

Penerapan Algoritma Naive Bayes Menentukan Klasifikasi Tingkat Kelulusan Siswa SMK Media Informatika Jakarta

Yufika Septiani^{1*}, Pipin Farida Ariyani²

^{1,2}Fakultas Teknologi Informasi, Teknik Informatika, Universitas Budi Luhur, Jakarta Selatan, Indonesia

Email: ^{1*}septianiyufika@gmail.com, ²pipinfaridaariyani@budiluhur.ac.id
(* : corresponding author)

Abstrak-Standar pendidikan sekolah dasar hingga Sekolah Menengah adalah Nilai Ujian Nasional dan Nilai Rapor sekolah, yang menunjukkan kemampuan siswa itu di Sekolah. Siswa mempunyai kewajiban untuk belajar, dalam hal ini harus lebih giat dalam belajar karena akan menghadapi Ujian Sekolah. Salah satu syarat untuk bisa melanjutkan ke jenjang pendidikan ke tingkat selanjutnya sekolah seharusnya mengetahui apa yang menjadi faktor yang menentukan tingkat kelulusan siswanya. Siswa SMK Media Informatika Jakarta saat akan mendaftar menuju ke jenjang selanjutnya selalu terkendala pada nilai yang digunakan sebagai persyaratan. Oleh karena itu diciptakan aplikasi yang akan mengklasifikasikan nilai mereka sehingga saat mendaftar memiliki kesempatan diterima yang lebih besar. Tujuan penelitian ini adalah untuk mengetahui tingkat kelulusan siswa dalam pelaksanaan data mining bagi siswa yang lulus dan tidak lulus, berdasarkan nilai Ujian Sekolah (US). Penelitian ini menggunakan teknik *data mining* dalam memprediksi status kelulusan siswa tersebut. Atribut yang akan digunakan dalam memprediksi kelulusan ini adalah nilai mtk, nilai bindo, nilai bing, dan nilai kejuruan. Data yang digunakan pada pengujian ini adalah data yang dapat dari arsip nilai pihak sekolah sebanyak 200 data. Dalam penelitian ini mendapatkan perhitungan persentase tingkat akurasi pada confusion matrix tingkat akurasi mencapai nilai persentase sebesar 90%. Diperoleh juga *recall* sebesar 94,44%. Dan diperoleh juga tingkat presisi sebesar 94,4%.

Kata Kunci: kelulusan siswa, metode naive bayes, klasifikasi, data mining

Application of The Naive Bayes Algorithm Determining Classification of Students' Graduation Level of Jakarta Media Informatika Vocational School

Abstract-The standard of education from elementary school to high school is the National Examination Score and school report card scores, which show the student's ability in school. Students have an obligation to study, in this case they must be more active in learning because they will face school exams. One of the requirements to be able to continue to the level of education to the next level, the school should know what are the factors that determine the level of graduation of students. Students of SMK Media Informatika Jakarta when registering for the next level are always constrained by the value used as a requirement, therefore an application was created that will classify their scores so that when registering they have a greater chance of being accepted. The purpose of this study was to determine the pass rate of students in the implementation of data mining for students who passed and did not pass, based on the value of the School Exam (US). This study uses data mining techniques in predicting the graduation status of these students. The attributes that will be used in predicting this graduation are mtk scores, bindo scores, bing grades, and vocational grades. The data used in this test is data that can be obtained from the school's value archive as much as 200 data. In this study, the calculation of the percentage level of accuracy in the confusion matrix level of accuracy reaches a percentage value of 90%. A recall of 94.44% was also obtained. And also obtained a level of precision of 94.4%.

Keywords: Student Graduation, Naive Bayes Method, Classification, Data Mining

1. PENDAHULUAN

Pendidikan saat ini memiliki ketentuan dalam mengukur kelulusan siswa ataupun peserta didik, standarisasi adalah bentuk kesesuaian dan komparabilitas yang mengikuti aturan, dalam hal ini pendidikan juga memiliki kesesuaian dan kriteria kelulusan siswa. Standar pendidikan sekolah dasar hingga Sekolah Menengah adalah Nilai Ujian Nasional dan Nilai Rapor sekolah, yang menunjukkan kemampuan siswa itu di Sekolah. Siswa mempunyai kewajiban untuk belajar, dalam hal ini harus lebih giat dalam belajar karena akan menghadapi Ujian Sekolah. Salah satu syarat untuk bisa melanjutkan ke jenjang pendidikan ke tingkat selanjutnya sekolah seharusnya mengetahui apa yang menjadi faktor yang menentukan tingkat kelulusan siswanya. Sistem yang akan dibuat akan membantu pihak sekolah mengetahui kriteria kelulusan dari siswa-siswinya dengan menggunakan data siswa dan data kelulusan, ini berarti bahwa sistem ini kemungkinan akan dapat menganalisis faktor-faktor yang memiliki dampak signifikan terhadap tingkat penyelesaian. Tujuan penelitian ini adalah untuk mengetahui tingkat kelulusan siswa dalam pelaksanaan data mining bagi siswa yang lulus dan tidak lulus, berdasarkan nilai Ujian Sekolah (US). Manfaat pada penelitian ini adalah menghasilkan aplikasi data mining untuk menentukan tingkat kelulusan dengan metode Naive Bayes dan sistem yang mempermudah perhitungan sehingga data yang dihasilkan valid.

Penelitian ini menggunakan metode *Naive Bayes* berbasis web dimana kelebihan program penelitian ini adanya autentikasi email Kata sandi sehingga tidak ada yang bisa mengaksesnya data-data yang terkandung didalam program, namun demikian program pada penelitian ini juga memiliki kekurangannya sendiri yaitu, semakin besar ukuran file *.xls* semakin lama prosesnya.

Data mining adalah proses mengekstraksi data dari informasi yang sangat penting. *Data mining* juga merupakan proses analisis pola dalam data. Pola diambil dari berbagai jenis *database*, termasuk *database* relasional, data *warehouse*, data transaksional, dan data berorientasi objek. *Data mining* memungkinkan pemilik bisnis untuk membuat keputusan dengan cepat dan akurat.

Algoritma Naive Bayes adalah salah satu dari 10 klasifikasi data mining paling umum dari teknik data mining, bersama dengan algoritma lainnya. Metode Naive Bayes juga dikatakan lebih cenderung mengklasifikasikan dokumen dari segi akurasi dan efisiensi komputasi dibandingkan dengan metode klasifikasi lainnya [1].

Pada Kajian Prediksi Kelulusan Berdasarkan Prestasi Akademik Menggunakan Teknik Data Mining Pada Program Studi S1 Sistem Informasi Universitas Brawijaya [3], Masalah pertama adalah Ketidakstabilan data tentu akan merugikan akademis dan mahasiswa. Dengan gunakan salah satu metode , algoritma klasifikasi yaitu *Naive Bayes*, menghasilkan banyak bentuk berdasarkan probabilitas satu item. Hasil pengujian menunjukkan bahwa sistem memiliki akurasi yang membutuhkan dan telah diuji sebelumnya. Hasil uji kegunaan dengan *System Usability Scale* adalah 57,5 persendinilai baik sebagai evaluasi kata sifat.

Kajian Klasifikasi orang miskin dengan metode Naive-Bayes [4], Masalah pertama adalah upaya menangani kemiskinan saat ini terkait dengan fakta bahwa pertumbuhan ekonomi tidak merata..Penelitian ini membuat klasifikasi berdasarkan data penduduk miskin yang diperoleh dari Kecamatan Tibawa dengan menggunakan teknik data mining. Atribut yang digunakan untuk klasifikasi penduduk adalah umur, pendidikan, pekerjaan, pendapatan, tanggungan, dan status (menikah/belum menikah). Metode yang digunakan adalah metode *Naive Bayes Classifier* adalah salah satu teknik klasifikasi dalam data mining. Berdasarkan investigasi yang dilakukan, disimpulkan bahwa sistem klasifikasi miskin di Wilayah Pemerintah Kabupaten Tibawa Kab. Gorontalo mungkin telah dimanipulasi.

Pada penelitian aplikasi opinion mining menggunakan algoritma *naive bayes classifier* pada indihome pt. telkom indonesia [5], Pada proses ini Kedua masalah ini tetap ada dan sangat mengecewakan Ketika data digunakan secara tidak tepat dan berdampak signifikan pada data. Selain perkembangan teknologi yang pesat, banyaknya permintaan pelanggan untuk mengubah data tersebut menjadi data informasi yang berguna kini menjadi kebutuhan yang mendesak. Ini adalah survei data yang tersedia bagi penulis untuk data tweet sentimen tinggi antara 1350 dan 1450.

Pada penelitian Penerapan Metode *Naive Bayes* pada Klasifikasi Pelanggan [6], pada penelitian ini mengurangi beban program perguruan tinggi untuk mencari data mahasiswa yang lulus awal dan akhir. STMIK Bina Nusantara Jaya Lubuklinggau merekomendasikan metode yang lebih akurat untuk menyelesaikan masalah prediksi kelulusan mahasiswa. Ketiga, hasil divalidasi menggunakan teknik *K-Fold Cross Validation*. Akhirnya, kami menggunakan Matriks Kebingungan untuk memvalidasi keakuratan hasil prediksi. Hasilnya ternyata menjadi metode algoritma C4. dapat digunakan untuk memprediksi status kelulusan siswa dengan akurasi 79,08%, sedangkan naive bayes classifier hanya 78,46%. Faktor yang dominan adalah variabel IPK-S4.

Pada penelitian Perbandingan Teknik *Data Mining* Untuk Prediksi Nilai dan Waktu Kelulusan Mahasiswa Ilmu Komputer Menggunakan Algoritma *C4.5*, *Nave Bayes*, *Knn*, dan *SVM* [7], Teknik yang digunakan untuk model klasifikasi data mining ini terdiri dari empat algoritma yaitu *C4.5* yang diuji keefektifannya menggunakan aplikasi rapidminer. Oleh karena itu, hasil dari keempat algoritma tersebut menunjukkan bahwa algoritma Naive Bayes merupakan algoritma yang paling baik untuk memprediksi ketuntasan siswa tepat waktu dan IPK 3 dengan akurasi (76,79%), *error* (23,17%), dan *AUC* (0,850) menunjukkan bahwa .

Pada penelitian Klasifikasikan nasabah asuransi jiwa menggunakan algoritma Naive Bayes berdasarkan eliminasi mundur. [8], Dalam penelitian ini, pendapatan perusahaan asuransi ditentukan oleh besarnya premi yang dibayarkan oleh pelanggannya. Banyaknya nasabah yang tidak mampu membayar premi asuransi dengan lancar mempengaruhi kinerja bisnis dan kelangsungan hidup perusahaan sehari-hari. Algoritma Naive Bayes berbasis *Backward Elimination* bertujuan untuk melakukan klasifikasi nasabah asuransi dengan hasil akurasi 85,89 %.

2. METODE PENELITIAN

2.1 Pengumpulan Data

Penelitian ini, penulis membutuhkan data yang berkaitan dengan topik yang dibahas. Metode pengumpulan data dilakukan dengan tujuan untuk memperoleh data, yaitu:

a. Observasi

Observasi dilakukan untuk meninjau langsung kegiatan dari proses transaksi dan tata letak SMK Media Informatika. Dari yang diamati berdasarkan nilai siswa dari mata pelajaran Bindo, Bing, Mtk, dan Kejuruan

b. Wawancara

Wawancara dilakukan dengan kepala sekolah SMK Media Informatika yang terkait dengan data yang diperlukan untuk memperoleh data dan informasi tentang nilai siswa. Informasi yang didapat akan digunakan sebagai acuan dalam pembuatan sistem aplikasi.

2.2 Knowledge Discovery Database (KDD)

Knowledge discovery in database (KDD) adalah aktivitas yang melibatkan pengumpulan dan penggunaan data historis untuk menemukan pola atau hubungan yang sering terjadi dalam manajemen data dalam jumlah banyak [9].

2.2.1 Data Selection

Proses pemilihan data dilakukan dengan tujuan untuk memilih atribut yang relevan dalam data sehingga dapat dilakukan analisis data [10]. Dari beberapa atribut yang telah didapatkan dari SMK Media Informatika Jakarta, dipilih 8 atribut seperti berikut:

- No Ujian
- Nama Siswa
- Kompetensi Keahlian
- Data nilai Bindo
- Data nilai Bing
- Data nilai Mtk
- Data nilai Jurusan
- Keterangan

2.2.2 Transformation

Transformasi data yaitu mentransformasikan Pilih data yang cocok untuk penambangan data. proses konversi di KDD merupakan proses yang kreatif dan bergantung sifat atau pola informasi yang diambil dari *database*. Pada tahap ini penulis tidak melakukan pada perubahan dikarenakan data yang didapat dari SMK Media Informatika Jakarta sudah sesuai untuk proses *data mining*.

2.2.3 Data Mining

Data yang telah dibersihkan dan dimodifikasi, kemudian diolah untuk dicari informasinya dalam data menggunakan algoritme *naive bayes*. Data disimpan pada *database* ditangani secara otomatis dengan aplikasi Menggunakan metode dan perhitungan algoritme *naive bayes*.

2.3 Tahapan Naive Bayes

Referensi menunjukan pada tahapan dalam melakukan classifier atau klasifikasi dengan metode naive bayes adalah sebagai berikut:

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

Dimana:

- X : Data kelas yang tidak diketahui
H : Hipotesis data adalah kelas tertentu
P(H|X) : Probabilitas hipotesis H dalam kondisi X (probabilitas posterior)
P(H) : Probabilitas hipotesis H (sebelum probabilitas)
P(X|H) : Probabilitas X di bawah kondisi hipotesis H
P (X) : Probabilitas X

2.3 Pengujian

Dalam penelitian ini digunakan uji menggunakan *confusion matrix* dimana ukuran kinerja untuk masalah klasifikasi pembelajaran mesin di mana outputnya adalah dua atau lebih kelas. Pada tabel 1, terdapat rumus *confusion matrix*

Tabel 1. Rumus *Confusion Matrix*

Klasifikasi		Prediksi	
		Yes	No
Aktual	Yes	TP Positif Benar	FN Salah Negatif
	No	FP Salah Positif	TN Benar Negatif

Keterangan :

TP : positif yang klasifikasi benar

TN : negatif yang klarifikasi benar

FP : positif yang klarifikasi salah

FN : negatif tapi klasifikasi salah

Untuk menghitung akurasi, presisi, recall

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{Presisi} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FP} \quad (4)$$

3. HASIL DAN PEMBAHASAN

3.1 Penerapan *Naive Bayes*

Setelah dilakukannya *preprocessing*, Hanya dengan begitu proses Klasifikasi Algoritma *Naive Bayes* dapat berjalan. Proses Algoritma *Naive Bayes* dijalankan menggunakan data latih sebesar 90% dari data nilai yang sudah dilakukan pembersihan, menggunakan data terhitung dari tahun 2021 - 2022 dengan 200 baris data.

3.1.1 Membaca Data Latih

Untuk memutuskan data mana yang akan dianalisis menggunakan metode *Naive Bayes*. Jadi langkah pertama merupakan membaca data pelatihan. Data pelatihan yang digunakan ditunjukkan Pada tabel 2, terdapat data latih berupa nilai siswa.

Tabel 2. Data Training

No	No Ujian	Nama	Kompetensi Keahlian	Nilai			Status	
				Bahasa Indonesia	Bahasa Inggris	MatematikaKejuruan		
1	23-101-463-2	AHMAD FAISAL	MM	74	80	90	90	LULUS
2	23-101-464-9	AKHAD FADLI	MM	81	82	90	70	LULUS
3	23-101-465-8	ALYA MAWR AZHARI	MM	76	72	70	98	LULUS
4	23-101-466-7	ARI FEDRI	MM	87	86	92	98	LULUS
5	23-101-467-6	ARSANA TIRTO WENING	MM	86	77	96	77	LULUS
6	23-101-468-5	BAGUS FARHANUDDIN	MM	82	82	90	82	LULUS
7	23-101-469-4	CHASELLA BUNGA ADISTIA	MM	82	72	96	72	LULUS
8	23-101-470-3	DEWI PARAMITA	MM	80	84	88	84	LULUS
9	23-101-471-2	AHMAD FAISAL	MM	74	80	90	90	LULUS
10	23-101-472-9	AKHMAD FADLI	MM	81	82	90	70	LULUS
.
.
.
19923-101-498-5		AJI SOLIHIN	TKJ	82	86	75	80	LULUS
20023-101-498-6		ARDHI PRAYOGA	TKJ	72	85	90	98	LULUS

3.1.2 Data Test

Selanjutnya dalam menentukan data adalah membaca data uji. Pada tabel 3, terdapat data test siswa yang berupa nilai dan hasil prediksi.

Table 3. Data Test

No	Nama	Kompetensi Keahlian	Bahasa		Nilai		Status	Prediksi
			Indonesia	Inggris	Matematika	Kejuruan		
1	AHMAD FAISAL	MM	74	80	90	90	Lulus	Lulus
2	AKHMAD FADLI	MM	81	82	90	70	Lulus	Lulus
3	ALYA MAWAR AZHARI	MM	76	72	70	98	Lulus	Lulus
4	ARI FEDRI	MM	87	86	92	98	Lulus	Lulus
5	ARSALNA TIRTO WENING	MM	86	77	96	77	Lulus	Lulus
6	BAGUS FARHANUDDIN	MM	82	82	90	82	Lulus	Lulus
7	CHASELLA BUNGA ADISTIA	MM	82	72	96	72	Lulus	Lulus
8	DEWI PARAMITA	MM	80	84	88	84	Lulus	Lulus
9	AHMAD FAISAL	MM	74	80	90	90	Lulus	Lulus
10	AKHMAD FADLI	MM	81	82	90	70	Lulus	Lulus
39	AJI SOLIHIN	TKJ	82	86	75	80	Lulus	Lulus
40	ARDHI PRAYOGA	TKJ	72	85	90	98	Lulus	Lulus

3.1.3 Probability

Nilai probabilitas untuk setiap atribut diambil dari data uji pada Tabel 4. Pada tabel 5, terdapat hasil dari nilai probabilitas untuk setiap atribut.

Table 4. Probabilitas

Type Data	Type Atribut	Lulus	Tidak Lulus
Kelas	P_Kelas	0.9371	0.0629
Kompetensi	Multimedia	0.598	0.273
Kompetensi	Teknik Komputer dan Jaringan	0.402	0.727

3.2 Hasil Naive Bayes

Proses ini menggunakan 38 data. Berdasarkan *matriks* konfusi, kita dapat melihat bahwa 34 data diprediksi lolos sebagai kelompok data yang lolos dan hingga 2 data diprediksi lolos sebagai kelompok data yang gagal. Anda juga dapat melihat bahwa 0 dari catatan yang diprediksi gagal karena grup data lulus, dan hingga 2 catatan yang diprediksi gagal karena grup data gagal. Bisa melihatnya di Tabel 5:

Table 5. Hasil Naive Bayes

	Nilai Sebenarnya	Nilai Sebenarnya	
		Lulus	Tidak Lulus
Nilai Prediksi	Lulus	34	2
	Tidak Lulus	0	2

- a. *Precision*: Persentase sampel yang benar berhasil diprediksi dengan benar.

$$Precision = \frac{34}{34+2} = \frac{34}{36} = 94,44\%$$

- b. *Recall*: Persentase sampel yang benar diprediksi dengan benar.

$$Recall = \frac{34}{34+2} = \frac{34}{36} = 94,44\%$$

- c. Nilai akurasi dapat dihitung dari hasil ini

$$Accuracy = \frac{34+2}{34+2+0+2} = \frac{36}{38} = 90\% = 90\%$$

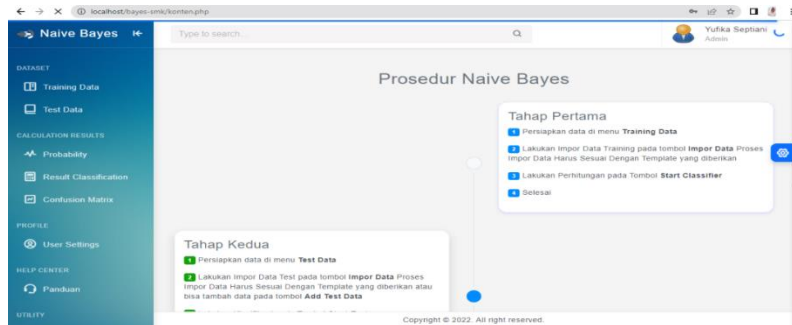
Hasil dari pengujian tersebut mengungkapkan nilai Akurasi, Presisi, dan Recall. Dari sini dapat disimpulkan bahwa perhitungan persentase tingkat akurasi pada *confusion matrix* mencapai nilai *persentase* 90%. Tingkat sensitivitas (*true positive rate (TP rate)* atau *i*) juga tercapai sebesar 94,44%. Selanjutnya dicapai tingkat

spesifisitas (tingkat negatif palsu (FN *rate*) atau presisi) sebesar 94,44%.

3.3 Tampilan Layar Aplikasi

3.3.1 Tampilan Layar Home

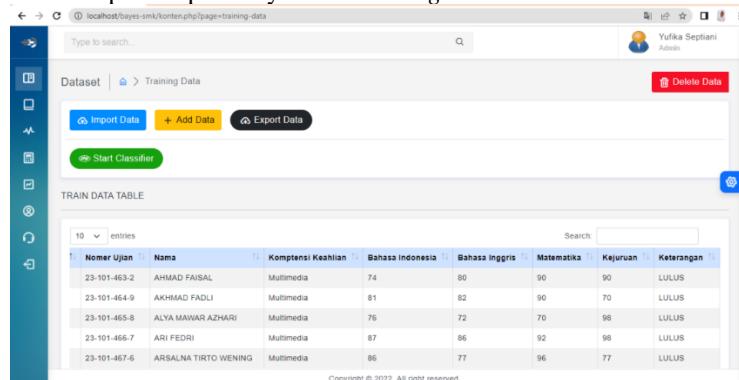
Tampilan layar *home* ini akan tampil pertama kali setelah berhasil masuk melewati halaman daftar. halaman ini menampilkan halaman *administator* untuk info kelulusan. Tampilan layar *home* yang ditunjukkan pada Gambar 1 terdapat tampilan dari layar home.



Gambar 1. Home

3.3.2 Tampilan Layar Training Data

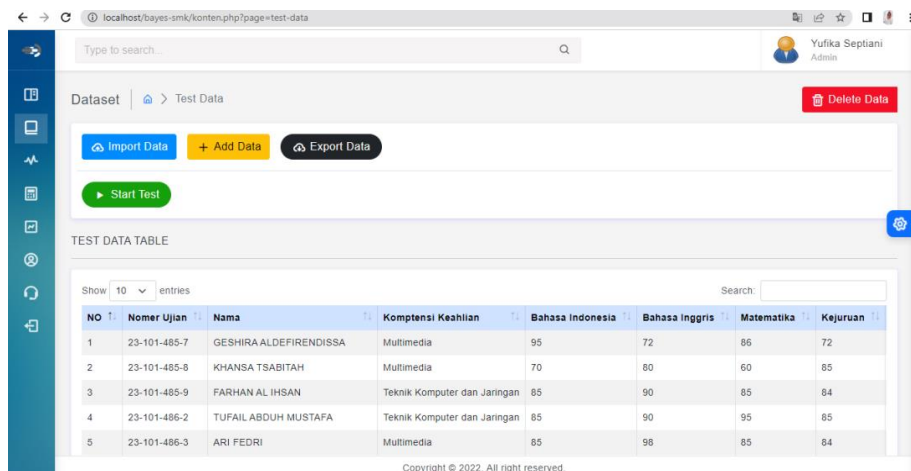
Tampilan Layar *training* data untuk memasukkan sebuah set data *training* yang sudah disiapkan datanya sampai selesai tahap pra-proses data. Pada halaman ini terdapat 4 tombol yaitu import data, *eksport* data, classifier, dan *delete* data Gambar 2 terdapat tampilan layar dari *training* data.



Gambar 2. Training Data

3.3.3 Tampilan Layar Test Data

Tampilan layar data ini menu yang dibuka setelah tahap-tahap yang ada di *menu training data* sudah dilakukan. Menu ini digunakan untuk menghitung tingkat akurasi model algoritma *naive bayes*. Serta terdapat 5 tombol. Tampilan layar data asli pada Gambar 3 terdapat tampilan layar *test* data.



Gambar 3. Test Data

3.3.4 Tampilan Layar *Result Probability*

Tampilan layar data ini menampilkan hasil perhitungan awal algoritma naive bayes dimana dia menampilkan nilai probabilitas dari atribut tersebut Gambar 4 terdapat tampilan layar dari *result probability*.



Type Data	Type Atribut	Lulus	Tidak Lulus
KELAS	P_Kelas	0.9371	0.0629
KOMPETENSI	Multimedia	0.598	0.273
KOMPETENSI	Teknik Komputer dan Jaringan	0.402	0.727

Gambar 4. Result Probability

4. KESIMPULAN

Melalui proses pengolahan dan pengujian yang dilakukan pada pengujian ini, dapat beberapa kesimpulan, antara lain: klasifikasi akan semakin akurat jika data latih yang digunakan berjumlah lebih banyak. Berdasarkan hasil pengujian diperoleh tingkat akurasi sebesar 90%. Diperoleh juga tingkat Sensitivitas (*true positive rate (TP rate) or recall*) sebesar 94,44%. Dan diperoleh juga tingkat *Specificity (false negative rate (FN rate or precision))* sebesar 94,44%. Untuk penelitian selanjutnya sebaiknya menambahkan form nilai berupa numerik dan untuk memperkuat hasil penelitian bisa menambahkan variabel lainnya.

DAFTAR PUSTAKA

- [1] S. Syarli dan A. Muin, "Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi)," *J. Ilm. Ilmu Komput.*, vol. 2, no. 1, hal. 22–26, 2017.
- [2] R. Dwi, Pambudi, A. Afif, Supianto, dan N. Y. Setiawan, "Prediksi Kelulusan Mahasiswa Berdasarkan Kinerja Akademik Menggunakan Pendekatan Data Mining Pada Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Brawijaya," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.* 2196, vol. 3, no. 3, hal. 2194–2200, 2019, [Daring]. Tersedia pada: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/download/4655/2154>.
- [3] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2, hal. 160–165, 2018, doi: 10.33096/ilkom.v10i2.303.160-165.
- [4] A. O. Murdiansyah dan I. Siswanto, "Algoritma Naive Baiyes Classsifier pada Aplikasi Data Mining Berbasis Web," vol. 1, no. 1, hal. 284–290, 2018.
- [5] H. F. Putro, R. T. Vulandari, dan W. L. Y. Saptomo, "Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan," *J. Teknol. Inf. dan Komun.*, vol. 8, no. 2, 2020, doi: 10.30646/tikomsin.v8i2.500.
- [6] S. Widaningsih, "Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naive Bayes, Knn Dan Svm," *J. Tekno Insentif*, vol. 13, no. 1, hal. 16–25, 2019, doi: 10.36787/jti.v13i1.78.
- [7] Betrisandi, "Klasifikasi Nasabah Asuransi Jiwa Menggunakan," vol. 9, no. April, hal. 96–101, 2019.
- [8] M. F. S. Wibowo, N. F. Puspitasari, dan B. Satya, "Penerapan Data Mining Dan Algoritma Naive Bayes Untuk Pemilihan Konsentrasi Mahasiswa Menggunakan Metode Klasifikasi," *J. Inf. Syst. Manag.*, vol. 3, no. 2, hal. 39–45, 2022, doi: 10.24076/joism.2022v3i2.680.
- [9] Y. B. Samponu dan K. Kusrini, "Optimasi Algoritma Naive Bayes Menggunakan Metode Cross Validation Untuk Meningkatkan Akurasi Prediksi Tingkat Kelulusan Tepat Waktu," *J. ELTIKOM*, vol. 1, no. 2, hal. 56–63, 2018, doi: 10.31961/eltikom.v1i2.29.
- [10] Y. B. Samponu dan K. Kusrini, "Optimasi Algoritma Naive Bayes Menggunakan Metode Cross Validation Untuk Meningkatkan Akurasi Prediksi Tingkat Kelulusan Tepat Waktu," *J. ELTIKOM*, vol. 1, no. 2, hal. 56–63, 2018, doi: 10.31961/eltikom.v1i2.29.