

IMPLEMENTASI METODE KLASIFIKASI NAIVE BAYES DALAM MEMPREDIKSI KELULUSAN MAHASISWA

IMPLEMENTATION OF NAIVE BAYES CLASSIFICATION METHOD IN PREDICTING STUDENT GRADUATION

Ananda Fiqri Firdaus¹, Rohmat Saedudin², Rachmadita Andeswari³

^{1,2,3} Universitas Telkom, Bandung

¹firdausfikri@student.telkomuniversity.ac.id, ²rdrohmat@telkomuniveristy.co.id,

³andreswari@telkomuniversity.ac.id

Abstrak

Persentase dari kelulusan tepat waktu mahasiswa merupakan salah satu aspek penting dalam akreditasi suatu perguruan tinggi. Oleh karena itu sangat penting untuk memastikan dan terus memotivasi mahasiswa agar dapat lulus tepat waktu. Dengan melakukan prediksi akan probabilitas kelulusan mahasiswa diharapkan para dosen dapat fokus memberikan perhatiannya kepada mahasiswa yang terprediksi kelulusannya rendah. Dengan mengimplementasikan klasifikasi data mining, memprediksi kelulusan mahasiswa akan jauh lebih mudah dan efisien. Tujuan Penelitian ini adalah untuk mengimplementasikan klasifikasi dengan metode naive bayes untuk memprediksi kelulusan mahasiswa. Penelitian ini juga menguji rasio antara data training dan data tes mana yang tepat untuk digunakan agar hasil yang diperoleh lebih optimal. Berdasarkan hasil penelitian ini diperoleh bahwa dengan menggunakan 80% data sebagai data training dihasilkan hasil akurasi paling optimal dengan tingkat akurasi 90.78%, presisi data sebesar 88% dan 88.4% untuk recall. Dan naive bayes memperoleh tingkat akurasi tertinggi sebesar 91.49% dengan menggunakan 70% data sebagai data training.

Kata kunci : klasifikasi, data mining, naive bayes, kelulusan mahasiswa

Abstract

The percentage of students graduating on time is one of the important aspects in the accreditation of a college. Therefore, it is very important to ensure and continue to motivate students to graduate on time. By predicting the probability of student graduation, it is hoped that lecturers can focus on paying attention to students who are predicted to have low graduation. By implementing data mining classification, predicting student graduation will be much easier and more efficient.

The purpose of this study was to implement a classification using the Naive Bayes method to predict student graduation. This study also examines the ratio between training data and test data which is appropriate to use so that the results obtained are more optimal. Based on the results of this study, it was found that by using 80% of the data as training data, the most optimal accuracy results were obtained with an accuracy rate of 90.78%, data precision of 88% and 88.4% for recall. And Naive Bayes obtained the highest accuracy rate of 91.49% by using 70% of the data as training data data

Keywords: classification, data mining, naive bayes, student graduation

1. Pendahuluan

Universitas Telkom adalah salah satu perguruan tinggi swasta yang ada di Indonesia. Universitas ini menawarkan layanan pendidikan yang berkualitas agar mahasiswanya dapat memiliki kemampuan yang kompeten dan berintegritas saat sudah lulus pada waktunya. Universitas Telkom memiliki 31 program studi yang dinaungi oleh tujuh fakultas salah satu diantaranya yaitu program studi Strata 1 (S1) Sistem Informasi yang berada dalam Fakultas Rekayasa Industri (FRI). Setiap tahunnya, program studi sistem informasi menerima mahasiswa baru dari berbagai daerah di Indonesia dengan jumlah yang cukup banyak. Dari banyaknya mahasiswa yang diterima masuk prodi ini hingga tahun 2018, program studi sistem informasi diperkirakan telah menghasilkan sekitar 1000 lulusan dihitung sejak berdirinya program studi sistem informasi pada tahun 2008. Dari jumlah tersebut diketahui bahwa tidak semua mahasiswa dapat menyelesaikan masa studinya dalam waktu yang ditentukan (empat tahun).

Bersumber dari data yang didapat pada website bachelor of information system telkom university, diperoleh bahwa telah terjadi penurunan persentase kelulusan mahasiswa hingga tahun 2018. Pada tahun 2018 persentase kelulusan mahasiswa program studi sistem informasi adalah 62%. Hal ini dapat menyebabkan turunnya kualitas dan mutu program studi dan mempengaruhi penilaian dan evaluasi terhadap program studi sistem informasi.

Mahasiswa dapat dinyatakan lulus atau tidak dapat dilihat dari Indeks Prestasi (IP), Indeks Prestasi Kumulatif (IPK) dan nilai huruf yang didapat dari setiap mata kuliah. Jika mahasiswa mendapatkan beberapa mata kuliah yang

nilainya dibawah D, mahasiswa harus mengulang matakuliah tersebut. Hal ini pun dapat memperlambat kelulusan mahasiswa karena harus mengulang mata kuliah[6]. Oleh karena itu, program studi perlu melakukan identifikasi sejak dini akan mahasiswa yang memiliki nilai yang kurang dan memberikan perhatian khusus agar mahasiswa itu dapat lulus tepat waktu, salah satu caranya adalah dengan melakukan prediksi kelulusan mahasiswa.

Dalam hal memprediksi kelulusan mahasiswa, penulis akan melakukan klasifikasi menggunakan data mining. Data mining digunakan untuk menemukan pola dalam data yang berjumlah besar, dalam data mining terdapat beberapa teknik seperti klasifikasi, clustering, asosiasi dll. Teknik yang cocok dengan kasus mengidentifikasi kebutuhan listrik ini dimana tujuannya yaitu memprediksi kebutuhan listrik berdasarkan pola data yang sudah ada maka dipilihlah teknik klasifikasi data mining.

Terdapat banyak algoritma yang dapat diterapkan pada klasifikasi data mining, dari sekian banyak algoritma penulis memutuskan untuk menggunakan algoritma Naïve Bayes. Alasan digunakannya algoritma ini adalah karena algoritma Naïve Bayes mudah untuk mengimplementasikannya dan juga metode ini cocok digunakan karena metode ini hanya membutuhkan jumlah data training yang kecil untuk menentukan parameter yang diperlukan dalam proses pengklasifikasian, Naïve Bayes memiliki tingkat akurasi yang tinggi berdasarkan penelitian sebelumnya [7].

Dalam hal memprediksi kelulusan mahasiswa ini maka ditetapkanlah untuk menggunakan teknik klasifikasi data mining dan menerapkan metode Naïve Bayes. Penelitian ini juga bertujuan untuk mengetahui proses pengklasifikasian dan juga mengukur keakuratan metode naive bayes. Dan memiliki manfaat dalam menjelaskan keakuratan dalam menggunakan metode Naïve Bayes untuk klasifikasi data mining bagi penelitian lain. Penelitian ini juga dapat dimanfaatkan untuk meningkatkan persentase kelulusan mahasiswa dengan mengambil tindakan pertama terhadap mahasiswa tahun awal yang terprediksi telat lulus

2. Dasar Teori /Material dan Metodologi/perancangan

2.1 Data Mining

Data Mining merupakan proses untuk menemukan sebuah pola yang unik berdasarkan atribut sebuah dataset yang berukuran besar dan dari data tersebut dapat disimpulkan sebuah pengetahuan. Data dapat didapat dari basis data, open-source API, website, ataupun dari hasil survey [1]. Menurut Rifai [3] data mining merupakan proses untuk menemukan hubungan antar data menggunakan teknik statistika dan matematika untuk mengenali pola pada data dari sekumpulan besar data yang tersimpan dalam suatu penyimpanan data. Tetapi data tersebut tidak bisa langsung digunakan untuk data mining, data tersebut harus melalui beberapa tahapan dalam penambangan data. Berikut proses tahapan dari awal hingga akhir dari penambangan data [5]:

a) Pembersihan data (Data Cleaning)

Pembersihan data merupakan proses menghilangkan noise ataupun data null, menghapus duplikasi data, memeriksa data inkonsisten, dan memperbaiki kesalahan pada data.

b) Integrasi data (Data Integration)

Adakalanya informasi yang ingin dihasilkan merupakan gabungan dari beberapa database. Untuk itu integrasi data diperlukan untuk penggabungan data dari berbagai database ke dalam satu database baru.

c) Seleksi data (Data Selection)

Pemilihan (Seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum diolah lebih lanjut. Dikarenakan tidak semua atribut data digunakan untuk suatu tujuan, oleh karena itu hanya data yang sesuai yang akan dianalisis yang diambil dari database. Data hasil seleksi itu akan disimpan kedalam berkas terpisah.

d) Transformasi data (Data Transformation)

Dengan melakukan proses coding dan mengubah bentuk data atau menggabungkan dengan data lain agar sesuai dengan format yang akan digunakan dalam data mining. Proses ini merupakan proses kreatif yang sangat bergantung pada jenis atau pola informasi yang ingin dicari.

e) Proses Mining

Merupakan suatu proses utama yaitu mencari pola atau informasi tertentu dengan menerapkan metode tertentu. Teknik, metode dan algoritma yang diterapkan pun bisa berbagai macam karna memiliki banyak variasi. Beberapa metode yang dapat digunakan berdasarkan tujuan dari informasi apa yang ingin ditarik dari data tersebut.

f) Evaluasi pola (Pattern Evaluation)

Untuk mengidentifikasi pola-pola menarik ke dalam knowledge based yang ditemukan. Dan memeriksa apakah pola yang didapat sesuai dengan fakta atau hipotesis yang ada sebelumnya.

g) Presentasi pengetahuan (Knowledge Presentation)

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.

2.2 Klasifikasi

Klasifikasi adalah proses penemuan model (fungsi) yang mendeskripsikan dan menyeleksi kelas data dan bertujuan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui [9]. Proses klasifikasi terdiri dari dua langkah yaitu model pelatihan (data training) dan model pengujian (data testing).

Langkah pertama disebut juga langkah pembelajaran, dimana algoritma klasifikasi digunakan untuk menganalisis data training kemudian direpresentasikan dalam bentuk rule klasifikasi. Proses ini juga bisa dibidang bertujuan untuk membangun proses klasifikasi. Kemudian langkah kedua adalah klasifikasi, dimana data testing digunakan untuk mengukur akurasi dari rule klasifikasi. Model klasifikasi yang dihasilkan bisa dipakai untuk menetapkan kelas dari data yang baru [2].

Teknik klasifikasi melakukan pengklasifikasian item data ke label kelas yang telah ditetapkan, membangun

model klasifikasi dari kumpulan data input, membangun model yang digunakan untuk memprediksi tren data masa depan. Algoritma yang umum digunakan antara lain yaitu *decision tree*, *Naïve Bayes classifier*, *k-nearest neighbor*, *k-means*, *C45*, dan *support vector machines (SVM)* [9].

2.3 Naïve Bayes

Naïve Bayes merupakan perhitungan teorema bayes yang paling sederhana, karena mampu mengurangi kompleksitas komputasi dengan asumsi independensi yang kuat (naif) antar satu kelas dengan kelas yang lain [9]. Naïve Bayes didasarkan pada asumsi simplifikasi bahwa nilai dari suatu atribut secara temporer saling bebas jika diberikan nilai output [4]. Dengan kata lain, dalam Naïve Bayes menggunakan model fitur independen, maksud independen yang kuat pada fitur adalah bahwa data tidak berkaitan dengan data yang lain dalam kasus yang sama ataupun atribut yang lain.

Berikut ini adalah rumus persamaan dari metode Naïve Bayes yaitu:

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \tag{1}$$

Dimana , X : Data dengan kelas yang belum diketahui.

H : Hipotesis data merupakan suatu kelas 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

2.4 Confusion Matrix

Confusion Matriks adalah salah satu cara untuk mengukur performa atas kinerja suatu klasifikasi yang mana mendapat keluaran dua kelas atau lebih [10]. Confusion matriks digambarkan berupa tabel dengan empat kombinasi berbeda dari nilai aktual dan nilai prediksi. Dari hasil klasifikasi dapat direpresentasikan kedalam tabel confusion matriks yaitu sebagai True Positif, True Negatif, False Positif, dan False Negatif. Dalam penjabarannya True Positif merupakan data positif yang diprediksi benar, True Negatif merupakan data negatif yang diprediksi benar, False Positif merupakan data negatif namun diprediksi sebagai data positif, dan False Negatif merupakan data positif namun diprediksi sebagai data negatif. Dengan menggunakan tabel confusion matriks ini kita dapat menggunakannya untuk mengukur beberapa performansi seperti akurasi, presisi, dan recall atau sensitivity. Persamaan untuk menghitungnya adalah sebagai berikut:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

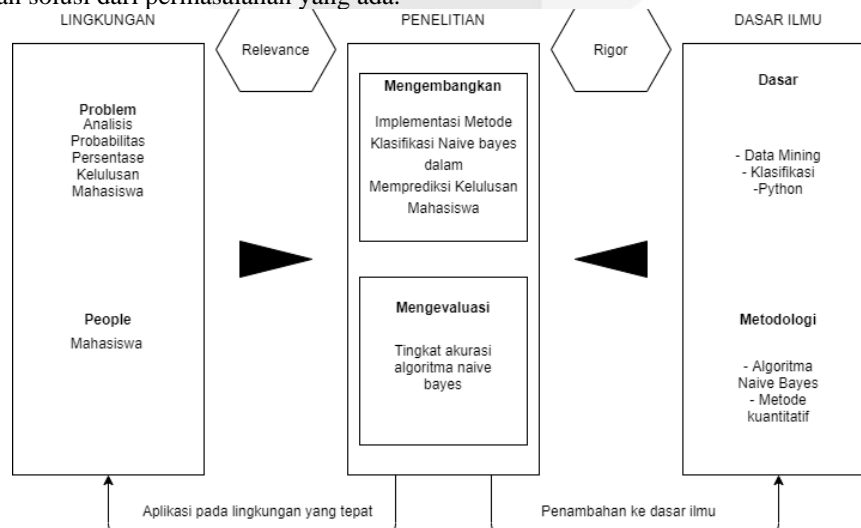
$$Presisi = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

3. Metode Penelitian

3.1 Model Konseptual

Model konseptual merupakan kerangka pemikiran yang dapat membantu memecahkan masalah dan membantu dalam merumuskan solusi dari permasalahan yang ada.



Gambar 1. Model Konseptual

Berdasarkan model konseptual pada gambar 1, pada lingkungan terdapat aspek problem, people dan teknologi. Aspek problem merupakan masalah yang diangkat untuk penelitian ini yaitu analisis probabilitas persentase kelulusan mahasiswa. Aspek people yaitu subjek yang terlibat dalam penelitian ini merupakan mahasiswa.

Selanjutnya pada penelitian terdapat aspek mengembangkan dan mengevaluasi. Aspek penelitian yaitu implementasi metode naive bayes dalam memprediksi kelulusan mahasiswa, yang menyesuaikan dengan studi kasus yang sedang diteliti. Aspek mengevaluasi merupakan hasil penelitian yang digunakan untuk mengukur tingkat akurasi algoritma naive bayes yang mana akurasi algoritma ini dihasilkan dengan membandingkan hasil yang diperoleh berdasarkan perhitungan manual, dan menggunakan perangkat lunak WEKA berdasarkan rasio data training dan data testing.

Pada dasar ilmu, aspek dasar memiliki penjabaran dasar ilmu yang terkait seperti data mining, klasifikasi, algoritma naive bayes dan python. Dasar yang dijabarkan merupakan dasar ilmu paling utama dalam penelitian ini. Pada aspek metodologi terdapat naive bayes yang merupakan metode klasifikasi yang digunakan dan metode kuantitatif yang merupakan metode penelitian yang diterapkan pada penelitian ini.

3.2 Sistematika Penyelesaian Masalah

Pada Penelitian ini sistematika penyelesaian masalah akan dibagi kedalam empat tahapan yaitu identifikasi masalah, preprocessing, processing, hasil dan kesimpulan.

3.2.1 Tahap Identifikasi Masalah

Pada tahapan ini terdapat beberapa aktivitas diantaranya yaitu melakukan identifikasi latar belakang permasalahan, perumusan masalah, menentukan tujuan, pemaparan manfaat penelitian dan identifikasi apa yang akan dikaji dalam penelitian ini. Pada tahapan ini juga dilakukan studi literatur mengenai teori-teori yang akan digunakan dalam penelitian ini, serta menentukan model konseptual untuk menggambarkan pengerjaan dan membantu penyelesaian masalah dengan merumuskan solusi yang akan digunakan. Pada tahap ini penulis juga menentukan sumber data serta menetapkan data-data yang akan digunakan.

3.2.2 Tahap Preprocessing

Pada tahapan ini data yang telah didapatkan akan diolah secara lebih lanjut agar memudahkan dalam melakukan klasifikasi. Selanjutnya data tersebut masuk dalam tahapan preprocessing, dalam tahapan tersebut data melalui proses data cleaning, data selection, dan data transformation. Data cleaning dimaksudkan untuk memeriksa kelengkapan data bahwa tidak terdapat data yang null. Kemudian data selection dimaksudkan untuk menyeleksi data yang ingin dipakai untuk penelitian. Dan terakhir yaitu data transformation yang dimaksudkan untuk mengubah format tabel data yang didapat agar lebih mudah dipahami dan diproses lebih lanjut.

3.2.3 Tahap Processing

Setelah dilakukan tahapan preprocessing pada data, selanjutnya data masuk ke tahapan processing. Pada tahapan ini perlu ditentukannya class dan probabilitas dari data yang ada supaya bisa diproses kedalam sebuah algoritma naive bayes, kemudian dilakukan perhitungan menggunakan rumus algoritma naive bayes. Selain menggunakan perhitungan menggunakan algoritma naive bayes, pengimplementasian juga menggunakan tools yang akan mendukung proses klasifikasi. Pertama yaitu menggunakan WEKA untuk melakukan perhitungan probabilitas secara otomatis oleh sistem. Perhitungan menggunakan WEKA dilakukan beberapa kali berdasarkan rasio yang ditentukan. Setelah hasil perhitungan tampil, hasil tersebut dilakukan uji akurasi menggunakan confusion matrix, selain untuk menguji akurasi confusion matrix juga dapat digunakan untuk menguji presisi dari data, dan recall atau kerelevanan dari data.

3.2.4 Tahap Kesimpulan dan Saran

Setelah mendapatkan probabilitas yang dihasilkan dengan algoritma naive bayes dan juga memeriksa keakuratannya menggunakan tools WEKA. Dengan membandingkan hasil dan keakuratan data menggunakan confusion matrix, selanjutnya dilakukan analisis terhadap hasil dan menyimpulkan apakah hasil yang didapat sesuai dengan rumusan masalah dan hipotesa awal. Dan membandingkan kinerja tool tersebut berdasarkan penentuan rasio perbandingannya. Juga memberikan saran terkait proses penelitian agar berguna untuk penelitian selanjutnya.

4. Pembahasan

4.1 Pengumpulan Data

Untuk data yang akan digunakan pada penelitian kali ini, data yang akan digunakan adalah data nilai mahasiswa dan probabilitas kelulusannya. Data ini diperoleh dari Fakultas Rekayasa Industri program studi Sistem Informasi. Data yang dikumpulkan adalah data nilai mahasiswa. Data yang dikumpulkan berjumlah 1249 baris data yang berisi tentang nilai UM, UAN, IP1, IP2, IP3, IP4 dan probabilitas kelulusan. Probabilitas kelulusan didapatkan berdasarkan perhitungan rata-rata nilai.

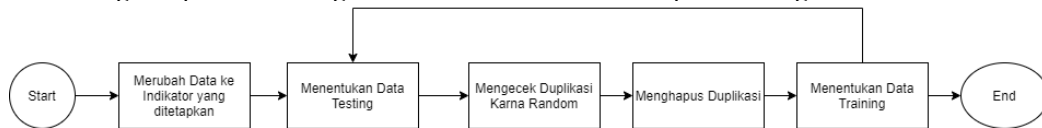
4.2 Pengolahan Data

Data yang didapatkan masih terdapat beberapa nilai kosong, karena itu harus dilakukan proses preprocessing. Dengan menggunakan tools jupyter notebook sebagai media untuk pengolahan data dilakukan proses data cleansing dengan alur sebagai berikut



Gambar 2. Flowchart Data Cleansing

Setelah data dibersihkan dan dihapuskan nilai kosong, data yang tersisa adalah 1138 baris data. Kemudian data perlu ditransformasikan agar dapat diolah dengan baik. Data tersebut melalui proses sebagai berikut.



Gambar 3. Flowchart Data Transformation dan Sampling

Setelah ditransformasikan, data dibagi menjadi data training dan data testing. Disini peneliti menguji keakuratan metode naive bayes berdasarkan perbandingan data trainingnya. Karna itu pada tahap ini rasio data training yang ditentukan adalah 80:20, 70:30, dan 60:40.

4.3 Implementasi Naive Bayes

Dengan menggunakan tools WEKA maka untuk rasio data training 80:20 didapatkanlah hasil 90.7895% data yang dapat diklarifikasikan dengan benar dan sebesar 9.2105% data yang tidak dapat diklarifikasikan dengan benar. Selanjutnya adalah melakukan perhitungan jika persentase data training dan data testing sebesar 70:30, dengan menggunakan rasio ini didapatkan hasil 91.4956% data yang dapat diklarifikasikan dengan benar dan sebesar 8.5044% data yang tidak dapat diklarifikasikan dengan benar. Dan untuk perhitungan naive bayes menggunakan rasio data training dan data testing sebesar 60:40, didapatkan hasil 89.011% data yang dapat diklarifikasikan dengan benar dan sebesar 10.989% data yang tidak dapat diklarifikasikan dengan benar.

4.4 Analisis Hasil

Setelah dilakukan perhitungan, selanjutnya adalah menguji performa dari hasil tersebut. Tahap pengujian tersebut dilakukan untuk mengecek kembali keakuratan perhitungan yang dilakukan oleh program yang digunakan. Adapun metode yang digunakan untuk pengujian adalah metode confusion matrix.

=== Confusion Matrix ===

```

a b c d e <-- classified as
76 4 2 0 0 | a = 75%
7 46 0 2 0 | b = 50%
2 0 61 0 0 | c = 100%
0 2 0 19 1 | d = 25%
0 0 0 1 5 | e = 0%
  
```

Gambar 4. Confusion Matriks 80:20

Dari hasil yang didapat menggunakan perbandingan data training 80:20 dapat disimpulkan bahwa hasil keakuratan data terhadap pengujian sebesar 90.78% , hasil presisi data terhadap pengujian sebesar 88% hasil kerelevanan data terhadap pengujian sebesar 88.4%.

=== Confusion Matrix ===

```

a b c d e <-- classified as
120 6 3 0 0 | a = 75%
7 71 0 3 0 | b = 50%
3 0 85 0 0 | c = 100%
0 2 0 30 2 | d = 25%
0 0 0 3 6 | e = 0%
  
```

Gambar 5. Confusion Matriks 70:30

Dari hasil yang didapat menggunakan perbandingan data training 70:30 dapat disimpulkan bahwa hasil keakuratan data terhadap pengujian sebesar 91.49% , hasil presisi data terhadap pengujian sebesar 86% hasil kerelevanan data terhadap pengujian sebesar 87%.

=== Confusion Matrix ===

```

a b c d e <-- classified as
156 13 4 0 0 | a = 75%
7 82 0 4 0 | b = 50%
5 0 117 0 0 | c = 100%
0 9 0 40 7 | d = 25%
0 0 0 1 10 | e = 0%
  
```

Gambar 6. Confusion Matriks 60:40

Dari hasil yang didapat menggunakan perbandingan data training 60:40 dapat disimpulkan bahwa hasil keakuratan data terhadap pengujian sebesar 89.01% , hasil presisi data terhadap pengujian sebesar 87% hasil kerelevanan data terhadap pengujian sebesar 82.4%.

5. Kesimpulan

- 1) Dengan mengimplementasikan metode naive bayes dalam memprediksi kelulusan mahasiswa didapatkan hasil akurasi sebesar 90.78%, presisi sebesar 88%, recall sebesar 88.4%. Dan dengan menggunakan aplikasi WEKA dalam mengimplementasikan metode ini waktu yang ditempuh untuk membangun model adalah 0.01 detik dan untuk mengetes model adalah 0.04 detik.
- 2) Dari rasio data training dan data testing yang telah diujikan yaitu 60:40, 70:30, dan 80:20. Didapatkan hasil bahwa dengan menggunakan rasio 80:20 dapat menghasilkan tingkat akurasi, presisi dan recall yang paling optimal secara keseluruhan.

Saran untuk penelitian selanjutnya adalah untuk menambahkan jumlah data sehingga dapat memaksimalkan hasil data training agar dapat meningkatkan nilai akurasi prediksi lebih maksimal. Melakukan penelitian dengan membandingkannya dengan beberapa metode lain dan aplikasi lain agar dapat hasil yang lebih pasti dan dapat menganalisa kelebihan masing-masing metode. Dan menggunakan data yang memiliki atribut yang lebih variatif agar atribut prediktor lebih jelas pengaruhnya terhadap kelas yang ingin diprediksi.

Referensi :

- [1] Fallo, D., & Sogen, M. 2018. Metode Naive Bayes untuk Memprediksi Penggunaan Listrik Rumah Tangga. Jurnal Pendidikan Teknologi Informasi JUKANTI.
- [2] Indriani, A. 2014. Klasifikasi Data Forum Dengan Menggunakan Metode Naive Bayes Classifier.
- [3] Rifai, M. F., Jatnika, H., & Valentino, B. 2019. Penerapan Algoritma Naive Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist MOS . Jurnal Pengkajian dan Penerapan Teknik Informatika.
- [4] Sadli, M., Fajriana, Fuadi, W., Ermatita, & Pahendra, I. 2018. Analisis Model Naive Bayes Untuk Identifikasi Penggolongan Daya Listrik Di Kota Lhoksumawe.
- [5] Saleh, A. 2015. Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. Citec Journal.
- [6] Salim, A. P. 2019. Prediksi Kelulusan Tepat Waktu Mahasiswa Setiap Tingkat Menggunakan Algoritma KNN .
- [7] Syarli, & Muin, A. A. 2016 . Metode Naive Bayes Untuk Prediksi Kelulusan Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi. Jurnal Ilmiah Ilmu Komputer.
- [8] Bachelor of Information System. 2021. Tentang Sistem Informasi. Diambil kembali dari <https://bis.telkomuniversity.ac.id/about-us-si/>
- [9] Widiawati, W. Y., & Atok, R. M. 2018. Analisis Klasifikasi Pelanggan Listrik Rumah Tangga Bersubsidi Kota Surabaya Menggunakan Support Vector Machine dan Naive Bayes Classifier. JURNAL SAINS DAN SENI ITS .
- [10] Wungo, S. L. 2017. Identifikasi Perilaku Pemakaian Energi Listrik Pelanggan Menggunakan Metode Boosting Naive Bayes.