.com*

Abstract

At this time, every organization or company needs a reliable information system to be able to produce information resources that are accurate, relevant, timely and up to date. to parties inside and outside the company. In inventory, there are still problems encountered at the Office of the Ministry of Public Works and Public Housing, Water Resources, Cimanuk Cisanggarung, especially in the inventory of goods, there are frequent errors in the calculation of inventory, stock records for incoming and outgoing goods have not been managed regularly so that there are still data discrepancies. there is. Naive Bayes is a statistical classification that can be used to predict the probability of membership of a class. The algorithm uses Bayes theorem and assumes all attributes are independent or not interdependent given by the value on the variable. Data taken from the purchase transaction for office stationery supplies at the Bbws Cimanuk Cimanuk Cisanggarung Office using the Naïve Bayes model yields an accuracy value of 80.00% with the following details: Prediction Ready and True Available at 69, Predictive available and True less than available at 7, Prediction less available and it turns out to be available by 13, Prediction Less available and Turns to True less than available by 10. With a proper class recall of 84.15% and an Infeasible class Recall of 58.82%.

Keyword: naïve bayes, akurasi, rapidminer

Abstrak

Pada masa ini, setiap organisasi atau perusahaan membutuhkan sistem informasi yang handal untuk dapat menghasilkan sumber daya informasi yang akurat, relevan, tepat waktu dan *up to date*. kepada pihak di dalam maupun di luar perusahaan. Dalam persediaan masih ada masalah yang dihadapi pada Kantor Kementerian Pekerjaan Umum dan Perumahan Rakyat Sumber Daya Air Cimanuk Cisanggarung khususnya dalam persediaan barang adalah sering terjadi kesalahan dalam perhitungan persediaan barang, pencatatan stok untuk barang masuk dan keluar belum dikelola secara teratur sehingga masih terjadi ketidaksesuaian data yang ada. *Naïve Bayes* Adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu *class*, Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel. Data yang diambil dari Transaksi pembelian persediaan barang Alat tulis kantor pada Kantor Bbws Cimanuk Cisanggarung menggunakan model Naïve bayes menghasilkan nilai akurasi sebesar 80.00% dengan rincian sebagai berikut: Prediksi Sedia dan True Sedia Sebesar 69, Prediksi sedia dan True kurang sedia sebesar 7, Prediksi kurang sedia dan Ternyata sedia Sebesar 13, Prediksi Kurang sedia dan Ternyata True kurang sedia Sebesar 10. Dengan class recall layak sebesar 84.15% dan Class Recall Tidak layak Sebesar 58.82%.

Kata Kunci: Naïve bayes, Akurasi Rapidminer

1. Pendahuluan

Saat ini merupakan era globalisasi, dimana penggunaan teknologi informasi sistem dan sistem informasi akuntansi yang baik sangat dibutuhkan. Pada masa ini, setiap organisasi atau perusahaan membutuhkan sistem informasi yang handal untuk dapat menghasilkan sumber daya informasi yang akurat, relevan, tepat waktu dan *up to date* kepada pihak di dalam maupun di luar perusahaan. Perkembangan teknologi informasi telah membawa dampak dalam kehidupan masyarakat. [1]

Persediaan adalah barang persediaan milik perusahaan yang ada di gudang, dalam perjalanan, maupun yang dititipkan pada pihak yang bebas dipergunakan dalam kegiatan operasi perusahaan. Dalam kegiatan perusahaan yang berhubungan dengan pengolahan, persediaan, penyimpanan data, sampai pembuatan laporan mengenai akuntansi biaya persediaan. [2] metode FIFO (*First in, First out*), FIFO adalah asumsi bahwa perhitungan harga pokok didasarkan atas urutan pembelian barang tersebut. Dengan demikian, persediaan yang tertinggal atau persediaan sisa dianggap berasal dari waktu pembelian terakhir. [3]

Dalam persediaan masih ada masalah yang dihadapi pada Kantor Kementerian Pekerjaan Umum dan Perumahan Rakyat Sumber Daya Air Cimanuk Cisanggarung khususnya dalam persediaan barang adalah sering terjadi kesalahan dalam perhitungan persediaan barang, pencatatan stok untuk barang masuk dan keluar belum dikelola secara teratur sehingga masih terjadi ketidaksesuaian data yang ada. Oleh karena itu berpengaruh dalam pembuatan laporan stok barang, sehingga membutuhkan waktu yang lama. Persediaan barang mempunyai beragam jenis barang. Karena kebutuhan perusahaan banyak maka perusahaan memiliki persediaan jenis barang alat tulis kantor. Penelitian yang dilakukan berdasarkan pokok permasalahan, yaitu mengetahui bagaimana sistem informasi akuntansi persediaan barang menggunakan metode Naïve Bayes pada Kantor Kementerian Pekerjaan Umum dan Perumahan Rakyat Direktorat Jenderal Sumber Daya Air Balai Besar Wilayah Sungai Cimanuk-Cisanggarung Satuan Kerja Operasi Dan Pemeliharaan Sumber Daya Air Cimanuk-Cisanggarung.

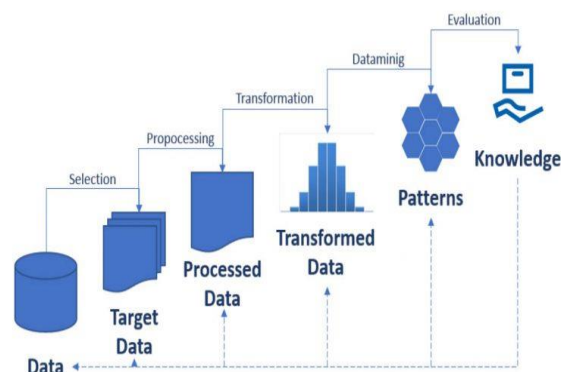
Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas. [4] Klasifikasi dapat didefinisikan sebagai proses untuk menyatakan suatu objek data sebagai salah satu kategori (kelas) yang telah didefinisikan sebelumnya. [5]

Tujuan penelitian ini berdasarkan rumusan masalah pada penelitian ini sehingga tujuan penelitian sebagai berikut :

1. Untuk Mengetahui pengelolaan prediksi persediaan barang menggunakan naïve bayes yang dapat digunakan supaya menjadi sistem persediaan barang pada Kantor Bbws Cimanuk Cisanggarung?
2. Untuk mengetahui data persediaan barang apakah terjadi aman atau tidak aman pada kantor Bbws Cimanuk Cisanggarung.
3. Untuk mengetahui berapa hasil tingkat akurasi dalam memprediksi data persediaan barang pada kantor Bbws Cimanuk Cisanggarung menggunakan naïve bayes.

2. Metode Penelitian

Alur penelitian ini mengacu pada Knowledge Discovery Data, alur tersebut dapat dilihat pada gambar berikut ini :



Gambar 1 Alur Penelitian

Berdasarkan Gambar 1 tentang Alur Penelitian dapat dijelaskan sebagai berikut :

1. Data Selection

Data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *data mining* disimpan dalam suatu berkas, terpisah dari basis data operasional. Data Selection proses seleksi data yang relevan terhadap analisa untuk diterima dari koleksi data yang ada. Karena tidak semua data yang didapat dari sumber objek penelitian akan digunakan. Data yang digunakan akan diseleksi dengan cara melihat kecenderungan data / kesesuaian data dengan topik / judul penelitian yang akan diteliti oleh penulis.

2. Pre-processing / Cleaning

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses pembersihan pada data yang menjadi fokus KDD. Proses pembersihan mencakup antara lain membuang *duplikasi data*, memeriksa data yang *inkonsisten*, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (*tipografi*). Data yang tidak konsisten atau kosong dan juga memperbaiki data yang rusak. Proses pembersihan data dilakukan untuk memastikan bahwa data yang telah dipilih layak untuk diolah berdasarkan proses *data mining*.

3. Transformation

Coding adalah transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data. Karena dalam penelitian ini akan dilakukan uji coba menggunakan *tools RapidMiner*, maka data yang telah melalui proses sebelumnya akan ditransformasi agar dapat sesuai dengan algoritma yang dipakai yaitu algoritma *naïve bayes*.

4. Data mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. Interpretation / Evaluation

Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.

6. Knowledge

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami data mining. Karenanya presentasi hasil data mining dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses data mining. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil data mining.

Literatur Review yang dilakukan oleh Ahmad turmudi dan Dede yusup menganalisis “Analisis Data Mining Menggunakan Algoritma Naïve Bayes Dalam Memprediksi Pembelian Material Plastik Injection” menyatakan Perusahaan Manufaktur adalah sebuah badan usaha yang mengoperasikan mesin, peralatan dan tenaga kerja dalam suatu medium proses untuk mengubah bahan- bahan mentah menjadi barang jadi yang memiliki nilai jual. Karena itulah dalam pembelian bahan baku plastik *injection* harus dilakukan secara obyektif dan efisien agar proses seleksi dapat bersaing dalam dunia indsutri yang diharapkan perusahaan. PT. Surya Teknologi Industri yang bergerak di bidang Manufacturing memiliki jenis bahan baku lebih dari 400 jenis sehingga perusahaan kesulitan untuk memprediksi kebutuhan yang akan datang dan dibutuhkan perusahaan[6]

Penelitian Deri lianda, Niko surya atmaja dalam jurnal pada naskah publikasih pada tahun 2021 yang berjudul “Prediksi data buku favorit Menggunakan metode naïve bayes” Mengatakan dalam data mining,

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam *database*, menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat bagi ilmu pengetahuan. *Data Mining* merupakan proses penggalian data yang sudah ada untuk mendapatkan informasi yang akan diolah kembali menjadi informasi yang baru. Klasifikasi *data mining* adalah metode pembelajaran data untuk meramalkan nilai dari beberapa atribut. Metode klasifikasi menghasilkan aturan-aturan yaitu *rule* untuk digunakan sebagai indikator dalam memprediksi sebuah tujuan.

Naïve Bayes merupakan salah satu algoritmayang terdapat pada teknik klasifikasi. Naïve Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan Naïve dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi Naïve Bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya.[7]

Knowledge discovery in database (KDD) adalah kegiatan yang melibatkan pengumpulan, pemakaian data, historis, untuk menemukan keteraturan, pola atau hubungan dalam set data yang berukuran besar. KDD adalah suatu proses pengganti data mentah menjadi lebih bermanfaat dan sebagai sumber pengetahuan.[8] RapidMiner merupakan perangkat lunak yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining text mining dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi. Tools yang digunakan dalam Rapidminer

1. Panel operator view

Panel operator view merupakan induk langkah kerja dari proses analisa.

2. Panel repository view

Merupakan komponen utama dalam Design Perspective selain Operator View. View ini dapat digunakan untuk mengelola dan menata proses Analisis menjadi proyek dan pada saat yang sama juga dapat digunakan sebagai sumber data dan yang berkaitan dengan meta data.

3. Process view

Menunjukkan langkah-langkah tertentu dalam proses analisis dan sebagai penghubung langkah-langkah tersebut.

4. Parameter view

Digunakan untuk mengatu fungsionalitas dari beberapa

operator yang membutuhkannya agar dapat dijalankan sesuai dengan proses analisa.

Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas. Definisi lain mengatakan Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. *Naive Bayes*, pengolahan data dan yang akan dijadikan dataset dalam penelitian ini. Dari data tersebut akan dibagi menjadi 80% *data training* dan 20% *data testing*. Memisahkan data menjadi training dan testingset dimaksudkan agar model yang diperoleh nantinya memiliki kemampuan generalisasi yang baik dalam melakukan klasifikasi data. Data training atau trainingset adalah bagian dataset yang dilatih untuk membuat klasifikasi atau menjalankan fungsi dari sebuah algoritma sesuai dengan tujuannya masing-masing. *Data testing* atau *test set* adalah bagian dataset yang digunakan untuk melihat keakuratan atau performa dalam hasil.[4]

Berikut ialah persamaan dari *teorema Bayes*:

$$P(H|X) = P(X|H) \cdot P(H)$$

$$P(X)$$

Keterangan :

X = Sampel data dengan class (label) yang belum diketahui

H = Hipotesis bahwa X adalah data suatu class (label) spesifik

$P(H|X)$ = Probabilitas Hipotesis H berdasar kondisi X (posteriori probabilitas)

$P(H)$ = Probabilitas hipotesis H (prior probabilitas)

$P(X|H)$ = Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$ = Probabilitas X

Berdasarkan uraian diatas, dapat dijelaskan bahwa $P(H|X)$ merupakan probabilitas dari H yang didalam X atau dalam bahasa lain bahwa $P(H|X)$ merupakan persentase dengan banyaknya H didalam X,

$P(X|H)$ merupakan probabilitas X didalam H, $P(H)$ merupakan probabilitas prior dari H dan $P(X)$ merupakan probabilitas prior dari X.

3. Hasil dan Pembahasan

Pada tahap ini menjelaskan tentang tahap awal yang dilakukan pada proses data minig. Data yang didapat dari BBWS Cimanuk-Cisanggarung akan diolah menjadi format yang dibutuhkan dan dilakukan penentuan atribut data yang akan digunakan. Berikut tahapan dalam pengolahan data awal :

a. Data Set Sebelum Preprocessing

Data Set Sebelum Preprocessing merupakan teknik awal data mining untuk mengubah data mentah

atau biasa dikenal dengan raw data yang dikumpulkan dari berbagai sumber menjadi informasi yang lebih bersih dan bisa digunakan untuk pengolahan selanjutnya. Berikut ini adalah data pembelian persediaan barang yang diambil pada BBWS Cimanuk-Cisanggarung yang kemudian akan dilakukan normalisasi.

b. Data Set Sesudah Preprocessing

Data Set Sesudah Preprocessing adalah pembersih data untuk membuang data yang missing value data tidak konsisten data yang rusak.

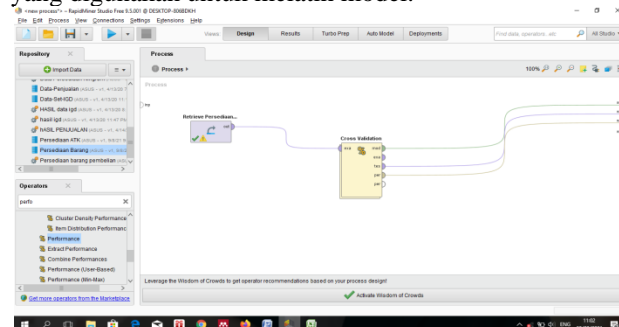
c. Model DataMining

Model Datamining dengan menggunakan algoritma Naive Bayes dalam menentukan Persediaan barang sebagai berikut:

a. Retrive Data Persediaan

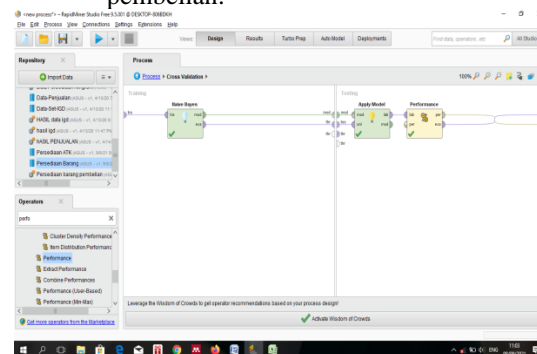
b. Cross Validation

Cross Validation adalah melakukan validasi berulang di mana dataset dibagi menjadi banyak subset (himpunan) data latih & validasi. Setiap iterasi memvalidasi (menguji) satu subset data dengan subset yang tersisa sebagai data latih. Pelatihan subproses yang digunakan untuk melatih model.



Gambar 2. Model Retrive Data dan Cross Validation

1. Pemodelan adalah tahapan (langkah) dalam membuat model dari suatu sistem nyata(realitas).
2. Rapidminer adalah sebuah lingkungan machine learning data mining.
3. Jumlah dataset mencapai 120 record maka penulis menggunakan tools Rapidminer untuk membantu proses perhitungan.
4. Validation untuk membantu menghasilkan tingkat keakurasian.....berdasarkan dataset pembelian.



Gambar 3 Proses Training & Testing

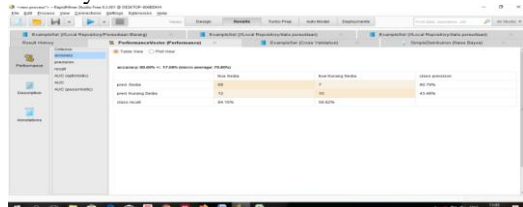
Berdasarkan semua gambar diatas menjelaskan bahwa dalam menerapkan model Naïve Bayes dibutuhkan aplikasi Rapidminer versi 9.3 dengan operator sebagai berikut Retrive, Cross Validation, Naïve Bayes, Apply Model dan Performance. Di dalam kolom trainingterdapat algoritma klasifikasi yang diterapkan yaitu Naive bayes. Sedangkan di dalam kolom testing terdapat Apply Model untuk menjalankan model naive bayes.

1. Hasil Performace Vector

Proses klasifikasi Rapidminer dengan metode Naïve bayes berdasarkan hasil Penerapan proses ini menggunakan Rapidminner maka menghasilkan akurasi Accuraey, Precision, Recall sabagai berikut:

a. Akurasi Accuraey

Akurasi Accuraey didefinisikan sebagai tingkat kedekatan antara nilai prediksi. Dengan jumlah data dapat diketahui nilai akurasi Accuraey hasil prediksi yaitu 80.00%

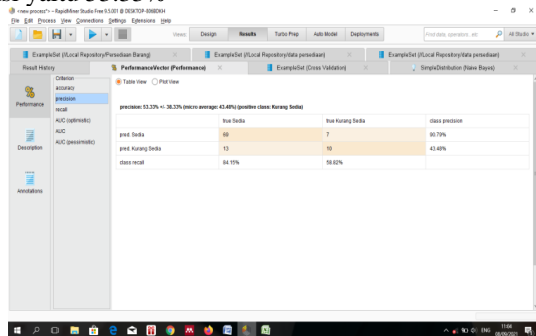


	True Sedia	True Kurang Sedia	Class Precision
pred. Sedia	69	7	80.79%
pred. Kurang Sedia	13	10	43.48%
class recall	84.15%	58.82%	

Gambar 4 .Hasil Akurasi Accuraey

b. Akurasi Precision

Akurasi Precision adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. Dengan jumlah data dapat diketahui nilai akurasi Precision hasil prediksi yaitu 53.33%

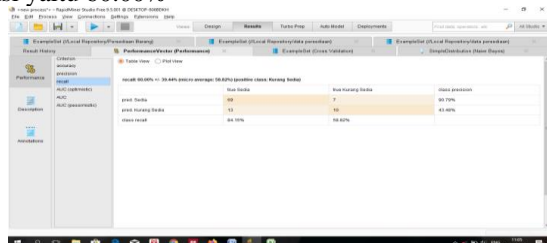


	True Sedia	True Kurang Sedia	Class Precision
pred. Sedia	69	7	80.79%
pred. Kurang Sedia	13	10	43.48%
class recall	84.15%	58.82%	

Gambar 5 Hasil Akurasi Precision

c. Akurasi Recall

Akurasi Recall adalah tingkat keberhasilan sistem dalam menemukan kembali informasi. Dengan jumlah data dapat diketahui nilai akurasi Recall hasil prediksi yaitu 60.00%



	True Sedia	True Kurang Sedia	Class Precision
pred. Sedia	69	7	80.79%
pred. Kurang Sedia	13	10	43.48%
class recall	84.15%	58.82%	

Gambar 6 Hasil Akurasi Recall

Berdasarkan gambar diatas menjelaskan nilai akurasi sebesar 80.00% dengan rincian sebagai berikut:Prediksi Sedia dan True Sedia Sebesar 69, Prediksi sedia dan True kurangsedia sebesar 7, Prediksi kurangsedia dan Ternyata sedia Sebesar 13, Prediksi Kurang sedia dan Ternyata True kurang sedia Sebesar 10. Dengan class recall layak sebesar 84.15% dan Class Recall Tidak layak Sebesar 58.82%.

4. Kesimpulan

Kesimpulan pada penelitian ini ialah berdasarkan dari data sudah di olah, menggunakan RapidMiner berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. Menjelaskan data yang diambil dari Transaksi pembelian persediaan barang Alat tulis kantor pada Kantor Bbws Cimanuk Cisanggarung menggunakan model Naïve bayes mengasilkans nilai akurasi sebesar 80.00% dengan rincian sebagai berikut:Prediksi Sedia dan True Sedia Sebesar 69, Prediksi sedia dan True kurangsedia sebesar 7, Prediksi kurang sedia dan Ternyata sedia Sebesar 13, Prediksi Kurang sedia dan Ternyata True kurang sedia Sebesar 10. Dengan class recall layak sebesar 84.15% dan Class Recall Tidak layak Sebesar 58.82%.

Daftar Rujukan

- [1] S. Maharsi, “Pengaruh perkembangan teknologi informasi terhadap bidang akuntansi manajemen,” pp. 127–137, 2018.
- [2] P. S. Informatika and H. S. Haryati, “SISTEM INVENTORY MENGGUNAKAN METODE FIFO,” 2020.
- [3] J. Teknologi and S. Informasi, “IMPLEMENTASI MODEL RAPID APPLICATION DEVELOPMENT PADA SISTEM INFORMASI PERSEDIAAN BARANG DENGAN METODE FIFO Iqbal Kamil Siregar Sistem Komputer , Sekolah Tinggi Manajemen Informatika dan Komputer Royal PENDAHULUAN Sistem informasi berlandaskan komputer me,” vol. 6, no. 2, pp. 187–192, 2020.
- [4] “Donn maulana,” vol. 10, pp. 191–197, 2019.
- [5] U. Optimasi and S. Pemasaran, “Jurnal Teknologi Informasi dan Komunikasi STMIK Subng, April 2018 ISSN: 2252-4517,” no. April, pp. 84–95, 2018.
- [6] “Ahmad Turmudi,” vol. 10, pp. 42–54, 2020.
- [7] Deri Lianda&Niko Atmaja, “PREDIKSI DATA BUKU FAVORIT,” vol. VIII, pp. 27–37, 2021.
- [8] F. Yunita, “Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru,” *Sistemasi*, vol. 7, no. 3, p. 238, 2018, doi: 10.32520/stmsi.v7i3.388.