

PENERAPAN ALGORITMA *NAIVE BAYES CLASSIFIER* UNTUK EVALUASI KINERJA AKADEMIK MAHASISWA UNIVERSITAS BINA DARMA

Muhammad Nasir¹⁾, Verawaty²⁾

Fakultas Ilmu Komputer Universitas Bina Darma Palembang
e-mail : nasir.binadarma.ac.id¹⁾, verawaty2608@gmail.com²⁾

ABSTRAK

Kinerja akademik mahasiswa merupakan hal yang sangat penting untuk di evaluasi sehingga dengan evaluasi tersebut dapat menemukan faktor-faktor yang mendukung peningkatan kinerja akademik. Pada penelitian ini dilakukan evaluasi kinerja akademik mahasiswa pada tahun 2010-2014. Dari data yang didapatkan penulis mengklasifikasi data menjadi beberapa kategori yaitu kinerja akademik tinggi, sedang, rendah dan sangat rendah, serta kategori mahasiswa yang lulus tepat waktu atau tidak tepat waktu. Kemudian dari klasifikasi tersebut dapat digunakan sebagai solusi untuk menentukan kinerja akademik yang optimal. Evaluasi dilakukan dengan menerapkan proses data mining dengan algoritma Naive Bayes Classifier untuk membuat tabel probabilitas sebagai dasar proses klasifikasi, atribut yang digunakan dalam penelitian ini : nomor induk mahasiswa, nama, jenis kelamin, asal sekolah, tempat tanggal lahir, kota asal, program studi, indeks prestasi semester, indeks prestasi kumulatif, target kelulusan.

Kata Kunci : *Kinerja Akademik, Data Mining, Naive Bayes Classifier.*

ABSTRACT

Student academic performance is a very important thing to evaluate so that the evaluation can find factors that support improving academic performance. In this study an evaluation of the academic performance of students in 2010-2014. From the data obtained the authors classify the data into several categories, namely high, medium, low and very low academic performance, and categories of students who graduate on time or not on time. Then from the classification can be used as a solution to determine optimal academic performance. Evaluation is done by applying the data mining process with the Naive Bayes Classifier algorithm to create a probability table as the basis of the classification process, the attributes used in this study: Student's parent number, name, gender, school of origin, place of birth date, hometown, study program, temporary achievement index, cumulative grade point, graduation target.

Keywords: *Academic Performance, Data Mining, Naive Bayes Classifier.*

I. PENDAHULUAN

Perguruan tinggi merupakan penyelenggara pendidikan akademik bagi mahasiswa, yang di sebutkan dalam peraturan Pemerintah Republik Indonesia nomor 66 tahun 2010 tentang perubahan atas peraturan pemerintah nomor 17 tahun 2010 tentang pengelolaan dan penyelenggaraan pendidikan, perkembangan perguruan tinggi di Indonesia saat ini semakin meningkat. Hal ini dapat dilihat dengan banyak munculnya perguruan-perguruan tinggi baru dari tahun ke tahun baik dalam bentuk universitas, akademi, institut, politeknik, dan sekolah tinggi yang berstatus negeri maupun swasta. Data dari PDDIKTI menyebutkan bahwa pada tahun 2018 terdapat 437 Perguruan Tinggi Negeri (PTN) di Indonesia, sedangkan untuk Perguruan Tinggi Swasta (PTS) tercatat sebanyak 4.253 perguruan tinggi (<http://forlap.ristekdikti.go.id>) [1], semakin bertambah jumlah perguruan tinggi maka semakin meningkat pula jumlah sumber daya manusia berkualitas yang dihasilkan perguruan tinggi.

Salah satu faktor yang menentukan kualitas perguruan tinggi adalah persentase kemampuan mahasiswa untuk menyelesaikan studi tepat waktu. Berdasarkan matriks penilaian instrument akreditasi program studi Badan Akreditasi Nasional Perguruan Tinggi bahwa persentase mahasiswa yang lulus tepat waktu merupakan salah satu elemen penilaian akreditasi universitas dan Mahasiswa merupakan salah satu aspek penting dalam evaluasi keberhasilan penyelenggaraan program studi pada suatu perguruan tinggi. Pemantauan mahasiswa yang masuk, peningkatan kemampuan mahasiswa, prestasi yang dicapai mahasiswa, rasio kelulusan terhadap jumlah total mahasiswa, dan kompetensi lulusan perguruan tinggi bisa mendapatkan perhatian yang serius untuk memperoleh kepercayaan *stakeholder* dalam menilai dan menetapkan penggunaan lulusannya.

Penelitian tentang kinerja akademik mahasiswa adalah topik penting dalam dunia pendidikan. Menemukan faktor-faktor yang berpengaruh terhadap kinerja akademik mahasiswa adalah hal yang sangat penting bagi universitas, dosen dan dalam beberapa hal untuk mahasiswa sendiri. Faktor tersebut akan sangat berpengaruh terhadap kebijakan akademik universitas, perbaikan terhadap kurikulum, penilaian kinerja dosen dan modifikasi cara dosen dalam mengajar (Guney 2009) [5].

Universitas Bina Darma Palembang adalah salah satu perguruan tinggi swasta di Sumatera Selatan, visi dari Universitas Bina Darma yaitu menjadi Universitas berstandar *international* berbasis teknologi informasi pada tahun 2025 dan dua dari misi

Universitas Bina Darma ialah “menyelenggarakan program pendidikan yang berstandar internasional dan menyelenggarakan proses pembelajaran yang berstandar internasional melalui pemanfaatan teknologi informasi”. Standar pendidikan disini terkait langsung dengan kemampuan mahasiswa yang merupakan aspek penting dalam menentukan keberhasilan penyelenggaraan program studi pada suatu perguruan tinggi. selain itu salah satu indikator dari keberhasilan suatu perguruan tinggi dapat di lihat dari kinerja akademik mahasiswa.

Setiap tahunnya selalu saja kinerja akademik mahasiswa menjadi topik yang harus dicari cara bagaimana menemukan faktor untuk meningkatkannya agar tidak ada lagi mahasiswa yang berstatus tanpa kabar, mahasiswa terhambat dalam menyelesaikan studinya maupun mahasiswa yang seringkali melakukan registrasi administrasi tetapi tidak pernah mengikuti perkuliahan. Beberapa penyebab yang mungkin mempengaruhi kinerja akademik mahasiswa diantaranya rendahnya kemampuan akademik, faktor pembiayaan, domisili saat menempuh studi dan faktor lainnya, dengan banyaknya permasalahan yang terjadi pada kinerja akademik di Universitas Bina Darma pasti berujung pada persentasi tingkat ketepatan kelulusan mahasiswa, semakin tingginya persentasi mahasiswa lulus tidak tepat waktu maka akan berakibat pada salah satu elemen penilaian akreditasi yang ditetapkan oleh Badan Akreditasi Nasional.

Sementara itu dibutuhkan solusi untuk mengevaluasi dalam kinerja akademik mahasiswa untuk mengklasifikasikan permasalahan yang timbul dikalangan mahasiswa dengan menggunakan atribut data yang dipilih dari data akademik mahasiswa dari tahun 2010 sampai dengan 2014 pada Program Studi Teknik Informatika, nomor induk mahasiswa, nama, jenis kelamin, asal sekolah, tempat tanggal lahir, kota asal, program studi, indeks prestasi sementara, indeks prestasi kumulatif, target kelulusan. Oleh karena itu perlu di terapkan Teknik *data mining* untuk mengklasifikasikan data serta memprediksi hasil evaluasi, dan untuk mengukur keakuratan akurasi klasifikasi data mining dapat diukur dengan menggunakan *Algoritma Naive Bayes Classifier*. *Algoritma Naive Bayes Classifier* yang merupakan sebuah pengklasifikasi probabilitas sederhana yang mengaplikasikan Teorema Bayes. Ide dasar dari Teorema Bayes adalah menangani masalah yang bersifat hipotesis yakni mendesain suatu klasifikasi untuk memisahkan objek [2]. Oleh karena itu dengan metode *Algoritma Naive Bayes Classifier* digunakan untuk mempermudah dalam proses pengklasifikasian evaluasi kinerja akademik mahasiswa.

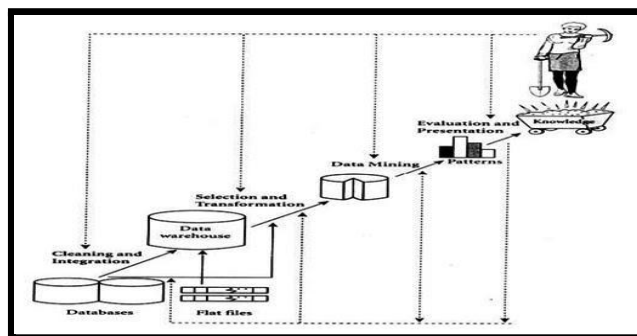
Berdasarkan latar belakang yang telah diuraikan diatas maka rumusan masalah pada penelitian ini adalah : mengevaluasi kinerja akademik mahasiswa Universitas Bina Darma dengan menerapkan algoritma *naive bayes classifier*. Permasalahan yang dibahas pada penelitian ini memiliki ruang lingkup yang cukup luas, sehingga permasalahannya dibatasi tentang bagaimana mengevaluasi kinerja akademik mahasiswa Universitas Bina Darma, memprediksi ketepatan waktu kelulusan mahasiswa menggunakan algoritma *naive bayes classifier* dengan menggunakan sampel data akademik mahasiswa tahun 2010 sampai dengan 2014

Tujuan yang ingin dicapai pada penelitian ini adalah untuk menganalisis permasalahan yang di alami mahasiswa Universitas Bina Darma Palembang dalam bidang akademik, mengetahui faktor yang menyebabkan mahasiswa yang sudah registrasi tapi tidak pernah mengikuti kelas perkuliahan, untuk mengetahui apa penyebab banyaknya mahasiswa yang berstatus tanpa kabar, serta untuk memprediksi penyebab sering terjadinya atau terhambatnya kelulusan mahasiswa tepat waktu dan mengklasifikasi tingkat kinerja dalam kategori tinggi, sedang dan rendah.

II. METODE

A. Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machinelearning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. Istilah *datamining* memiliki hakikat sebagai disiplin ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki. *Data mining*, sering juga disebut sebagai *Knowledge Discovery in Database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar[2].



Gambar 1 Tahap-tahap Data Mining[2]

Tahap-tahap *Data Mining* Sebagai suatu rangkaian proses, *data mining* dapat dibagi menjadi beberapa tahap proses yang diilustrasikan pada Gambar 1. Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantara *knowledge base*. Tahapan *data mining* adalah sebagai berikut:

1. Pembersihan data (*data cleaning*)
Pembersihan data merupakan proses menghilangkan-kan *noise* dan data yang tidak konsisten atau data tidak relevan.
2. Integrasi data (*data integration*)
Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru.
3. Seleksi data (*data selection*)
Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*.
4. Transformasi data (*data transformation*)
Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*.
5. Proses *mining*
Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.
6. Evaluasi pola (*pattern evaluation*)
Untuk mengidentifikasi pola-pola menarik ke dalam *knowledge based* yang ditemukan.
7. Presentasi pengetahuan (*knowledge presentation*)
Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.

B. Naive Bayes Classifier (NBC)

Menurut Santosa (2007) dalam buku *data mining teori dan aplikasi Rapid miner* (2017) “teori keputusan *bayes* adalah pendekatan statistika yang *fundamental* dalam *data mining*. Pendekatan ini didasarkan pada kuantifikasi *trade-off* antara berbagai keputusan klasifikasi dengan menggunakan probabilitas .

Menurut Suyanto (2017) klasifikasi *bayes* menggunakan metode teorema *bayes* di abad ke-18. Dalam teorema *bayes*, probabilitas atau peluang bersyarat dinyatakan sebagai berikut :

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

Jadi *naive bayes* atau sering dikenal dengan nama lain teorema *bayes* maupun klasifikasi *bayes* merupakan pendekatan statistika yang *fundamental* dalam *data mining*. Pendekatan yang didasarkan pada kuantifikasi *trade-off* antara berbagai keputusan klasifikasi dengan menggunakan probabilitas atau peluang dengan pernyataan sebagai berikut :

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

Keterangan :

P (H|X) : Probabilitas
H : Hipotesis
X : Bukti

C. RapidMiner

RapidMiner adalah *platform* perangkat lunak untuk tim ilmu data yang menyatukan persiapan data, pembelajaran mesin dan penyebaran model prediktif[4]. *Rapidminer* digunakan untuk melakukan ekstrak dari data set yang digunakan dengan mengkombinasikannya ke basis data, perangkat lunak ini digunakan untuk mengelola data mining dalam penelitian termasuk proses ekstraksi data, transformasi data dan penyimpanan data.

D. Data Penelitian

Data yang digunakan adalah data akademik mahasiswa dari tahun 2010 - 2014 di Universitas Bina Darma Palembang. Terdiri dari data *attribut nomor induk mahasiswa, nama, jenis kelamin, asal sekolah, tempat tanggal lahir, kota asal, program studi, indeks prestasi sementara, indeks prestasi kumulatif, target kelulusan*.

III. HASIL

A. Analisis Data

Data awal atau data mentah yang didapat dari pengambilan data pada Universitas Bina Darma sebelum dilakukannya *preprocessing* data, dengan data keseluruhan yang dapat diakses kembali Universitas Bina Darma untuk tahun 2010 sebanyak 360 data Mahasiswa Teknik Informatika, tahun 2011 sebanyak 403 data Mahasiswa Teknik Informatika, tahun 2012 sebanyak 329 data Mahasiswa Teknik Informatika, tahun 2013 sebanyak 392 data Mahasiswa Teknik Informatika, sedangkan untuk tahun 2014 sebanyak 393 data Mahasiswa Teknik informatika, jadi jumlah seluruh data yang digunakan pada penelitian ini sebanyak 1877 data Mahasiswa Teknik Informatika.

Tabel 1. Dataset

NIM	Semester	IPK Akhir	Status	Kinerja	Target
101423xx	10	3.220.000	AKTIF	Sedang	Tepat Waktu
101422xx	10	3.140.000	AKTIF	Sedang	Tepat Waktu
101422xx	10	3.640.000	AKTIF	Tinggi	Tepat Waktu
101423xx	10	3.390.000	AKTIF	Sedang	Tepat Waktu
101422xx	10	3.140.000	AKTIF	Sedang	Tepat Waktu
101423xx	10	3.030.000	AKTIF	Sedang	Tepat Waktu
101422xx	10	3.250.000	AKTIF	Sedang	Tepat Waktu
101422xx	5	3.480.000	WISUDA	Sedang	Tepat Waktu
101422xx	5	3.460.000	WISUDA	Sedang	Tepat Waktu
1014223x					
x	5	3.210.000	WISUDA	Sedang	Tepat Waktu
101422xx	6	3.040.000	WISUDA	Sedang	Tepat Waktu
101422xx	6	3.360.000	WISUDA	Sedang	Tepat Waktu
101421xx	6	3.030.000	WISUDA	Sedang	Tepat Waktu
101422xx	7	3.470.000	WISUDA	Sedang	Tepat Waktu
101422xx	7	3.420.000	WISUDA	Sedang	Tepat Waktu
1014220x					
x	7	3.610.000	WISUDA	Tinggi	Tepat Waktu
101422xx	7	3.530.000	WISUDA	Tinggi	Tepat Waktu
101423xx	7	3.160.000	WISUDA	Sedang	Tepat Waktu
101422xx	7	3.640.000	WISUDA	Tinggi	Tepat Waktu
1014225x					
x	7	3.610.000	WISUDA	Tinggi	Tepat Waktu
101421xx	7	3.720.000	WISUDA	Tinggi	Tepat Waktu
101423xx	7	3.310.000	WISUDA	Sedang	Tepat Waktu

Selanjutnya diproses menggunakan algoritma *naive bayes classifier* dengan urutan langkah berikut. (1) *data cleaning* (2) *data integration* (3) *data selection* (4) *data transformation* (5) *Proses mining* (6) *pattern evaluation* (7) *knowledge presentation*

B. *Rule Target Akademik Mahasiswa*

Sebelum melakukan perhitungan manual, ditentukan dulu target dari setiap Indeks Prestasi Kumulatif dan kinerja dengan melakukan perhitungan dan membuat aturan untuk mendapatkan tingkat kinerja, maka aturan *rules* yang terbentuk adalah sebagai berikut :

1. *If* IPK = 4.00 – 3.51 *then* Kinerja = Tinggi
2. *If* IPK = 3.50 – 3.01 *then* Kinerja = Sedang
3. *If* IPK = 3.00 – 2.75 *then* Kinerja = Rendah
4. *If* IPK = 2.74 – 0.00 *then* Kinerja = Sangat Rendah
5. *If* Semester = 8 *and* Kinerja Tinggi *then* Target = Tepat Waktu
6. *If* Semester = 8 *and* Kinerja Sedang *then* Target = Tepat Waktu
7. *If* Semester > 8 *and* Kinerja Rendah *then* Target = Tidak Tepat Waktu
8. *If* Semester > 8 *and* Kinerja Sangat Rendah *then* Target = Tidak Tepat Waktu

C. Menghitung P(Ci) Untuk Setiap Kelas

Menghitung jumlah kelas dari target berdasarkan klasifikasi yang terbentuk (*prior probability*), dimana jumlah tiap kelas masing-masing target yang ada dibagi dengan jumlah seluruh data. C1 untuk kelas target tepat waktu, C2 untuk kelas tidak tepat waktu :

1. $C1(\text{Class Target} = \text{'Tepat Waktu'}) = \frac{\text{jumlah "tepat waktu" pada kolom Target}}{112/200} = 0,56$
2. $C2(\text{Class Estimasi} = \text{'Tidak Tepat Waktu'}) = \frac{\text{jumlah "tidak tepat waktu" pada kolom Target}}{88/200} = 0,44$

D. Perhitungan P(X|Ci) Setiap Kasus

Menghitung jumlah kasus yang sama pada setiap atribut dari kelas target (tepat waktu, tidak tepat waktu) berdasarkan masing-masing nilai IPK Akhir yang terpilih secara acak (1.00, 2.50, 2.55, 2.67, 2.74, 2.75, 2.80, 2.90, 2.93, 2.97, 3.00, 3.01, 3.10, 3.20, 3.25, 3.27, 3.34, 3.35, 3.38, 3.40, 3.45, 3.48, 3.50, 3.54, 3.60, 3.64, 3.70, 3.73, 3.80) kemudian dibagi dengan seluruh jumlah atribut IPK Akhir masing-masing.

1. $P(\text{IPK} = "1.00" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 7/7 = 1$
2. $P(\text{IPK} = "2.50" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 3/3 = 1$
3. $P(\text{IPK} = "2.55" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 3/3 = 1$
4. $P(\text{IPK} = "2.67" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 6/6 = 1$
5. $P(\text{IPK} = "2.74" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 6/6 = 1$
6. $P(\text{IPK} = "2.75" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 5/5 = 1$
7. $P(\text{IPK} = "2.80" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 2/2 = 1$
8. $P(\text{IPK} = "2.90" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 10/10 = 1$
9. $P(\text{IPK} = "2.93" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 8/8 = 1$
10. $P(\text{IPK} = "2.97" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 8/8 = 1$
11. $P(\text{IPK} = "3.00" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 20/20 = 1$
12. $P(\text{IPK} = "3.01" \mid \text{Class Target} = \text{"tepat waktu"}) = 6/8 = 0.75$
13. $P(\text{IPK} = "3.01" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 2/8 = 0.25$
14. $P(\text{IPK} = "3.10" \mid \text{Class Target} = \text{"tepat waktu"}) = 4/9 = 0.44$
15. $P(\text{IPK} = "3.10" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 5/9 = 0.55$
16. $P(\text{IPK} = "3.20" \mid \text{Class Target} = \text{"tepat waktu"}) = 5/7 = 0.71$
17. $P(\text{IPK} = "3.20" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 2/7 = 0.28$
18. $P(\text{IPK} = "3.25" \mid \text{Class Target} = \text{"tepat waktu"}) = 13/17 = 0.76$
19. $P(\text{IPK} = "3.25" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 4/17 = 0.23$
20. $P(\text{IPK} = "3.27" \mid \text{Class Target} = \text{"tepat waktu"}) = 6/7 = 0.85$
21. $P(\text{IPK} = "3.27" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 1/7 = 0.14$
22. $P(\text{IPK} = "3.34" \mid \text{Class Target} = \text{"tepat waktu"}) = 11/11 = 1$
23. $P(\text{IPK} = "3.35" \mid \text{Class Target} = \text{"tepat waktu"}) = 1/1 = 1$
24. $P(\text{IPK} = "3.38" \mid \text{Class Target} = \text{"tepat waktu"}) = 9/10 = 0.9$
25. $P(\text{IPK} = "3.38" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 1/10 = 0.1$
26. $P(\text{IPK} = "3.40" \mid \text{Class Target} = \text{"tepat waktu"}) = 4/4 = 1$
27. $P(\text{IPK} = "3.45" \mid \text{Class Target} = \text{"tepat waktu"}) = 11/11 = 1$
28. $P(\text{IPK} = "3.48" \mid \text{Class Target} = \text{"tepat waktu"}) = 3/4 = 0.75$
29. $P(\text{IPK} = "3.48" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 1/4 = 0.25$
30. $P(\text{IPK} = "3.50" \mid \text{Class Target} = \text{"tepat waktu"}) = 11/11 = 1$
31. $P(\text{IPK} = "3.54" \mid \text{Class Target} = \text{"tepat waktu"}) = 9/10 = 0.9$
32. $P(\text{IPK} = "3.54" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 1/10 = 0.1$
33. $P(\text{IPK} = "3.60" \mid \text{Class Target} = \text{"tepat waktu"}) = 1/1 = 1$
34. $P(\text{IPK} = "3.64" \mid \text{Class Target} = \text{"tepat waktu"}) = 6/7 = 0.85$
35. $P(\text{IPK} = "3.64" \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 1/6 = 0.14$
36. $P(\text{IPK} = "3.70" \mid \text{Class Target} = \text{"tepat waktu"}) = 9/9 = 1$
37. $P(\text{IPK} = "3.73" \mid \text{Class Target} = \text{"tepat waktu"}) = 1/1 = 1$
38. $P(\text{IPK} = "3.80" \mid \text{Class Target} = \text{"tepat waktu"}) = 2/2 = 1$

E. Perhitungan Perkalian Semua Variabel

Kalikan Semua Hasil *Variable* yang didapat dari hasil $P(\text{IPK} \mid \text{Class Target})$ dengan sesama kelas targetnya (C1, C2) , dapat dilihat seperti dibawah ini :

a. Untuk semua atribut *Class Target* = "tepat waktu"

$$P(X \mid \text{Class Target} = \text{"tepat waktu"}) = 0.75 \times 0.44 \times 0.71 \times 0.76 \times 0.85 \times 1 \times 1 \times 0.9 \times 1 \times 1 \times 0.75 \times 1 \times 0.9 \times 1 \times 0.85 \times 1 \times 1 \times 1 = 0.078157$$

b. Untuk semua atribut *Class Target* = "tidak tepat waktu"

$$P(X \mid \text{Class Target} = \text{"tidak tepat waktu"}) = 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 0.25 \times 0.55 \times 0.28 \times 0.23 \times 0.14 \times 0.1 \times 0.25 \times 0.1 \times 0.14 = 0.004339$$

F. Perhitungan $P(X \mid C_i) \cdot P(C_i)$

Kalikan hasil yang didapat dari $P(\text{IPK} \mid \text{kelas Target})$ dengan hasil yang didapat dari hasil perkalian semua *variable* (C1, C2) :

1. Perkalian *prior probability* dengan semua atribut *Class Target* = "tepat waktu"

a. $P(C_i) \mid \text{Class target} = \text{"tepat waktu"}) \times P(X \mid \text{Class target} = \text{"tepat waktu"})$

b. = $0,56 \times 0,078157$

c. = $0,04376792$

2. Perkalian *prior probability* dengan semua atribut *Class Target* = "tidak tepat waktu"

- a. $P(C_i | \text{Class target} = \text{"tidak tepat waktu"}) \times P(X | \text{Class target} = \text{"tidak tepat waktu"})$
- b. $= 0,44 \times 0,004339$
- c. $= 0,00190916$

IV PEMBAHASAN

Dari hasil analisis *data mining* secara manual dan dengan implementasi pada *tools rapidminer* didapatkan hasil evaluasi tingkat kinerja akademik mahasiswa Universitas Bina Darma dalam tingkat kinerja sedang, didapatkan target kelulusan mahasiswa Universitas Bina Darma Palembang rata-rata pertahunnya lulus dengan tepat waktu.

Tabel 2. Hasil Pengujian Kinerja kelas tinggi

Kinerja	Tahun	<i>class NBC</i>	Jumlah	Akurasi
Tinggi	2010	0.093	34	94.22%
	2011	0.116	46	87.67%
	2012	0.085	28	92.89%
	2013	0.109	43	93.03%
	2014	0.152	59	93.33%

Dari tabel 2 diatas didapatkan hasil kelas evaluasi kinerja tertinggi tingkat tinggi dengan hasil pengujian pada tahun 2014 dari jumlah seluruh data *class* tingkat tinggi sebanyak 210 data akademik mahasiswa Universitas Bina Darma.

Tabel 3 Hasil Pengujian Kinerja kelas sedang

Kinerja	Tahun	<i>class NBC</i>	Jumlah	Akurasi
Sedang	2010	0.627	225	94.22%
	2011	0.356	145	87.67%
	2012	0.640	211	92.89%
	2013	0.512	201	93.03%
	2014	0.418	165	93.33%

Dari tabel 3 diatas didapatkan hasil kelas evaluasi kinerja tertinggi tingkat sedang dengan hasil pengujian pada tahun 2010 dari jumlah seluruh data *class* tingkat sedang sebanyak 947 data akademik mahasiswa Universitas Bina Darma.

Tabel 4. Hasil Pengujian Kinerja kelas rendah

Kinerja	Tahun	<i>class NBC</i>	Jumlah	Akurasi
Rendah	2010	0.173	62	94.22%
	2011	0.164	66	87.67%
	2012	0.152	50	92.89%
	2013	0.119	47	93.03%
	2014	0.188	73	93.33%

Dari tabel 4 diatas didapatkan hasil kelas evaluasi kinerja tertinggi tingkat rendah dengan hasil pengujian pada tahun 2014 dari jumlah seluruh data *class* tingkat rendah sebanyak 298 data akademik mahasiswa Universitas Bina Darma.

Tabel 5. Hasil Pengujian Kinerja kelas sangat rendah

Kinerja	Tahun	<i>class NBC</i>	Jumlah	Akurasi
Sangat Rendah	2010	0.107	39	94.22%
	2011	0.363	146	87.67%
	2012	0.123	40	92.89%
	2013	0.259	101	93.03%
	2014	0.242	96	93.33%

Dari tabel 5 diatas didapatkan hasil kelas evaluasi kinerja tertinggi tingkat sangat rendah dengan hasil pengujian pada tahun 2011 dari jumlah seluruh data *class* tingkat sangat rendah sebanyak 422 data akademik mahasiswa Universitas Bina Darma.

Tabel 6. Hasil Pengujian Target kelas tepat waktu

Kinerja	Tahun	<i>class NBC</i>	Jumlah	Akurasi
---------	-------	------------------	--------	---------

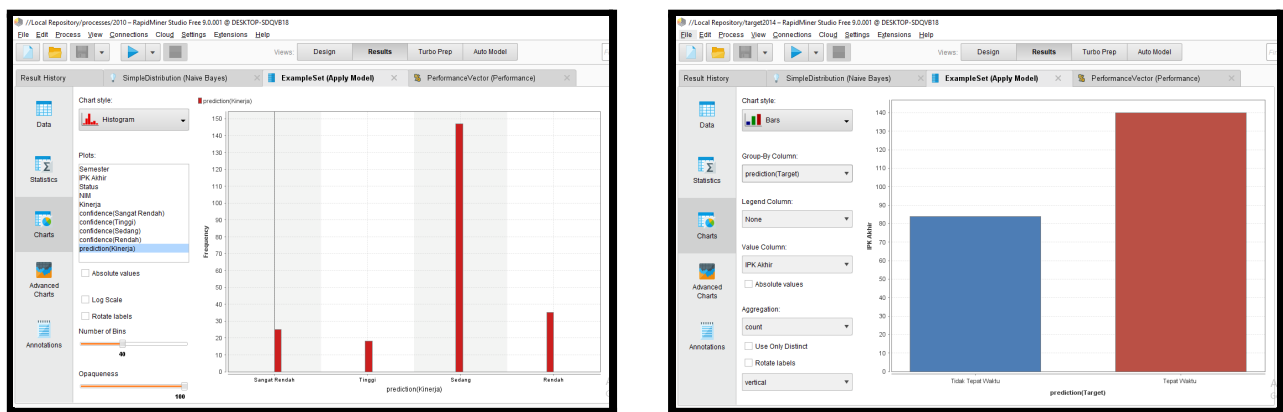
	2010	0.718	259	91.51%
Tepat Waktu	2011	0.473	191	85.63%
	2012	0.728	239	89.54%
	2013	0.623	244	91.39%
	2014	0.571	224	91.07%

Dari tabel6 diatas didapatkan hasil kelas ketepatan waktu kelulusan tertinggi tepat waktu dengan hasil pengujian pada tahun 2010 dari jumlah seluruh data *class* tepat waktu sebanyak 1.157 data akademik mahasiswa Universitas Bina Darma.

Tabel 7. Hasil Pengujian Target *class* tidak tepat waktu

Kinerja	Tahun	<i>class NBC</i>	Jumlah	Akurasi
Tidak Tepat Waktu	2010	0.282	101	94.22%
	2011	0.527	212	87.67%
	2012	0.272	90	92.89%
	2013	0.377	148	93.03%
	2014	0.429	169	93.33%

Dari tabel 7 diatas didapatkan hasil kelas ketepatan waktu kelulusan tertinggi tidak tepat waktu dengan hasil pengujian pada tahun 2011 dari jumlah seluruh data *class* tepat waktu sebanyak 720 data akademik mahasiswa Universitas Bina Darma.



Gambar 2. Grafik hasil penelitian

Dari semua tabel hasil diatas dapat dilihat juga pada grafik di gambar 2 bahwa hasil evaluasi kinerja tertinggi terdapat pada tingkat *class* sedang dengan data sebanyak 947 menempati *calss* evaluasi kinerja tingkat sedang dari seluruh jumlah data sebanyak 1877 data. Sedangkan untuk ketepatan waktu kelulusan mahasiswa dapat dilihat bahwa hasil perhitungan target ketepatan waktu kelulusan mahasiswa tertinggi terdapat pada *class* tepat waktu dengan data sebanyak 1.157 menempati *class* tepat waktu dari seluruh jumlah data sebanyak 1877 data.

V. SIMPULAN DAN SARAN

Berdasarkan penelitian yang penulis lakukan, dapat diambil kesimpulan bahwa dari perhitungan data *mining* menggunakan algoritma *naive bayes* pada teknik *classification* dapat ditarik kesimpulan bahwa kelas kinerja akademik tingkat sedang lebih besar daripada kelas kinerja tinggi, rendah dan sangat rendah sedangkan untuk kelas target kelulusan mahasiswa tingkat lulus tepat waktu lebih besar dari pada kelas tidak tepat waktu. Proses data *training* dapat mempengaruhi hasil pegujian, karena pola data *training* tersebut akan dijadikan sebagi *rule* untuk menentukan kelas pada data *testing*. Sehingga besar atau kecilnya tingkat *presentase accuracy* data dipengaruhi juga oleh penentuan data *training*. Hasil proses data *mining naive bayes classifier* dapat digunakan untuk mengklasifikasi kinerja akademik mahasiswa tahun 2010 sampai dengan tahun 2014 dalam penelitian ini dijadikan data kinerja dan target.

Diharapkan dapat dilakukan penelitian lebih lanjut terkait kinerja akademik mahasiswa, sehingga dapat diketahui faktor-faktor lain yang berpengaruh terhadap kinerja akademik mahasiswa. Kekurangan dari penelitian ini antara lain masih terbatasnya data yang digunakan sehingga dengan menggali data yang lebih banyak dapat diketahui faktor-faktor yang berpengaruh terhadap evaluasi kinerja mahasiswa khususnya factor penghambat studi yang menjadi aspek penting dalam penyelenggaraan program studi pada perguruan tinggi.

REFERENSI

- [1] SRV5 PDDIKTI : Pangkalan Data Pendidikan Tinggi. (t.t.). Diambil 17 Desember 2018, dari <https://forlap.ristekdikti.go.id/perguruan tinggi/homegraphpt>
- [2] Ridwan, M., & Suyono, H. (2013). Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier, 7(1), 6.
- [3] Visi dan Misi. (2016, Maret 24). Diambil 17 Desember 2018, dari <https://www.binadarma.ac.id/visi-dan-misi/>
- [4] Lightning Fast Data Science Platform for Teams | RapidMiner®. (t.t.). Diambil 17 Desember 2018, dari <https://rapidminer.com/>
- [5] Wijaya Langgeng Anggita. (2012). Pengaruh Tingkat Partisipasi Kelas Terhadap Kinerja Akademik Mahasiswa.