

Prediksi Tingkat Kelulusan Tepat Waktu Mahasiswa Menggunakan Algoritma *Naive Bayes* pada Universitas XYZ

Nurhayati¹, Nuraeny Septianti², Nani Retnowaty³, Arief Wibowo⁴

^{1,2,3,4} Fakultas Teknologi Informasi Magister Ilmu Komputer Universitas Budi Luhur, Jakarta, Indonesia

¹ hayatinur10@gmail.com

² septiantireny@gmail.com

³ retnowatinani88@gmail.com

⁴ arief.wibowo@budiluhur.ac.id

Diterima 20 Juli 2020

Disetujui 23 November 2020

Abstract—Data processing is imperative for the development of information technology. Almost any field of work has information about data. The data is made use of the analysis of the job. Nowadays, information data is imperatively processed to help workers in making decisions. This study discusses student prediction graduation rates by using the naïve Bayes method. That aims at providing information to college if they can use it properly to utilize the data of students who graduated by processing data mining. Based on the data mining process, steps founded that used producing information, namely predicting student graduation on time. The method of this study is Naïve Bayes with classification techniques. At this study, researchers used a six-phase data mining process of industry crossing standards in data mining known as CRISP-DM. The results of research concluded that the application of the Naive Bayes algorithm uses 4 (four) parameters namely ips, ipk, the number of credits, and graduation by getting an accuracy value of 80.95%.

Index Terms—classification, graduation, Naive Bayes, student data

I. PENDAHULUAN

Teknologi dan informasi yang semakin pesat, membuat tidak terhitungnya data informasi dalam kehidupan manusia, diamati secara jelas pada bidang pengolahan data yang berjumlah besar dalam penyimpanan datanya. Hal ini menjadi daya tarik besar pada perusahaan dan organisasi baik negeri ataupun swasta untuk memiliki penyimpanan data yang cukup besar kapasitasnya. Kemampuan teknologi dan informasi untuk mengumpulkan data menyimpan berbagai tipe data yang jumlahnya sangat besar, umumnya mendukung dalam segi pengolahan data internal maupun eksternal dalam transaksi perusahaan serta layanan yang didukung dan dikelola oleh teknologi informasi [1].

Pemanfaatan data dalam sebuah perusahaan untuk menunjang pengambilan keputusan tidak cukup dalam sistem operasi saja, diperlukan analisis data

perusahaan untuk mendapatkan hasil kajian yang tepat dan akurat. Hal ini menjadi daya tarik dalam pemanfaatan ilmu yang dapat menyelesaikan masalah data dengan jumlah besar menjadi sebuah informasi. Data mining dapat dilakukan dengan sebuah aplikasi seperti *weka*, *spss clementine*, *matlab*, *dataminer*, *orangeanvas*, *dataengine*, *dbminer*, *webminer* maupun *rapidminer* yang dapat mempermudah dalam penggalian data. Menurut Witten, Frank & Hall, 2011. “*The Waikato Environment for Knowledge Analysis (WEKA)* adalah perkumpulan data yang lengkap dan diimplementasikan *State-of-the-art* dalam pembelajaran dan algoritma di *data mining*”. *Statistical Product and Service Solution (SPSS)* dipilih sebagai aplikasi yang digunakan untuk pengolahan data dengan prosedur statistik yang digunakan dalam bidang bisnis mulai dari tingkat sederhana.

Didalam dunia pendidikan sangat penting bagi pendidik dan mahasiswa, dalam menentukan kelulusan pada mata kuliah yang ditempuh pada setiap semester. Data kelulusan mahasiswa menjadi sangat penting dikarenakan data tersebut dapat menjadi tolak ukur instansi. sehingga penelitian ini dapat digunakan oleh program studi dalam menentukan kelulusan tepat waktu Pada penelitian prediksi kelulusan ada beberapa metode yang biasanya digunakan. Seperti, dengan menggunakan metode *Naive Bayes* yaitu teknik klasifikasi dengan metode probabilitas dan statistik, metode *K-Nearest Neighbor (KNN)* yaitu metode klasifikasi dengan data terbaru dan data terdekatnya (Gorunescu, 2011) serta algoritma C4.5 dalam proses pembuatan pohon keputusan.

Oleh karena itu, dari data yang diperoleh pada penelitian ini akan dilakukan pemodelan klasifikasi data mahasiswa untuk memprediksi kelulusan nilai dengan *Naive Bayes* [2]. Sumber data penelitian diperoleh dari data lulusan dari Universitas XYZ yang diwisuda pada tahun 2019. Data bersumber dari *database* akademik mahasiswa menjadi lulusan di

sepuluh program studi, terdiri dari atribut nim, nama, ips, ipk, jumlah sks, dan kelulusan.

II. TINJAUAN PUSTAKA

A. Analisis

Menurut Nasution dalam Sugiyono (2010:244) analisis adalah kemampuan yang menganalisis dataset dalam pekerjaan yang sulit dan memerlukan kerja keras. Tidak ada cara tertentu yang dapat diikuti untuk mengadakan analisis, sehingga setiap peneliti harus mencari sendiri metode yang baik sesuai dengan sifat penelitiannya. Bahan yang sama bisa diklasifikasikan berbeda [3].

B. Data Mining

Menurut Larose, (2005) *data mining* didefinisikan sebagai sebuah proses untuk menentukan hubungan pola dan tren baru yang bermakna dengan menyaring, memakai data dengan metode pola [4].

Menurut Turban, dkk. (2005) *data mining* adalah kata lain dari menjabarkan database yang telah didapatkan. *Data mining* adalah proses dengan menggunakan metode statistik dan matematika agar database bisa digolongkan untuk mendapatkan data yang dibutuhkan dari sumber *database* [5].

Menurut Jefri, (2013) *data mining* adalah identifikasi data dalam jumlah data yang cukup besar untuk menentukan hubungan yang tidak diketahui sebelumnya dan dua metode baru untuk dalam menyerderhanakan data agar mudah diproses serta digunakan untuk memilih data [4].

Menurut hasil studi Priati, (2018) *data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar [4].

Jadi dapat dipahami, *data mining* menurut peneliti adalah cara memperoleh beragam informasi dari banyaknya data yang tersimpan dilakukan dengan domain aplikasi yang diinginkan.

C. Teknik Klasifikasi

Klasifikasi adalah pengelompokan data untuk menemukan model bertujuan data mempunyai kelas untuk memprediksi perbedaan kelas objek yang akan dianalisis [6].

D. Algoritma Naïve Bayes

Algoritma *Naive Bayes* digunakan untuk prediksi peluang atau kemungkinan suatu kelas. mempunyai asumsi yang lebih kuat untuk tidak terkait dari masing-masing kondisi [2].

E. Tools Rapidminer

Rapidminer adalah salah satu alat bantu atau aplikasi yang digunakan dalam pengolahan data mining dan salah satu paket data mining yang dapat digunakan untuk perhitungan dan analisis secara lengkap. Peneliti ini menggunakan *rapidminer studio*.

III. METODOLOGI PENELITIAN

Dalam penelitian ini kami sebelumnya telah menentukan objek penelitian yaitu mahasiswa Universitas XYZ sebanyak 1157 mahasiswa, dengan data nilai dalam satu semester perkuliahan. Dalam proses pengumpulan data kami menghubungi pihak kampus terlebih dahulu. Dan proses selanjutnya adalah menyeleksi data, menentukan metode yang akan digunakan dan memproses data. Setelah data selesai diproses, kami lakukan evaluasi untuk hasil penelitian [1]. Berikut adalah beberapa metode yang kami gunakan dalam penelitian ini.

A. Pengumpulan Data

Dalam tahap ini kami mengumpulkan data mahasiswa dari Universitas XYZ dari semester 1 (satu) sampai 8 (delapan) dari beberapa mata kuliah yang diambil dalam semester tersebut. Atribut data terdiri dari nama, NIM, program studi, nilai-nilai IPS, dan IPK serta total jumlah SKS.

B. Selection

Setelah dilakukan pengumpulan data maka tahap selanjutnya yaitu menyeleksi data karena terdapat data yang muncul berulang, hasil seleksi yang didapat untuk selanjutnya digunakan dalam pengolahan penelitian.

C. Cleaning

Cleaning adalah proses pembersihan data-data mahasiswa yang tidak diperlukan dalam penelitian.

D. Transformasi

Perubahan data yang sesuai untuk diproses dalam penelitian selanjutnya.

E. Pengolahan Data

Dalam tahap pengolahan data dilakukan proses data seleksi dan transformasi.

F. Evaluasi Data

Dalam tahap evaluasi mengetahui apakah model sudah sesuai dengan tujuan penelitian. tujuan dilakukan penelitian ini adalah untuk mengetahui nilai akurasi yang tinggi, sehingga dapat membuktikan bahwa penelitian yang dilakukan sudah berhasil.

IV. HASIL DAN PEMBAHASAN

Pada bab ini akan dibahas deskripsi dari data yang sudah diperoleh, serta analisa prediksi kelulusan pada

data mahasiswa dengan menggunakan tools rapidminer menggunakan metode klasifikasi algoritma *Naïve Bayes*. Pembahasan dari hasil analisis akan memberikan penjelasan yang lebih detail.

A. Data Mahasiswa

Berikut ini data mahasiswa yang akan dianalisis oleh peneliti dengan menggunakan atribut nim, nama, program studi, status mahasiswa, ips, ipk, dan jumlah sks yang di ambil oleh mahasiswa [7]. Sampel dataset penelitian terlihat sebagaimana pada Gambar 1.

| NIM | Nama | Program Studi | IPS 1 | IPS 2 | IPS 3 | IPS 4 | IPS 5 | IPS 6 | IPS 7 | IPS 8 | IPK | Jumlah sks total | Kelulusan |
|----------------|----------------------|---------------|-------|-------|-------|-------|-------|-------|-------|-------|------|------------------|-------------------|
| 15416274201091 | Muallif Ismail | 74201 | 3.63 | 1 | 0 | 0 | 0 | 3.81 | 3.71 | 0 | 1.41 | 80 | Tidak Tepat Waktu |
| 15416274201092 | Aphanius | 74201 | 3.25 | 3.33 | 3.18 | 3.25 | 3.11 | 3.64 | 0 | 3.25 | 3.29 | 118 | Tidak Tepat Waktu |
| 15416274201093 | Muhammad Fadhri Huma | 74201 | 3.54 | 2.87 | 3.15 | 3.05 | 0.91 | 3.25 | 0 | 3.05 | 2.80 | 145 | Tidak Tepat Waktu |
| 15416274201094 | Sis Sarah | 74201 | 3.28 | 3.32 | 3.39 | 0 | 3.29 | 0 | 1.14 | 0 | 2.21 | 145 | Tidak Tepat Waktu |
| 15416274201095 | Purnadi | 74201 | 3.39 | 3.11 | 3.14 | 3.17 | 0 | 1 | 3.8 | 3.17 | 2.30 | 145 | Tidak Tepat Waktu |
| 15416274201096 | Tiwad | 74201 | 3.57 | 3.58 | 3.47 | 0 | 3.55 | 1.19 | 0 | 0 | 2.56 | 126 | Tidak Tepat Waktu |
| 15416274201098 | Tawardana | 74201 | 3.7 | 3.53 | 3.38 | 3.55 | 3.58 | 1.7 | 3.92 | 3.53 | 3.24 | 145 | Tidak Tepat Waktu |
| 15416274201101 | Hari Azera | 74201 | 3.36 | 2.41 | 0.95 | 1.62 | 0.87 | 1.62 | 0 | 1.62 | 1.81 | 80 | Tidak Tepat Waktu |
| 15416274201102 | Prawira Indra | 74201 | 2.98 | 2.24 | 2.3 | 1.85 | 3.51 | 2.04 | 1.14 | 1.85 | 2.49 | 142 | Tidak Tepat Waktu |
| 15416274201105 | Sufian | 74201 | 3.28 | 3.05 | 0 | 0 | 0 | 2.11 | 0 | 0 | 1.07 | 145 | Tidak Tepat Waktu |
| 15416274201106 | Rizki Ratna | 74201 | 3.21 | 3.66 | 3.18 | 3.52 | 3.63 | 3.79 | 4 | 3.52 | 3.50 | 144 | Tepat waktu |
| 15416274201104 | Darmawan | 74201 | 3.08 | 2.64 | 3.22 | 3 | 3.61 | 3.37 | 3.8 | 3 | 3.15 | 144 | Tepat waktu |
| 15416274201100 | Nurul Ichwan | 74201 | 3.16 | 3.37 | 3.44 | 3.38 | 3 | 1.19 | 3.71 | 3.38 | 2.92 | 144 | Tepat waktu |
| 15416274201099 | Hadyoto | 74201 | 3.3 | 3.25 | 3.29 | 3.33 | 3.39 | 3.79 | 3.9 | 3.33 | 3.39 | 144 | Tepat waktu |
| 15416274201087 | Rizki Maulana | 74201 | 3.4 | 2.6 | 3.31 | 2.74 | 3.72 | 3.37 | 4 | 2.74 | 3.19 | 144 | Tepat waktu |
| 15416274201088 | Sudarmawan | 74201 | 3.26 | 3.03 | 3.22 | 3.26 | 3.36 | 0.39 | 3.66 | 3.36 | 2.77 | 144 | Tepat waktu |
| 15416274201089 | Purni Riyadi | 74201 | 2.83 | 3.46 | 2.74 | 2.63 | 0.82 | 3 | 3 | 2.63 | 2.41 | 144 | Tepat waktu |

Gambar 1. Contoh dataset penelitian

B. Pembersihan Data

Proses pembersihan data tahap pertama memilih jenis data dan atribut yang akan dianalisis dan *cleaning* tahap 2 memilih atribut yang akan dipakai seperti ips 1 sampai dengan IPS 6, IPK, dan kelulusan. Pada tahap *cleaning* tahap 1, atribut yang tidak dipakai oleh peneliti dieleminasi. Hasil dari *cleaning* tahap 1 sebanyak 1165 data mahasiswa dengan atribut nomor NIM, nama, program studi, IPS, IPK, jumlah SKS, kelulusan. Sedangkan hasil dari hasil *cleaning* tahap 2 berjumlah 1157 data mahasiswa. Sampel data dapat dilihat di Gambar 2 dan Gambar 3.

| A | B | C | D | E | F | G | H | I | J | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------------------|-------------------|-------------------|
| IPS 1 | IPS 2 | IPS 3 | IPS 4 | IPS 5 | IPS 6 | IPS 7 | IPS 8 | IPK | Kelulusan | |
| 3.66 | 3.51 | 3.63 | 3.59 | 3.34 | 3.56 | 3.95 | 3.59 | 3.55 | Tepat Waktu | |
| 3.3 | 3.55 | 3.25 | 0 | 0 | 0 | 0 | 0 | 1.68 | Tepat Waktu | |
| 3.05 | 2.3 | 1.72 | 0 | 0 | 0 | 0 | 0 | 1.18 | Tidak Tepat Waktu | |
| 3.18 | 2.55 | 3.03 | 2.42 | 2 | 2.86 | 3.51 | 2.42 | 2.67 | Tepat Waktu | |
| 3.17 | 3.1 | 3.39 | 3.07 | 0 | 0 | 0 | 0 | 3.07 | 2.12 | Tidak Tepat Waktu |
| 2.92 | 2.86 | 3.24 | 2.64 | 2.83 | 3 | 3.52 | 2.64 | 2.92 | Tepat Waktu | |
| 3.24 | 1.62 | 0 | 0 | 0 | 0 | 0 | 0 | 0.81 | Tepat Waktu | |
| 3.08 | 1.71 | 1.65 | 1.05 | 2.75 | 2.76 | 3.6 | 1.05 | 2.17 | Tidak Tepat Waktu | |
| 3.13 | 2.72 | 3.11 | 2.71 | 3.07 | 2.61 | 0 | 2.71 | 2.89 | Tepat Waktu | |
| 3.05 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0.84 | Tidak Tepat Waktu | |
| 3.5 | 1.24 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0.79 | Tidak Tepat Waktu |
| 3.19 | 2.8 | 2.89 | 2.68 | 0 | 1.33 | 0 | 2.68 | 2.15 | Tidak Tepat Waktu | |
| 2.65 | 1.69 | 1.68 | 2.4 | 0.79 | 0 | 0 | 2.4 | 1.54 | Tidak Tepat Waktu | |
| 3.44 | 2.58 | 3.08 | 2.84 | 0 | 0 | 3 | 2.84 | 1.99 | Tidak Tepat Waktu | |
| 3.28 | 2.86 | 3.19 | 2.5 | 0 | 1.51 | 3.38 | 2.5 | 2.22 | Tidak Tepat Waktu | |
| 3.23 | 2.78 | 3.28 | 3.18 | 3.2 | 3.47 | 3.75 | 3.18 | 3.19 | Tepat Waktu | |
| 3.01 | 3.05 | 3.08 | 2.42 | 0 | 3.92 | 2.42 | 1.93 | Tidak Tepat Waktu | | |
| 3.33 | 2.91 | 0.88 | 0 | 3.19 | 3.53 | 3.84 | 0 | 2.31 | Tidak Tepat Waktu | |
| 3.26 | 2.9 | 3.04 | 2.88 | 3.19 | 3.33 | 3.61 | 2.88 | 3.10 | Tepat Waktu | |
| 3.26 | 3.42 | 3.49 | 3.24 | 0 | 0 | 0 | 3.24 | 2.25 | Tidak Tepat Waktu | |
| 3.47 | 2.98 | 3.14 | 0.25 | 0 | 0 | 0 | 0.25 | 1.64 | Tidak Tepat Waktu | |
| 3 | 3.09 | 2.94 | 2.15 | 2.13 | 1.81 | 4 | 2.15 | 2.52 | Tidak Tepat Waktu | |
| 2.63 | 2.36 | 2.7 | 1.61 | 3.4 | 3.39 | 3.37 | 1.61 | 2.68 | Tidak Tepat Waktu | |
| 3.15 | 2.7 | 2.74 | 2.2 | 2.74 | 2.72 | 3.37 | 2.2 | 2.71 | Tidak Tepat Waktu | |
| 2.6 | 0.4 | 0 | 0 | 3.17 | 3.43 | 3.93 | 0 | 1.60 | Tidak Tepat Waktu | |
| 3.21 | 3.51 | 3.46 | 3.65 | 3.39 | 3.62 | 3.85 | 3.65 | 3.51 | Tepat Waktu | |

Gambar 2. Pembersihan data tahap pertama

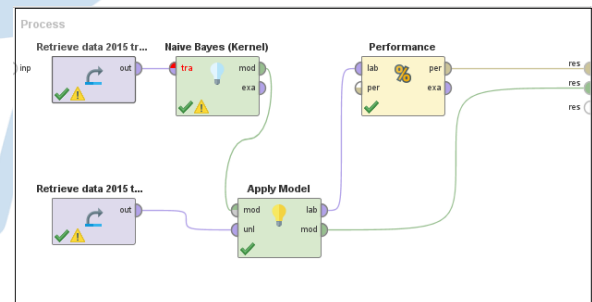
| A | B | C | D | E | F | G | H | I | J | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------------------|-------------------|-------------------|
| IPS 1 | IPS 2 | IPS 3 | IPS 4 | IPS 5 | IPS 6 | IPS 7 | IPS 8 | IPK | Kelulusan | |
| 3.66 | 3.51 | 3.63 | 3.59 | 3.34 | 3.56 | 3.95 | 3.59 | 3.55 | Tepat Waktu | |
| 3.3 | 3.55 | 3.25 | 0 | 0 | 0 | 0 | 0 | 1.68 | Tepat Waktu | |
| 3.05 | 2.3 | 1.72 | 0 | 0 | 0 | 0 | 0 | 1.18 | Tidak Tepat Waktu | |
| 3.18 | 2.55 | 3.03 | 2.42 | 2 | 2.86 | 3.51 | 2.42 | 2.67 | Tepat Waktu | |
| 3.17 | 3.1 | 3.39 | 3.07 | 0 | 0 | 0 | 0 | 3.07 | 2.12 | Tidak Tepat Waktu |
| 2.92 | 2.86 | 3.24 | 2.64 | 2.83 | 3 | 3.52 | 2.64 | 2.92 | Tepat Waktu | |
| 3.24 | 1.62 | 0 | 0 | 0 | 0 | 0 | 0 | 0.81 | Tepat Waktu | |
| 3.08 | 1.71 | 1.65 | 1.05 | 2.75 | 2.76 | 3.6 | 1.05 | 2.17 | Tidak Tepat Waktu | |
| 3.13 | 2.72 | 3.11 | 2.71 | 3.07 | 2.61 | 0 | 2.71 | 2.89 | Tepat Waktu | |
| 3.05 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0.84 | Tidak Tepat Waktu | |
| 3.5 | 1.24 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0.79 | Tidak Tepat Waktu |
| 3.19 | 2.8 | 2.89 | 2.68 | 0 | 1.33 | 0 | 2.68 | 2.15 | Tidak Tepat Waktu | |
| 2.65 | 1.69 | 1.68 | 2.4 | 0.79 | 0 | 0 | 2.4 | 1.54 | Tidak Tepat Waktu | |
| 3.44 | 2.58 | 3.08 | 2.84 | 0 | 0 | 3 | 2.84 | 1.99 | Tidak Tepat Waktu | |
| 3.28 | 2.86 | 3.19 | 2.5 | 0 | 1.51 | 3.38 | 2.5 | 2.22 | Tidak Tepat Waktu | |
| 3.23 | 2.78 | 3.28 | 3.18 | 3.2 | 3.47 | 3.75 | 3.18 | 3.19 | Tepat Waktu | |
| 3.01 | 3.05 | 3.08 | 2.42 | 0 | 3.92 | 2.42 | 1.93 | Tidak Tepat Waktu | | |
| 3.33 | 2.91 | 0.88 | 0 | 3.19 | 3.53 | 3.84 | 0 | 2.31 | Tidak Tepat Waktu | |
| 3.26 | 2.9 | 3.04 | 2.88 | 3.19 | 3.33 | 3.61 | 2.88 | 3.10 | Tepat Waktu | |
| 3.26 | 3.42 | 3.49 | 3.24 | 0 | 0 | 0 | 3.24 | 2.25 | Tidak Tepat Waktu | |
| 3.47 | 2.98 | 3.14 | 0.25 | 0 | 0 | 0 | 0.25 | 1.64 | Tidak Tepat Waktu | |
| 3 | 3.09 | 2.94 | 2.15 | 2.13 | 1.81 | 4 | 2.15 | 2.52 | Tidak Tepat Waktu | |
| 2.63 | 2.36 | 2.7 | 1.61 | 3.4 | 3.39 | 3.37 | 1.61 | 2.68 | Tidak Tepat Waktu | |
| 3.15 | 2.7 | 2.74 | 2.2 | 2.74 | 2.72 | 3.37 | 2.2 | 2.71 | Tidak Tepat Waktu | |
| 2.6 | 0.4 | 0 | 0 | 3.17 | 3.43 | 3.93 | 0 | 1.60 | Tidak Tepat Waktu | |
| 3.21 | 3.51 | 3.46 | 3.65 | 3.39 | 3.62 | 3.85 | 3.65 | 3.51 | Tepat Waktu | |

Gambar 3. Pembersihan data tahap kedua

Pada gambar 3, data siap diolah untuk prediksi kelulusan mahasiswa. Dari 1157 mahasiswa angkatan 2015, peneliti membagi dua data tersebut menjadi 80% data *training*, 20% data *testing* dari data yang sama kelulusan mahasiswa secara acak.

C. Pemodelan Data

Pada tahap ini dilakukan pemodelan data mining dengan aplikasi rapidminer dan menguji keakuratan klasifikasi kelulusan mahasiswa Universitas XYZ dengan menggunakan algoritma *Naïve Bayes*. Pengaturan pada perangkat lunak RapidMiner terlihat di Gambar 4.



Gambar 4. Pemodelan prediksi dengan Naïve Bayes

D. Evaluasi Naïve Bayes

Peneliti melakukan uji akurasi, presisi dan *recall* terhadap data lulusan mahasiswa dengan mendapatkan nilai-nilai sebagaimana terlihat pada Gambar 5.

| accuracy: 80.95% | | | |
|-------------------------|------------------|------------------------|-----------------|
| | true Tepat Waktu | true Tidak Tepat Waktu | class precision |
| pred. Tepat Waktu | 121 | 16 | 88.32% |
| pred. Tidak Tepat Waktu | 28 | 66 | 70.21% |
| class recall | 81.21% | 80.49% | |

Gambar 5. Hasil evaluasi metode Naïve Bayes

Dari hasil prediksi menggunakan rapidminer prediksi kelulusan mahasiswa dan menunjukkan nilai akurasi sebesar 80,95% dengan nilai recall pada rentang 81,21% - 80,49%. Sementara dalam *precision* kelas tepat waktu, mendapatkan nilai pada rentang 70,21% -88,32%. Hasil akurasi yang diperoleh dapat dikategorikan relatif baik. Kontribusi hasil penelitian ini diharapkan dapat menjadi tolak ukur manajemen terutama pengelola program studi dalam menentukan strategi untuk mendapatkan keberhasilan dalam kelulusan tepat waktu.

V. KESIMPULAN DAN SARAN

Berdasarkan hasil pemahaman analisis oleh peneliti, didapat ditarik beberapa kesimpulan, yaitu pertama, jumlah kategori mempengaruhi kinerja klasifikasi data mahasiswa lulusan tahun 2019 menggunakan metode *Naive bayes*. Klasifikasi data mahasiswa lulusan tahun 2019 dengan data 889 data training dari 80% jumlah keseluruhan data mahasiswa dan 231 data testing dari 20% jumlah keseluruhan, diperoleh data mahasiswa benar benar tepat waktu ada sekitar 121 orang, sedangkan yang tidak tepat waktu ada 16 orang.

Dengan nilai akurasi yang didapat dapat disimpulkan bahwa, data prediksi kelulusan tepat waktu para mahasiswa menjadi sangat penting dikarenakan dapat menjadi tolak ukur kinerja instansi. Sehingga penelitian ini dapat dimanfaatkan oleh program studi dalam menentukan strategi mencapai kelulusan tepat waktu dan membantu perguruan tinggi dalam membuat kebijakan kelulusan pada mahasiswa.

Untuk pengembangan penelitian selanjutnya dapat melakukan perbandingan hasil dengan menggunakan metode lain seperti C.4.5 atau KNN untuk mengetahui hasil kelebihan dari masing-masing metode yang digunakan.

DAFTAR PUSTAKA

- [1] M. A. Nurrohmah and Y. S. Nugroho, "khazanah informatika Aplikasi Pemrediksi Masa Studi dan Predikat Kelulusan Mahasiswa Informatika Universitas Muhammadiyah Surakarta Menggunakan Metode *Naive Bayes*," vol. I, no. 1, pp. 29–34, 2015.
- [2] D. L. Fithri, E. Darmanto, P. Studi, S. Informasi, F. Teknik, and U. M. Kudus, "Informasi Fakultas Teknik 1. Universitas Muria Kudus untuk memprediksi kelulusan mahasiswa.," pp. 319–324, 2014.
- [3] J. Jtik, J. Teknologi, and T. Iqbal, "Prediksi Kelulusan Mahasiswa menggunakan Algoritma *Naive Bayes* (Studi Kasus 5 PTS di Banda Aceh)," vol. 3, no. 2, pp. 1–5, 2019.
- [4] Priati, "Penerapan Data Mining Pada Data Transaksi Superstore Untuk Mengetahui Kemungkinan Pelanggan Membeli Product Category Dan Product Container Secara Bersamaan Dengan Teknik Asosiasi Menggunakan Algoritma *Apriori*," *Techno Xplore J. Ilmu Komput. dan Teknol. Inf.*, vol. 1, no. 2, pp. 11–18, 2017, doi: 10.36805/technoexplore.v1i2.106.
- [5] A. Trimanto, F. Faqih, I. M. Irfani, and S. Timur, "Penerapan Data Mining Untuk Evaluasi Status Kelulusan Mahasiswa Fakultas Teknologi Pertanian Tahun 2015 Menggunakan Algoritma *Naive Bayes Classifier*," 2015.
- [6] Bustami, "Penerapan Algoritma *Naive Bayes*," *J. Inform.*, vol. 8, no. 1, pp. 884–898, 2014.
- [7] S. Salmu, "Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu Menggunakan *Naive Bayes*: Studi Kasus UIN Syarif Hidayatullah Jakarta Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu Menggunakan *Naive Bayes*: Studi Kasus UIN Syarif Hidayatullah Jakarta Prediction of Tim," no. April, 2017.
- [8] A. A. Murtopo, "Prediksi Kelulusan Tepat Waktu Mahasiswa STMIK YMI Tegal Menggunakan Algoritma *Naive Bayes Time Graduation Prediction by Using Naive Bayes Algorithm at STMIK YMI Tegal*," pp. 145–154.
- [9] M. F. Nugroho and S. Wibowo, "Fitur Seleksi Forward Selection Untuk Menentukan Atribut Yang Berpengaruh Pada Klasifikasi Kelulusan Mahasiswa Fakultas Ilmu Komputer UNAKI Semarang Menggunakan Algoritma *Naive Bayes*," vol. 3, no. 1, pp. 63–70, 2017.