

PREDIKSI HASIL PANEN PADI DI KABUPATEN JEMBRANA MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER

Ida Bagus Kade Dwi Suta Negara¹⁾, I Putu Kusuma Negara²⁾, Norsa Yudhi Arso³⁾

Program Studi Teknik Informatika K. Jembrana^{1) 2) 3)}

Universitas Triatma Mulya, Jembrana, Bali^{1) 2) 3)}

suta.negara@triatmamulya.ac.id

ABSTRACT

The agricultural sector plays an important role in the economy of Indonesia. One of the common challenges faced in the rice farming sector is the low quality of harvest. The low quality of rice harvest is caused by factors such as the use of poor-quality seeds, improper use of fertilizers and pesticides, as well as insufficient care and monitoring of rice plants. With the advancement of information technology, there are many solutions provided to farmers for predicting rice harvest results more accurately and effectively. One of the solutions that can be used is the Naive Bayes Classifier (NBC) method. With a rice harvest prediction system, farmers can make better decisions regarding the use of fertilizers and irrigation, as well as anticipate and overcome potential issues such as pest attacks and plant diseases. Furthermore, with a rice harvest prediction system, farmers can improve resource utilization efficiency and optimize their harvest results. The model used to conduct this research is CRISP-DM (Cross Industry Standard Process for Data Mining). Based on the research findings, several conclusions can be drawn, including: this research has successfully produced rice harvest predictions in Jembrana district using the Naive Bayes Classifier (NBC) method. The naive Bayes classifier method provides reasonably accurate calculation results in predicting rice harvest with an accuracy rate of 89% using 9 test data.

Keywords: prediction, rice harvest, nbc.

ABSTRAK

Sektor pertanian memegang peranan penting dalam perekonomian di Indonesia. Salah satu kendala yang sering dihadapi dalam sektor pertanian padi adalah rendahnya kualitas hasil panen. Kualitas panen padi yang rendah disebabkan oleh faktor-faktor seperti penggunaan bibit yang kurang berkualitas, penggunaan pupuk dan pestisida yang tidak tepat, serta kurangnya perawatan dan pengawasan terhadap tanaman padi. Dengan kemajuan teknologi informasi, terdapat banyak solusi yang diberikan bagi petani untuk memprediksi hasil panen padi dengan lebih akurat dan efektif. Salah satu solusi yang dapat digunakan adalah metode *Naive Bayes Classifier* (NBC). Dengan sistem prediksi panen padi, petani dapat mengambil keputusan yang lebih baik dalam hal penggunaan pupuk dan irigasi. serta mengantisipasi dan mengatasi permasalahan yang mungkin terjadi seperti serangan hama dan penyakit pada tanaman. Selain itu, dengan sistem prediksi panen padi, petani dapat meningkatkan efisiensi penggunaan sumber daya dan mengoptimalkan hasil panen mereka. Model yang digunakan untuk menyelesaikan penelitian ini adalah *CRISP-DM* (*Cross Industry Standard Process for Data Mining*). Berdasarkan hasil penelitian, dapat diambil beberapa kesimpulan di antaranya : penelitian ini telah berhasil menghasilkan prediksi panen padi di kabupaten Jembrana menggunakan metode Naive Bayes Classifier (NBC). Metode naive bayes classifier memberikan hasil perhitungan yang cukup akurat dalam memprediksi hasil panen padi dengan tingkat akurasi sebesar 89% dengan data uji sebanyak 9 data.

Kata Kunci : prediksi, panen padi, nbc.

PENDAHULUAN

Sektor pertanian memegang peranan penting dalam perekonomian di Indonesia. Salah satu komoditas unggulan pertanian adalah padi yang merupakan tanaman pangan utama yang menjadi sumber penghasilan bagi petani dan masyarakat di Indonesia, termasuk di kabupaten Jembrana. Salah satu kendala yang sering dihadapi dalam sektor pertanian padi adalah rendahnya kualitas hasil panen. Rendahnya hasil panen padi dapat menurunkan nilai jual dan mengurangi keuntungan bagi petani. Selain itu, kualitas hasil panen yang rendah juga dapat berdampak negatif pada pangan dan gizi masyarakat yang bergantung pada kualitas produksi padi.

Kualitas panen padi yang rendah disebabkan oleh faktor-faktor seperti penggunaan bibit yang kurang berkualitas, penggunaan pupuk dan pestisida yang tidak tepat, serta kurangnya perawatan dan pengawasan terhadap tanaman padi. Selain itu, faktor lingkungan seperti curah hujan yang tidak teratur dan kondisi tanah yang tidak mendukung juga dapat berkontribusi terhadap rendahnya kualitas hasil panen. Oleh karena itu, perlu dilakukan upaya untuk meningkatkan kemampuan petani dalam memprediksi hasil panen serta pengelolaan pertanian yang lebih efektif dan berkelanjutan untuk meningkatkan kualitas hasil panen padi.

Dengan pesatnya kemajuan teknologi informasi, terdapat banyak solusi yang dapat diberikan bagi petani untuk memprediksi hasil panen padi dengan lebih akurat dan efektif. Salah satu solusi yang dapat digunakan adalah metode Naive Bayes, yang merupakan metode klasifikasi data yang berbasis pada teori probabilitas. Dengan menggunakan metode Naive Bayes, petani dapat mengolah data historis seperti kondisi tanah, musim, dan teknik budidaya yang digunakan, sehingga dapat memperoleh prediksi hasil panen padi yang lebih akurat dan dapat membantu petani dalam merencanakan dan mengambil keputusan yang lebih baik dalam mengelola produksi padi.

Naive Bayes merupakan suatu bentuk klasifikasi data dengan menggunakan metode probabilitas dan statistik. Bayes rule digunakan untuk menghitung probabilitas

suatu class. Algoritma Naive Bayes memberikan suatu cara mengkombinasikan peluang terdahulu dengan syarat kemungkinan menjadi sebuah formula yang dapat digunakan untuk menghitung peluang dari tiap kemungkinan yang terjadi. Naive Bayes biasanya digunakan untuk memprediksi data, terutama untuk melakukan klasifikasi data. Metode ini dapat memprediksi kategori atau label yang paling mungkin untuk suatu data berdasarkan pada kemunculan fitur atau atribut tertentu di dalamnya (1).

Panen padi di Kabupaten Jembrana dipengaruhi oleh berbagai faktor seperti cuaca, penggunaan pupuk, dan kepadatan tanaman, maka sangat penting bagi petani untuk mengetahui prediksi panen padi yang dapat membantu mereka merencanakan kegiatan bercocok tanam dengan lebih baik. Dengan sistem prediksi panen padi, petani dapat mengambil keputusan yang lebih baik dalam hal penggunaan pupuk dan irigasi, serta mengantisipasi dan mengatasi permasalahan yang mungkin terjadi seperti serangan hama dan penyakit pada tanaman. Selain itu, dengan sistem prediksi panen padi, petani dapat meningkatkan efisiensi penggunaan sumber daya dan mengoptimalkan hasil panen mereka. Oleh karena itu, sangat penting untuk mengembangkan sistem prediksi panen padi bagi petani di Kabupaten Jembrana.

TINJAUAN PUSTAKA

Penelitian terkait prediksi data menggunakan metode naive bayes telah banyak dilakukan, seperti yang dalam penelitian dengan judul Implementasi Algoritma Data Mining Naive Bayes untuk Prediksi Kelulusan Mahasiswa. Dalam penelitian ini dilakukan prediksi kelulusan mahasiswa dengan menggunakan Metode Data Mining dan Algoritma Naive Bayes bertujuan untuk menemukan informasi kelulusan mahasiswa dan membantu pihak manajemen dalam membuat kebijakan untuk prestasi mahasiswa agar dapat menyelesaikan studinya dengan hasil yang memuaskan (2).

Penelitian lain membahas tentang penggunaan model pendekatan Bayesian dengan klasifikasi Naive Bayes dan HMAP untuk memprediksi kejadian resiko tinggi pada

kehamilan ibu. Data yang digunakan adalah karakteristik usia ibu, tinggi badan, jumlah Hb, tekanan darah, riwayat kehamilan lalu, dan penyakit bawaan, yang kemudian didiskritkan berdasarkan batasan yang dipakai oleh Departemen Kesehatan. Hasil prediksi berupa probabilitas terjadinya resiko, yang dapat digunakan sebagai rujukan tempat melahirkan atau penilaian kinerja dari penyelenggara jasa persalinan. Penggunaan algoritma Naïve Bayes ini diharapkan dapat membantu menurunkan Angka Kematian Ibu (AKI) dan Angka Kematian Bayi (AKB) dengan memprediksi resiko secara dini sehingga dapat ditanggulangi sejak awal kehamilan. Fungsi klasifNB dalam bahasa R digunakan untuk melakukan training dan prediksi dengan memperhitungkan karakteristik ibu hamil secara dinamis sesuai wilayah yang dipilih (3).

Penelitian terkait prediksi panen padi sudah pernah dilakukan pada penelitian dengan judul “Sistem Pakar Penyakit Padi Menggunakan Metode Certainty Factor Di Desa Giling, Pati Jawa Tengah”. Penelitian ini bertujuan untuk membuat sebuah aplikasi sistem pakar berbasis web yang dapat membantu para petani dalam menentukan cara pengendalian penyakit padi yang sedang dihadapi. Tujuan aplikasi ini adalah untuk meminimalisir kerusakan yang disebabkan oleh penyakit padi yang akan menurunkan hasil panen. Selain itu, aplikasi ini juga memberikan pengetahuan tentang penyakit padi dan cara pengendaliannya kepada para petani (4).

Penelitian berikutnya membahas tentang pencarian solusi atas rendahnya produksi beras di beberapa wilayah di Indonesia, dengan menganalisis daerah mana saja yang mempunyai potensi peningkatan atau percepatan produksi. Metode yang digunakan adalah pendekatan data mining konsep K-means pada rekaman data produksi padi dari tahun 1993 sampai dengan 2012 dari 33 provinsi di Indonesia. Dari analisis tersebut, didapatkan urutan pengembangan produksi padi yang terdiri dari enam kelompok. Produksi tertinggi padi nasional masih didominasi oleh Pulau Jawa dan Bali, namun pengembangan lebih lanjut dapat difokuskan pada wilayah Sulawesi, NTB, dan Sumatera

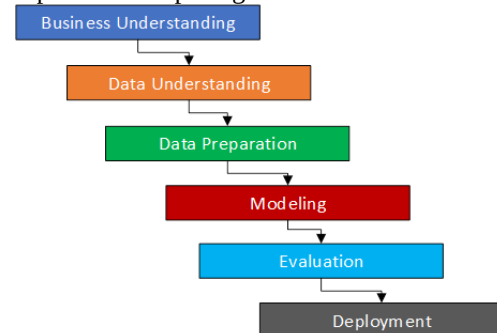
Barat. Simulasi hasil optimasi telah mengakomodir riwayat produksi pada dimensi masa lampau, saat ini, dan perkiraan produksi mendatang berdasarkan peningkatan yang terjadi. Dengan demikian, penelitian ini memberikan rekomendasi untuk pengembangan produksi beras yang lebih efektif dan efisien di Indonesia (5).

Penelitian lainnya berjudul “Sistem Cerdas dalam Klasifikasi Kematangan Buah Jeruk Berdasarkan Fitur Ekstraksi GLCM dengan Metode Naïve Bayes”. Penelitian ini membahas tentang penggunaan teknologi informasi dan komunikasi dalam bidang pertanian dan perkebunan, khususnya pada penggunaan metode klasifikasi dengan algoritma Naïve Bayes pada sistem cerdas untuk mengklasifikasikan kematangan buah jeruk keprok. Para petani kebun sering mengalami kesulitan dalam mengelompokkan buah jeruk yang matang karena keterbatasan fisik, sehingga penggunaan sistem cerdas dapat meningkatkan efektivitas dan efisiensi dalam penyortiran buah jeruk yang matang (6).

METODE PENELITIAN

Metode Penelitian

Model yang digunakan untuk menyelesaikan penelitian ini adalah CRISP-DM (Cross Industry Standard Process for Data Mining). Model CRISP-DM diperkenalkan pertengahan tahun 1990 oleh sebuah perusahaan konsorsium Eropa. Dalam CRISP-DM, sebuah proyek Data Mining memiliki siklus hidup yang terbagi dalam 6 fase (7), seperti terlihat pada gambar :



Gambar 1. Fase CRISP-DM (Cross Industry Standard Process for Data Mining)

1. Fase Pemahaman Bisnis (Business Understanding Phase), meliputi penentuan tujuan bisnis, menilai situasi saat ini, menetapkan tujuan data mining, dan mengembangkan rencana proyek. Tujuan bisnis yang akan dilakukan pada penelitian ini adalah menghasilkan prediksi hasil panen padi dengan metode klasifikasi naïve bayes.
2. Fase Pemahaman Data (Data Understanding Phase), setelah tujuan bisnis dan rencana proyek ditetapkan, langkah selanjutnya melakukan pengumpulan data awal, deskripsi data, eksplorasi data, dan verifikasi kualitas data. Penelitian yang diusulkan ini menggunakan data primer, dengan data yang digunakan adalah data hasil pertanian padi periode tahun 2021-2022 di wilayah kabupaten Jembrana.
3. Fase Persiapan Data (Data Preparation Phase), pada tahap ini dilakukan identifikasi dan pembangunan jawaban dari data yang telah dikumpulkan untuk bisa melakukan pengelompokan dan pemilihan ke dalam kelompok-kelompok yang telah ditentukan. Atribut yang digunakan dalam penelitian ini berupa atribut lahan, musim, isi buah padi, air, daun, dan batang.
4. Fase Pemodelan (Modeling Phase), pada fase ini dilakukan pemilihan model yang akan digunakan untuk melakukan klasifikasi hasil panen padi. Model atau metode yang akan digunakan pada penelitian ini adalah metode Naïve Bayes. Jumlah data latih yang akan digunakan terdiri dari 50 record yang diambil dari 19 desa di wilayah Kabupaten Jembrana.
5. Fase Evaluasi (Evaluation Phase), pengujian akan dilakukan dengan membandingkan prediksi yang dilakukan oleh algoritma Naïve Bayes Classifier dengan prediksi yang dilakukan oleh Penyuluh Pertanian di kabupaten Jembrana.
6. Fase Penyebaran (Deployment Phase), fase ini dilakukan untuk menghasilkan penemuan pengetahuan (identifikasi hubungan yang tak terduga dan berguna) untuk kemudian diterapkan pada operasi bisnis di berbagai tujuan.

IMPLEMENTASI SISTEM

Pada hasil penelitian ini, penulis memberikan penjelasan hasil dari penelitian yang dilakukan pada metodologi penelitian. Berikut ini penjelasan dari hasil yang dilakukan pada penelitian ini:

Analisa Perhitungan Naive Bayes Classifier

Algoritma Naïve Bayes Classifier adalah salah satu algoritma yang terdapat pada teknik klasifikasi. Naive Bayes Classifier merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes yaitu memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan Naive dimana diasumsikan kondisi antara atribut saling bebas. Dasar formula teorema bayes yang digunakan adalah :

$$P(H|X) = \frac{P(H) \times P(X|H)}{P(X)}$$

Keterangan :

- X = Data dengan kelas yang belum diketahui
- H = Label Kelas
- P(H) = Probabilitas dari Hipotesa H
- P(X) = Probabilitas X
- P(H|X) = Probabilitas Hipotesis H berdasarkan kondisi X.
- P(X|H) = Probabilitas X berdasarkan kondisi hipotesis H

Dalam penerapan metode Naïve Bayes Classifier ada tahapan sebagai berikut :

1. Hitung probabilitas dari setiap kelas: Dalam hal ini, probabilitas dari setiap kelas adalah jumlah data dalam kelas tersebut dibagi dengan jumlah data keseluruhan. Ini dapat dinyatakan sebagai: $P(C) = (\text{jumlah data dalam kelas } C) / (\text{jumlah data keseluruhan})$.
2. Hitung probabilitas dari setiap fitur: Dalam hal ini, probabilitas dari setiap fitur dalam suatu kelas adalah jumlah data dalam kelas C yang memiliki fitur X dibagi dengan jumlah data dalam kelas C. Ini dapat dinyatakan sebagai: $P(X | C) = (\text{jumlah data dalam kelas } C \text{ yang memiliki fitur } X) / (\text{jumlah data dalam kelas } C)$.

data dalam kelas C dengan fitur X) / (jumlah data dalam kelas C). Nilai X menunjukkan kriteria yang digunakan yaitu sebagai berikut :

Tabel 1 Kriteria

Kriteria	Intances
Luas Tanah	• Luas jika >5000 ha • Sedang jika 3000-5000 ha • Kecil jika < 20000 ha
Ketersediaan alat	• Baik jika >200 • Cukup jika 50-200 • Kurang jika <50
Curah hujan	• Rendah jika 0-100 mm • Cukup jika 100 mm – 300 mm • Tinggi jika >300

3. Tentukan probabilitas priori, dalam hal ini, probabilitas priori adalah probabilitas dari suatu kelas sebelum melihat data. Ini dapat ditentukan dengan menggunakan probabilitas kelas seperti yang dihitung pada langkah nomor 1.
4. Hitung probabilitas posteriori: Setelah menentukan probabilitas priori, langkah selanjutnya adalah menghitung probabilitas posteriori dari suatu kelas C dan fitur X. Ini dapat dinyatakan sebagai: $P(C | X) = (P(X | C) * P(C)) / P(X)$.
5. Tentukan kelas dengan probabilitas posteriori paling tinggi: Setelah menghitung probabilitas posteriori, tahap terakhir adalah menentukan kelas dengan probabilitas posteriori paling tinggi. Kelas ini akan menjadi kelas yang diberikan pada data uji.

Implementasi Naive Bayes Classifier pada Prediksi Hasil Panen Padi

Metode naive bayes classifier diimplementasikan pada sistem untuk melakukan klasifikasi hasil seleksi. Dikarenakan metode naive bayes classifier merupakan metode yang memiliki konsep terawasi, maka membutuhkan inputan seorang pakar untuk mengolah data latih atau data acuan. Contoh perhitungan manual Naive Bayes Classifier jika dimasukkan data uji sebagai berikut :

Tabel 2 Data latih

No	Luas	Alat	Hujan	Hasil
01	sedang	tinggi	cukup	baik
02	kecil	sedang	cukup	kurang
03	sedang	sedang	cukup	baik
04	luas	rendah	tinggi	baik
05	luas	sedang	tinggi	baik
06	luas	sedang	tinggi	baik
07	luas	sedang	tinggi	baik
08	luas	tinggi	tinggi	baik
09	kecil	sedang	tinggi	kurang
10	luas	rendah	rendah	kurang
11	sedang	rendah	cukup	baik
12	luas	sedang	cukup	kurang

Sedangkan data yang akan diuji adalah sebagai berikut :

Tabel 3 Data Uji

No	Luas	Alat	Hujan	Hasil
01	luas	tinggi	cukup	?

Untuk menentukan data uji masuk ke dalam kelas BAIK atau KURANG, maka dilakukan proses perhitungan menggunakan metode Naive Bayes yaitu sebagai berikut :

$$P(\text{Baik}) = 8 / 12 = 0.7$$

$$P(\text{Kurang}) = 4 / 12 = 0.3$$

Luas :

$$P_{\text{luas}}(\text{Luas} | \text{Baik}) = 5 / 8 = 0.6$$

$$P_{\text{luas}}(\text{Luas} | \text{Kurang}) = 2 / 4 = 0.5$$

Alat :

$$P_{\text{alat}}(\text{Tinggi} | \text{Baik}) = 5 / 8 = 0.6$$

$$P_{\text{alat}}(\text{Tinggi} | \text{Kurang}) = 2 / 4 = 0.5$$

Hujan :

$$P_{\text{hujan}}(\text{Cukup} | \text{Baik}) = 3 / 8 = 0.4$$

$$P_{\text{hujan}}(\text{Cukup} | \text{Kurang}) = 2 / 4 = 0.5$$

$P(\text{Hasil} | \text{Baik}) :$

$$\begin{aligned}
 &= P(\text{Baik}) \times P_{\text{luas}}(\text{Luas} | \text{Baik}) \times P_{\text{alat}}(\text{Tinggi} | \text{Baik}) \times P_{\text{hujan}}(\text{Cukup} | \text{Baik}) \\
 &= 0.7 \times 0.6 \times 0.6 \times 0.4 \\
 &= 0.10
 \end{aligned}$$

$$\begin{aligned}
 &P(\text{Hasil} \mid \text{Kurang}) : \\
 &= P(\text{Kurang}) \times P_{\text{luas}}(\text{Luas} \mid \text{Kurang}) \times P_{\text{alat}} \\
 &\quad (\text{Tinggi} \mid \text{Kurang}) \times P_{\text{hujan}}(\text{Cukup} \mid \text{Kurang}) \\
 &= 0.3 \times 0.5 \times 0.5 \times 0.5 \\
 &= 0.04
 \end{aligned}$$

Berdasarkan perhitungan di atas didapatkan hasil $P(\text{Hasil} \mid \text{Baik}) > P(\text{Hasil} \mid \text{Kurang})$, dengan demikian maka hasil prediksi panen padi adalah BAIK.

Evaluasi Analisa Metode Naive Bayes

Dari implementasi prediksi panen padi menggunakan metode bayes, maka hasil analisis pengolahan data tersebut harus dilakukan pengujian akurasi data. Dengan adanya proses pengujian ini dapat diketahui akurasi hasil analisis yang dibuat, berikut ini adalah hasilnya :

Tabel 3 Hasil Analisa Pakar dan Analisa dengan Metode Naive Bayes

No	Hasil Pakar	Hasil Metode Naive Bayes	Keterangan
13	Baik	Baik	Sesuai
14	Kurang	Diterima	Tidak Sesuai
15	Diterima	Kurang	Tidak Sesuai
16	Diterima	Diterima	Sesuai
17	Kurang	Kurang	Sesuai
18	Diterima	Diterima	Sesuai
19	Kurang	Kurang	Sesuai
20	Diterima	Kurang	Tidak Sesuai
21	Diterima	Diterima	Sesuai

Nilainya akurasinya sebagai berikut :

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{Jumlah Data yang Sesuai}}{\text{Jumlah Data Uji}} \times 100\% \\
 &= \frac{8}{9} \times 100\% \\
 &= 0.89 \times 100\% = 89\%
 \end{aligned}$$

Hasil pengujian perhitungan dengan metode Naive Bayes diperoleh akurasi sebesar 89 % dari total 9 data uji.

SIMPULAN

Penelitian ini telah berhasil menghasilkan prediksi panen padi di kabupaten Jembrana menggunakan metode Naive Bayes Classifier (NBC). Metode naive bayes classifier memberikan hasil perhitungan yang cukup akurat dalam memprediksi hasil

panen padi dengan tingkat akurasi sebesar 89% dengan data uji sebanyak 9 data.

DAFTAR PUSTAKA

- [1.] Rifai MF, Jatnika H, Purwanto YSS, Karmila S. Pengaruh Kondisi Cuaca Terhadap Serangan Hama Penggerek Batang Pada Tanaman Padi Di Desa Ciaruteun Ilir, Kec. Bungbulang, Kab. Bogor. PETIR. 2020;13(2).
- [2.] Shiddieq DF, Patricia. Implementasi Algoritma Data Mining Naive Bayes untuk Prediksi Kelulusan Mahasiswa. Jurnal Komputer Bisnis. 2020;(456).
- [3.] Nugroho A, Subanar S. Klasifikasi Naive Bayes untuk Prediksi Kelahiran pada Data Ibu Hamil. Bimipa. 2013;23(3).
- [4.] Santoso S, Julianti MR, Winarto AH. Sistem Pakar Penyakit Padi Menggunakan Metode Certainty Factor Di Desa Giling, Pati Jawa Tengah. JURNAL SISFOTEK GLOBAL. 2018;8(2).
- [5.] Sudono I, Utami W, Lestari S. Pengelompokan Produksi Padi Nasional dengan Pendekatan Data Mining Konsep K-Means. Jurnal Irigasi. 2016;8(2).
- [6.] Haba ARK, Pelangi KC. Sistem Cerdas Dalam Klasifikasi Kematangan Buah Jeruk Berdasarkan Fitur Ekstraksi GLCM Dengan Metode Naive Bayes. Jurnal Teknologi dan Manajemen Informatika. 2019;5(2).
- [7.] Kecerdasan Buatan J, dan Teknologi Informasi K, Hasyim F, Artikel R, Kunci Strategi Promosi K. Implementasi Data Mining Dalam Menentukan Strategi Promosi Program KB Menggunakan Algoritma K-Means Clustering. 2022;3(1). Available from: <https://ejournal.unuja.ac.id/index.php/core>