

## PREDIKSI MASA TUNGGU KERJA ALUMNI MENGGUNAKAN *NAÏVE BAYES CLASSIFIER* PADA PROGRAM STUDI ILMU KOMPUTER UNIVERSITAS NUSA CENDANA

Rachmadiansyah<sup>1</sup>, Nelci Dessy Rumlaklak<sup>2</sup> dan Arfan Yeheskiel Mauko<sup>3</sup>

<sup>1,2,3</sup>Program Studi Ilmu Komputer, Fakultas Sains dan Teknik, Universitas Nusa Cendana

<sup>1</sup>Email: [rachmadiansyah2000@gmail.com](mailto:rachmadiansyah2000@gmail.com)

<sup>2</sup>Email: [dessyrumlaklak@staf.undana.ac.id](mailto:dessyrumlaklak@staf.undana.ac.id)

<sup>3</sup>Email: [arfanmauko@staf.undana.ac.id](mailto:arfanmauko@staf.undana.ac.id)

### ABSTRAK

Di era global yang penuh dengan tantangan, perguruan tinggi diharapkan dapat menghasilkan lulusan yang berkualitas agar dapat bersaing di dunia kerja. Salah satu indikator yang dapat digunakan untuk menilai kualitas lulusan adalah masa tunggu kerja. Pada penelitian ini penulis mengimplementasikan metode *Naïve Bayes Classifier* dengan menggunakan aplikasi RapidMiner 7.3 untuk menghasilkan prediksi masa tunggu kerja dan tingkat akurasi dari hasil prediksi yang diperoleh. Data pada penelitian ini diperoleh dari hasil kuesioner *tracer study* yang disebar oleh Program Studi Ilmu Komputer Universitas Nusa Cendana untuk mengetahui pencapaian karier para alumni. Atribut yang digunakan pada penelitian ini adalah Masa Studi, Indeks Prestasi Kumulatif (IPK), Keikutsertaan Organisasi, dan Penguasaan Kompetensi dengan kelas Masa Tunggu yang terbagi menjadi 4, yaitu  $\leq 10$  bulan, 11 bulan - 2 tahun 1 bulan, 2 tahun 2 bulan - 3 tahun 4 bulan dan  $> 3$  tahun 4 bulan. Hasil prediksi masa tunggu kerja yang diperoleh, disajikan dalam bentuk *confusion matrix* dengan tingkat akurasi sebesar 81,82%.  
Kata kunci: *Naïve Bayes Classifier*, RapidMiner 7.3, *Confusion Matrix*, Masa Tunggu Kerja

### ABSTRACT

In a global era that is full of challenges, universities are expected to produce quality graduates in order to compete in the world of work. One indicator that can be used to assess the quality of graduates is the job waiting period. In this research, the researcher implements *Naïve Bayes Classifier* method using the RapidMiner 7.3 app to generate predictions for the job waiting period and the accuracy rate of the prediction results obtained. The data in this research were obtained from the results of the Tracer Study questionnaire distributed by Computer Science Study Program at The University of Nusa Cendana to determine the career achievements of alumni. The attributes used in this research are Study Period, Grade Point Average (GPA), Organizational Participation, and Competency Mastery with Waiting Period classes which are divided into 4, namely  $\leq 10$  months, 11 months - 2 years 1 month, 2 years 2 months - 3 years 4 months, and  $> 3$  years 4 months. The prediction results of the job waiting period obtained are presented in the form of a confusion matrix with an accuracy rate of 81.82%.

Keywords: *Naïve Bayes Classifier*, RapidMiner 7.3, *Confusion Matrix*, Job Waiting Period

### 1. PENDAHULUAN

Di era global seperti saat ini, tantangan bagi perguruan tinggi ditandai dengan pesatnya perkembangan ilmu pengetahuan dan teknologi yang penuh dengan persaingan. Dengan demikian, cepat atau lambat perguruan tinggi di Indonesia harus meningkatkan mutu pendidikan agar dapat menghasilkan sumber daya manusia yang berkualitas dan berdaya saing serta memiliki keunggulan tidak hanya dari segi intelektual saja, tetapi juga cerdas dalam berbagai aspek kehidupan [1].

Universitas Nusa Cendana merupakan salah satu perguruan tinggi negeri yang berlokasi di Provinsi Nusa Tenggara Timur. Berdasarkan data yang diperoleh dari sistem informasi *tracer study* alumni Universitas Nusa Cendana sejak tahun 2000 hingga tahun 2020 terdapat 45861 alumni yang berhasil menyelesaikan pendidikannya di Universitas Nusa Cendana. Dengan demikian, diperlukan studi pelacakan jejak alumni untuk keperluan evaluasi capaian manajemen terhadap proses pembelajaran yang dilakukan. Data yang diperlukan untuk studi tersebut diperoleh melalui *tracer study* atau pelacakan data alumni khususnya dalam hal pencarian kerja, situasi kerja, dan pemanfaatan pemerolehan kompetensi selama berkuliah di Universitas Nusa Cendana.

Salah satu fakultas yang didirikan di Universitas Nusa Cendana adalah Fakultas Sains dan Teknik, fakultas ini adalah fakultas ke-7 di Universitas Nusa Cendana yang berfokus pada pengembangan keilmuan

di bidang sains dan teknik, fakultas ini telah menghasilkan 3973 alumni sejak tahun 2000 hingga tahun 2020. Salah satu program studi yang paling diminati di Fakultas Sains dan Teknik adalah Program Studi Ilmu Komputer. Hal ini sejalan dengan bidang keilmuan yang ditawarkan berupa sains dan komputasi yang makin dibutuhkan di dunia kerja. Sejak bulan Juli 2021 Program Studi Ilmu Komputer telah menyebarkan kuesioner *tracer study* bagi para alumni. Hingga bulan November tahun 2021 Program Studi Ilmu Komputer berhasil mendapatkan 157 hasil respons kuesioner dari total 259 alumni yang telah diwisuda hingga tahun 2020.

Salah satu aspek penilaian pada kuesioner *tracer study* dilihat dari masa tunggu alumni sejak menyelesaikan studinya hingga mendapatkan pekerjaan. Pada penelitian ini, penulis menggunakan metode NBC karena metode ini hanya membutuhkan *training data* dalam jumlah yang kecil untuk proses klasifikasi [2]. Metode ini akan diimplementasikan dengan bantuan aplikasi RapidMiner 7.3 untuk menghasilkan prediksi masa tunggu kerja dan tingkat akurasi dari hasil prediksi yang diperoleh. Atribut yang digunakan pada penelitian ini adalah Masa Studi, Indeks Prestasi Kumulatif (IPK), Keikutsertaan Organisasi, dan Penguasaan Kompetensi. Hasil dari penelitian ini dapat digunakan oleh pihak program studi sebagai bahan analisis untuk meningkatkan kualitas pembelajaran bagi para mahasiswa yang tentunya akan berdampak pada kualitas lulusan.

Penelitian lain yang juga menggunakan metode NBC di antaranya adalah prediksi masa studi mahasiswa Program Studi Teknik Informatika Universitas Mercu Buana [3], penentuan strategi promosi penerimaan mahasiswa baru di Universitas Bina Darma [4], diagnosis hepatitis [5], klasifikasi masyarakat miskin di Kabupaten Gorontalo [6], prediksi partisipasi pemilihan Gubernur Jawa Timur [7] dan prediksi pengambil mata kuliah di Fakultas Ilmu Komputer Universitas Brawijaya [8]. Penelitian sejenis untuk memprediksi masa tunggu kerja juga dilakukan dengan metode yang berbeda seperti penelitian [9] dengan metode C4.5 dan penelitian [10] dengan metode *Neural Network*.

## 2. MATERI DAN METODE

### Klasifikasi

Salah satu teknik yang dapat digunakan untuk *data mining* adalah klasifikasi. Beberapa algoritma atau metode yang dapat digunakan untuk klasifikasi diantaranya adalah *Naïve Bayes*, *Decision Tree*, *Neural Network*, *K-Nearest Neighbor*, dan lain-lain. Klasifikasi menggunakan *dataset* awal atau *training data* yang memiliki beberapa atribut dengan nilai yang lengkap sehingga terbentuk sebuah model klasifikasi yang dapat digunakan untuk menemukan nilai atribut yang masih kosong pada *testing data*. Klasifikasi berguna untuk menyajikan informasi yang tidak diketahui sebelumnya dengan menganalisis data yang telah ada [11].

### *Naïve Bayes Classifier* (NBC)

NBC menggunakan metode probabilitas dan statistik untuk memprediksi peluang di masa depan berdasarkan peristiwa yang ada di masa sebelumnya [12]. Metode ini mengasumsikan bahwa semua atribut independen atau tidak saling berketertgantungan satu dengan yang lainnya [13]. Kelebihan dari NBC adalah bahwa metode ini hanya membutuhkan *training data* dalam jumlah kecil yang dibutuhkan dalam proses klasifikasi dan bekerja lebih baik pada banyak situasi yang kompleks [2]. Hasil klasifikasi dengan NBC diperoleh menggunakan persamaan 1.

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)} \dots\dots\dots (1)$$

di mana:

- B = data dengan kelas yang belum diketahui
- A = hipotesis data berupa suatu kelas spesifik
- P(A|B) = probabilitas hipotesis A berdasarkan kondisi B (*posterior probability*)
- P(A) = probabilitas hipotesis A (*prior probability*)
- P(B|A) = probabilitas B berdasarkan kondisi hipotesis A
- P(B) = probabilitas B

### Masa Tunggu Kerja

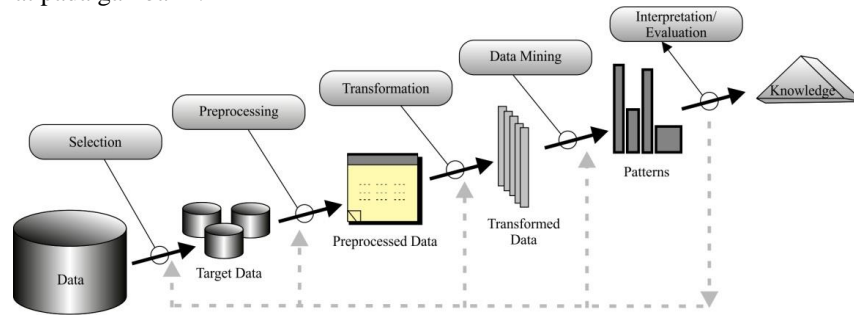
Kerja adalah suatu instrumen yang dapat digunakan seseorang untuk memperoleh uang dan membayar kebutuhan, dengan bekerja seseorang dapat memperoleh pengalaman untuk berinteraksi dengan orang lain dan kesempatan untuk berbagi keterampilan [14]. Dengan demikian, masa tunggu kerja bagi alumni dapat didefinisikan sebagai rentang waktu yang akan dilalui sejak lulus kuliah hingga mendapatkan pekerjaan pertama agar kompetensi yang dimiliki dapat segera dimanfaatkan demi pemenuhan kebutuhan ekonomi.

**RapidMiner**

RapidMiner adalah perangkat lunak yang bersifat *open source* untuk menerapkan proses *data mining* yang dikembangkan oleh RapidMiner Inc. Sebelumnya perangkat lunak ini dikenal dengan nama Yet Another Learning Environment (YALE) dan dikembangkan di Jerman tepatnya di University of Dortmund [15]. RapidMiner adalah perangkat lunak dengan *graphical user interface* (GUI), yang mudah digunakan untuk desain proses *data mining*, visualisasi, validasi, dan optimasi dari proses yang telah didesain sebelumnya [16].

**Metodologi Penelitian**

Data sekunder yang diperoleh melalui studi dokumen berupa 157 respons kuesioner *tracer study* Program Studi Ilmu Komputer Universitas Nusa Cendana akan diproses menggunakan tahapan *data mining* yang dapat dilihat pada gambar 1.



Gambar 1. Tahapan *Data Mining* [17]

**Seleksi Data**

Data mentah yang telah dikumpulkan akan diseleksi sesuai dengan kebutuhan penelitian. Pada tahap ini akan dipilih data tertentu yang akan diolah untuk menghasilkan atribut dan kelas. Data mentah yang dimaksud dapat dilihat pada gambar 2.

	A	B	C	D
1	Timestamp	Jenis Kelamin	Nama Tempat Kerja Sekarang	Jenis instansi/bidang usaha/indus Jabatan/Posisi
2	7-15-2021 15:51:16	Laki-Laki	Kantor Lurah Merdeka	Pemerintah (daerah) Staff Administr
3	7-15-2021 15:54:23	Laki-Laki	KANWIL KEMENTERIAN AGAMA PROVINSI NTT	Pemerintah (pusat/departemen) ADMINISTRAS
4	7-15-2021 16:38:44	Laki-Laki	PLAN INTERNASIONAL	LEMBAGA SWADAYA MASYAR.ADMIN & PRO
5	7-15-2021 20:21:44	Laki-Laki	BP PAUD DIKMAS PROV. NTT	Pemerintah (pusat/departemen) Pengelola Situs
6	7-15-2021 20:59:19	Laki-Laki	Univ. Nusa Cendana	Pemerintah (pusat/departemen) Staf
7	7-16-2021 14:54:26	Laki-Laki	ANGKASA PURA SUPPORT	Swasta (jasa) TEKNISI
8	7-17-2021 16:08:38	Perempuan	SMKN 1 Amarasi Selatan	Pemerintah (daerah) Guru
9	7-22-2021 12:49:55	Perempuan	PT. Bank Mandiri (Persero) Tbk	
10	10-7-2021 15:56:11	Perempuan	Hotel new sasando international	Pemerintah (pusat/departemen) admin
11	10-7-2021 20:54:55	Perempuan	PT. Batutua Khanisma Permai	Swasta (Pertambangan) Coordinator Ge
12	11-1-2021 12:49:19	Laki-Laki	KANTOR IMIGRASI KELAS II TPI ATAMBUA	Pemerintah (pusat/departemen) Analis Keimigr
13	7-15-2021 16:18:50	Laki-Laki	PT Pegadaian	Pemerintah (BUMN,BHMN) Staf
14	7-15-2021 18:29:57	Perempuan	SMK negeri noemuti	Pemerintah (daerah) Guru
15	7-15-2021 18:50:31	Laki-Laki	Samsat Malaka	Pemerintah (daerah) Operator
16	7-15-2021 19:29:13	Perempuan	UKAW Kupang	Swasta (pendidikan) Tenaga Kepenc
17	7-15-2021 21:26:13	Perempuan	PT. Sapta Sari Tama	Perusahaan Besar Farmasi Kasir & EDP
18	7-15-2021 21:27:25	Laki-Laki	Bagian Perekonomian Setda Kota Kupang	Pemerintah (daerah) Staf
19	7-16-2021 8:50:25	Perempuan	PT. Pegadaian (Persero)	Pemerintah (BUMN,BHMN) Penaksir Muda
20	7-16-2021 9:16:02	Laki-Laki	DINKES Kab. TTS	Pemerintah (daerah) Staf

Gambar 2. Data Mentah

**Preprocessing**

Data yang telah diseleksi kemudian akan dibersihkan dari isian yang kosong, respons yang berulang dan isian yang tidak relevan. Gambar 3 menampilkan contoh *record* yang akan dibersihkan pada tahap *preprocessing*.

Q1	R	S	T	U	V	W	X
Saat baru lulus, menurut	Saat baru lulus, menurut	Saat baru lulus, menurut	Saat baru lulus, menurut	Saat baru lulus, menurut	Saat baru lulus, menurut	Saat baru lulus, menurut	Saat baru lulus, menurut Sa
2	4	3	4	3	4	3	4
3	4	3	4	4	4	4	4
4	3	3	3	3	4	4	4
5	3	3	3	3	3	3	3
6	2	2	4	3	4	4	4
7	3	2	4	3	3	3	4
8	2	2	3	2	3	2	2
9	4	4	4	3	4	4	3
10	2	3	2	2	3	2	2
11	4	3	3	3	3	3	3
12	3	2	4	2	3	3	2
13	3	3	2	2	3	3	2
14			4				
15	3	3	3	3	3	3	3
16	3	3	2	2	3	2	3
17	3	3	4	4	3	3	3
18	3	3	3	3	3	3	3
19	3	3	3	3	3	3	3
20	4	3	4	3	4	3	3

Gambar 3. Record dengan Missing Values

Pada gambar 3 terdapat *record* dengan isian yang kosong, opsi untuk menghapus *record* tersebut tentu akan mengurangi jumlah data yang kita miliki. Namun, opsi ini diambil oleh penulis untuk memudahkan proses *preprocessing*. Dari 157 data mentah yang ada terdapat 42 data yang dihapus sehingga tersisa 115 data bersih akan masuk pada tahap transformasi data.

**Transformasi Data**

Data yang telah diseleksi kemudian akan ditransformasikan agar dapat diproses oleh RapidMiner. Tabel 1 dan 2 menampilkan atribut dan kelas yang akan digunakan untuk mentransformasikan data setelah melalui tahap *preprocessing*.

Tabel 1. Atribut

No	Atribut	Nilai
1	Masa Studi	< 5 tahun 5 - 6 tahun > 6 tahun
2	IPK	2,50 - 2,99 3,00 - 3,49 3,50 - 4,00
3	Keikutsertaan Organisasi (KO)	Ya Tidak
4	Penguasaan Kompetensi Pengetahuan Umum (PU), Bahasa Inggris (BI), Komputer (KOM), Metodologi Penelitian (MP), Kerja Sama Tim (KST), Keterampilan Komunikasi Lisan (KKL), Keterampilan Komunikasi Tertulis (KKT), Proses Pemberdayaan Masyarakat (PPM), Pengetahuan Teoretis Spesifik Fakultas (PTSF), Pengetahuan Praktis Spesifik Fakultas (PPSF), Manajemen Organisasi (MO), dan Kepemimpinan (KEP).	1: Tidak Menguasai 2: Kurang Menguasai 3: Menguasai 4: Sangat Menguasai

Tabel 2. Kelas

Kelas	Nilai
Masa Tunggu	≤ 10 bulan
	11 bulan - 2 tahun 1 bulan
	2 tahun 2 bulan - 3 tahun 4 bulan
	> 3 tahun 4 bulan

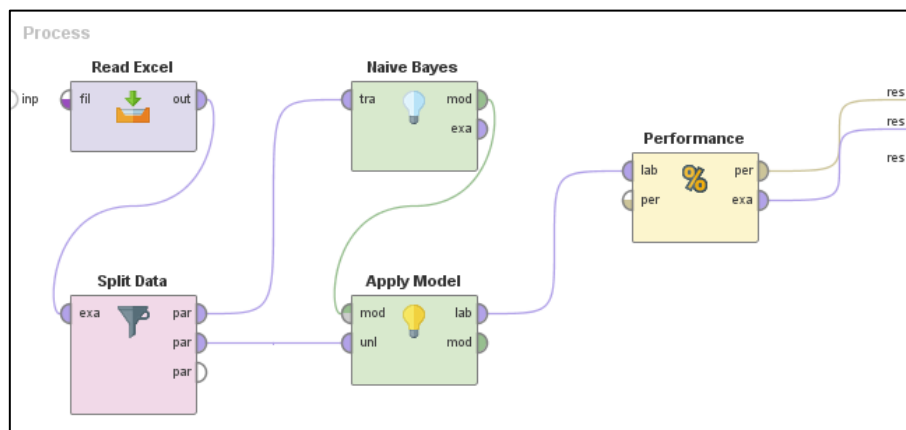
*Dataset* yang dihasilkan pada tahap transformasi adalah *dataset* yang siap digunakan pada proses prediksi, atribut dengan tipe data numerik seperti Penguasaan Kompetensi dapat diubah menjadi nominal pada saat *user* mengimpor data ke dalam aplikasi RapidMiner 7.3, sedangkan atribut lainnya akan otomatis diidentifikasi sebagai atribut dengan tipe data nominal. *Dataset* yang dimaksud dapat dilihat pada tabel 3.

Tabel 3. *Dataset*

No	Masa Studi	IPK	KO	PU	BI	KOM	MP	KST	KKL	KKT	PPM	PTSF	PPSF	MO	KEP	Masa Tunggu
1	> 6 tahun	3,00 - 3,49	Ya	3	3	2	2	3	3	3	2	3	3	2	3	≤ 10 bulan
2	> 6 tahun	2,50 - 2,99	Ya	4	3	4	3	4	3	3	4	3	3	3	3	11 bulan - 2 tahun 1 bulan
3	> 6 tahun	3,00 - 3,49	Ya	3	2	4	3	3	3	3	4	3	3	3	3	2 tahun 2 bulan - 3 tahun 4 bulan
4	5 - 6 tahun	2,50 - 2,99	Tidak	2	2	3	2	3	2	2	2	2	2	2	2	> 3 tahun 4 bulan
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
115	> 6 tahun	3,00 - 3,49	Tidak	2	4	4	2	3	2	3	2	2	2	2	2	> 3 tahun 4 bulan

**Implementasi Metode Naïve Bayes Classifier**

Pada tahap ini, metode *Naïve Bayes Classifier* akan diimplementasikan menggunakan aplikasi RapidMiner 7.3 dengan membuat sebuah model *data mining* yang dapat dilihat pada gambar 4.



Gambar 4. Model *Data Mining* pada RapidMiner 7.3

Terdapat 5 buah operator yang digunakan untuk membuat model *data mining* yaitu:

1. *Read Excel*: User dapat mengimpor data dari Microsoft Excel pada operator ini, kemudian menentukan tipe data pada setiap kolom serta menentukan kolom yang menjadi atribut dan kelas pada tabel yang telah diimpor.
2. *Split Data*: Pada operator ini, data akan dibagi menjadi *training data* dan *testing data* berdasarkan rasio yang ditentukan oleh user agar dapat menghasilkan model klasifikasi dengan nilai akurasi terbaik.
3. *Naive Bayes*: Metode *Naive Bayes Classifier* akan bekerja pada operator ini untuk menghasilkan pola pada *training data* yang selanjutnya akan digunakan pada *testing data*.
4. *Apply Model*: Operator ini akan menerapkan model yang sudah dipelajari atau dilatih pada tahap sebelumnya.
5. *Performance*: Nilai-nilai yang menunjukkan kinerja dari model klasifikasi yang ada akan dihasilkan oleh operator ini. Nilai-nilai yang dimaksud adalah *accuracy*, *precision*, dan *recall*.

Semua operator dihubungkan satu dengan yang lainnya seperti pada gambar 2, kemudian hasil yang diperoleh akan tampil ketika user mengeklik tombol *Run* yang berada pada bagian atas lembar kerja. Hasil yang dimaksud adalah prediksi yang diperoleh pada setiap *record* yang ada pada *data testing* dan *confusion matrix* beserta nilai-nilai yang merepresentasikan kinerja dari metode yang digunakan.

**Evaluasi**

Hasil prediksi masa tunggu yang diperoleh akan dievaluasi dengan menggunakan *confusion matrix* untuk mengukur kinerja yang diberikan oleh metode yang digunakan terhadap dataset yang ada. *Confusion matrix* dapat dilihat pada tabel 4.

Tabel 4. *Confusion Matrix* [18]

		Predicted class		
		yes	no	Total
Actual class	yes	TP	FN	P
	no	FP	TN	N
Total		P'	N'	P+N

keterangan:

*True Positive* (TP) : Kelas *yes* yang dilabeli dengan benar oleh metode klasifikasi.

*True Negative* (TN) : Kelas *no* yang dilabeli dengan benar oleh metode klasifikasi.

*False Positive* (FP) : Kelas *no* yang salah dilabeli dengan *yes* oleh metode klasifikasi.

*False Negative* (FN) : Kelas *yes* yang salah dilabeli dengan *no* oleh metode klasifikasi.

Kriteria evaluasi yang digunakan dalam pengujian *confusion matrix* adalah:

1. *Accuracy*

*Accuracy* dari sebuah metode klasifikasi adalah persentase kelas yang dilabeli dengan benar, rumus *accuracy* dapat dilihat pada persamaan 2.

$$Accuracy = \frac{TP+TN}{P+N} \dots\dots\dots (2)$$

2. *Precision*

*Precision* dapat diartikan sebagai kadar ketepatan, contohnya: Persentase kelas yang dilabeli dengan *yes* dan juga *yes* pada kenyataannya, rumus *precision* dapat dilihat pada persamaan 3.

$$Precision = \frac{TP}{TP+FP} \dots\dots\dots (3)$$

3. *Recall*

*Recall* dapat juga diartikan sebagai kadar kelengkapan, contohnya: Persentase kelas *yes* yang juga dilabeli dengan *yes*, rumus *recall* dapat dilihat pada persamaan 4.

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots (4)$$

**3. HASIL DAN PEMBAHASAN**

**Hasil**

Hasil yang diperoleh pada penelitian ini adalah tingkat akurasi yang dapat diberikan oleh metode NBC untuk memprediksi masa tunggu kerja alumni pada Program Studi Ilmu Komputer. *Dataset* yang berjumlah 115 data akan dibagi oleh operator *split data* pada aplikasi RapidMiner 7.3 dengan rasio 9 : 1 (104 *training data* dan 11 *testing data*). Rasio 9 : 1 digunakan pada penelitian ini karena dapat menghasilkan model klasifikasi dengan nilai akurasi terbaik. Hasil prediksi yang diperoleh akan ditampilkan dalam bentuk *confusion matrix* yang dapat dilihat pada tabel 5.

Tabel 5. *Confusion Matrix* Hasil

*Accuracy*: 81,82%

	<i>True</i> ≤ 10 bulan	<i>True</i> 11 bulan - 2 tahun 1 bulan	<i>True</i> 2 tahun 2 bulan - 3 tahun 4 bulan	<i>True</i> > 3 tahun 4 bulan	<i>Class</i> <i>precision</i>
<i>Prediction</i> ≤ 10 bulan	6	0	1	0	85,71%
<i>Prediction</i> 11 bulan - 2 tahun 1 bulan	0	2	0	0	100,00%
<i>Prediction</i> 2 tahun 2 bulan - 3 tahun 4 bulan	0	0	0	0	0,00%
<i>Prediction</i> > 3 tahun 4 bulan	0	0	1	1	50,00%
<i>Class recall</i>	100,00%	100,00%	0,00%	100,00%	

Nilai-nilai pada tabel 5 yang merepresentasikan performa dari metode yang digunakan dapat dihitung secara manual. *Accuracy* dapat dihitung menggunakan persamaan 2, *precision* dapat dihitung menggunakan persamaan 3 dan *recall* dapat dihitung menggunakan persamaan 4:

1.  $Accuracy = \left( \frac{6+2+0+1}{6+0+1+0+0+2+0+0+0+0+0+0+0+0+1+1} \right) \times 100\% = 81,82\%$

2. *Precision*

$P(\leq 10 \text{ bulan}) = \left( \frac{6}{6+0+1+0} \right) \times 100\% = 85,71\%$

$P(11 \text{ bulan} - 2 \text{ tahun} 1 \text{ bulan}) = \left( \frac{2}{0+2+0+0} \right) \times 100\% = 100\%$

$P(2 \text{ tahun} 2 \text{ bulan} - 3 \text{ tahun} 4 \text{ bulan}) = \left( \frac{0}{0+0+0+0} \right) \times 100\% = 0\%$

$P(> 3 \text{ tahun} 4 \text{ bulan}) = \left( \frac{1}{0+0+1+1} \right) \times 100\% = 50\%$

3. *Recall*

$R(\leq 10 \text{ bulan}) = \left( \frac{6}{6+0+0+0} \right) \times 100\% = 100\%$

$R(11 \text{ bulan} - 2 \text{ tahun} 1 \text{ bulan}) = \left( \frac{2}{0+2+0+0} \right) \times 100\% = 100\%$

$R(2 \text{ tahun} 2 \text{ bulan} - 3 \text{ tahun} 4 \text{ bulan}) = \left( \frac{0}{1+0+0+1} \right) \times 100\% = 0\%$

$R(> 3 \text{ tahun} 4 \text{ bulan}) = \left( \frac{1}{0+0+0+1} \right) \times 100\% = 100\%$

**Pembahasan**

Berdasarkan *confusion matrix* pada tabel 5, nilai akurasi yang diperoleh adalah 81,82%. Percobaan ke-2 dan ke-3 juga dilakukan dengan menggunakan kombinasi *training data* dan *testing data* yang berbeda, tetapi tetap dengan rasio yang sama, yaitu 9 : 1. Hasil prediksi yang diperoleh memiliki tingkat akurasi yang sama dengan percobaan sebelumnya, yaitu 81,82%. Selain itu, penulis juga menganalisis data yang ada untuk mencari atribut yang paling signifikan dalam memengaruhi masa tunggu kerja seorang alumni, tabel 6 menampilkan persentase atribut terhadap kelas yang akan digunakan untuk proses analisis.

Tabel 6. Persentase Atribut terhadap Kelas

		≤ 10 bulan	11 bulan - 2 tahun	2 tahun 2 bulan - 3 tahun	> 3 tahun 4 bulan
<b>Masa Studi</b>	< 5 tahun	16%	0%	5,3%	26,1%
	5 - 6 tahun	58%	47,8%	31,6%	21,7%
	> 6 tahun	26%	52,2%	63,2%	52,2%
<b>IPK</b>	2,50 - 2,99	14%	8,7%	26,3%	30,4%
	3,00 - 3,49	60%	78,3%	57,9%	56,5%
	3,50 - 4,00	26%	13%	15,8%	13%
<b>Keikutsertaan Organisasi</b>	Ya	76%	65%	58%	43%
	Tidak	24%	35%	42%	57%
<b>Penguasaan Kompetensi</b>	4	32,8%	18,1%	29,4%	23,9%
	3	54,2%	74,6%	55,7%	51,8%
	2	13%	7%	13%	24%
	1	0%	0%	2%	0%

Persentase atribut terhadap kelas yang ditampilkan pada tabel 6 terlihat sangat bervariasi. Namun, apabila diperhatikan dengan baik, persentase “Ya” pada atribut “Keikutsertaan Organisasi” makin menurun seiring bertambahnya masa tunggu, sedangkan persentase “Tidak” pada atribut “Keikutsertaan Organisasi” makin meningkat seiring bertambahnya masa tunggu. Dengan demikian, atribut “Keikutsertaan Organisasi” adalah atribut yang paling memengaruhi masa tunggu kerja seorang alumni, jika semasa berkuliah seorang alumni aktif berorganisasi, maka akan makin tinggi kemungkinan alumni tersebut untuk mendapatkan pekerjaan dengan waktu yang lebih singkat begitu juga sebaliknya.

#### 4. KESIMPULAN DAN SARAN

##### Kesimpulan

Metode NBC memiliki performa yang baik untuk memprediksi masa tunggu kerja alumni Program Studi Ilmu Komputer dengan tingkat akurasi yang mencapai 81,82% pada ketiga percobaan yang dilakukan. Dengan demikian, atribut Masa Studi, IPK, Keikutsertaan Organisasi dan Penguasaan Kompetensi dapat digunakan oleh pihak program studi dan mahasiswa aktif sebagai indikator yang berperan penting dalam menentukan masa tunggu seorang alumni sebelum memperoleh pekerjaan.

##### Saran

Penelitian serupa dapat dilakukan beberapa tahun ke depan agar data yang digunakan dapat bertambah secara signifikan. Selain itu, peneliti selanjutnya dapat menggunakan metode klasifikasi yang berbeda atau dapat menggunakan dua atau lebih metode klasifikasi untuk mengetahui metode manakah yang dapat memberikan hasil prediksi yang lebih akurat. Penilaian Penguasaan Kompetensi alumni juga dapat dilakukan oleh pihak lain yang mampu memberikan nilai secara lebih objektif seperti atasan langsung atau pimpinan kantor pada tempat alumni tersebut bekerja.

##### DAFTAR PUSTAKA

- [1] M. Kadarisman, “Tantangan Perguruan Tinggi dalam Era Persaingan Global,” *Sociae Polites*, pp. 3–20, Oct. 2017, doi: [10.33541/sp.v1i1.459](https://doi.org/10.33541/sp.v1i1.459).
- [2] S. A. Pattekari and A. Parveen, “Prediction System for Heart Disease Using Naïve Bayes,” *International Journal of Advanced Computer and Mathematical Sciences*, vol. 3, no. 3, pp. 290–294, 2012.
- [3] F. E. Prabowo and A. Kodar, “Analisis Prediksi Masa Studi Mahasiswa Menggunakan Algoritma Naïve Bayes,” *Jurnal Ilmu Teknik dan Komputer*, vol. 3, no. 2, pp. 147–151, Jul. 2019, doi: [10.22441/jitkom.2020.v3.i2.008](https://doi.org/10.22441/jitkom.2020.v3.i2.008).
- [4] A. H. Mirza, “Penerapan Algoritma Naive Bayes Classifier dalam Menentukan Strategi Promosi Penerimaan Mahasiswa Baru,” *Journal of Information Systems and Informatics*, vol. 1, no. 1, pp. 14–28, Mar. 2019.
- [5] D. Novianti, “Implementasi Algoritma Naïve Bayes Pada Data Set Hepatitis Menggunakan Rapid Miner,” *Jurnal Sistem Informasi, Teknik Informatika, Software Engineering, dan Multimedia*, vol. 21, no. 1, pp. 49–54, Mar. 2019, doi: [10.31294/p.v21i1.4979](https://doi.org/10.31294/p.v21i1.4979).
- [6] H. Annur, “Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes,” *ILKOM Jurnal Ilmiah.*, vol. 10, no. 2, pp. 160–165, Aug. 2018, doi: [10.33096/ilkom.v10i2.303.160-165](https://doi.org/10.33096/ilkom.v10i2.303.160-165).
- [7] A. S. Fitriani, “Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur,” *JTAM (Jurnal Teori dan Aplikasi Matematika)*, vol. 3, no. 2, pp. 98–104, Oct. 2019.

- [8] I. K. Syahputra, F. A. Bachtiar, and S. A. Wicaksono, "Implementasi Data Mining untuk Prediksi Mahasiswa Pengambil Mata Kuliah dengan Algoritme Naive Bayes," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 11, pp. 5902–5910, Nov. 2018.
- [9] R. Cahyaningtyas, Luqman, and R. Y. I. Heriyanto, "Klasifikasi Kompetensi Alumni Berdasarkan Masa Tunggu Alumni untuk Mendapatkan Pekerjaan Menggunakan Metode Algoritma C4.5," *Jurnal KILAT*, vol. 9, no. 2, pp. 297–310, Oct. 2020.
- [10] Khoirudin, S. Hadi, and A. Nugroho, "Analisa dan Penerapan Metode Neural Networks dalam Mengidentifikasi Faktor-Faktor Masa Tunggu Kerja Lulusan," *Jurnal Pengembangan Rekayasa dan Teknologi*, vol. 16, no. 1, pp. 17–22, Jun. 2020, doi: [10.26623/jprt.v16i1.2399](https://doi.org/10.26623/jprt.v16i1.2399).
- [11] M. J. Zaki and W. Meira, *Data Mining and Analysis: Fundamental Concepts and Algorithms*. New York, NY: Cambridge University Press, 2014.
- [12] Bustami, "Penerapan Algoritma Naive Bayes untuk Mengklasifikasi Data Nasabah Asuransi," *JURNAL INFORMATIKA*, vol. 8, no. 1, pp. 884–898, Jan. 2014.
- [13] T. R. Patil and M. S. S. Sherekar, "Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification," *International Journal of Computer Science and Applications*, vol. 6, no. 2, pp. 256–261, Apr. 2013.
- [14] A. H. Wiltshire, "The Meanings of Work in A Public Work Scheme in South Africa," *International Journal of Sociology and Social Policy*, vol. 36, no. 1/2, pp. 2–17, Mar. 2016, doi: [10.1108/IJSSP-02-2015-0014](https://doi.org/10.1108/IJSSP-02-2015-0014).
- [15] I. Mierswa, M. Wurst, R. Klinkenberg, M. Scholz, and T. Euler, "YALE: Rapid Prototyping for Complex Data Mining Tasks," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, USA, 2006, pp. 935–940. doi: [10.1145/1150402.1150531](https://doi.org/10.1145/1150402.1150531).
- [16] M. Hofmann and R. Klinkenberg, *RapidMiner: Data Mining Use Cases and Business Analytics Applications*. USA: CRC Press, 2016. [Online]. Available: [https://books.google.co.id/books?id=Y\\_wYCwAAQBAJ](https://books.google.co.id/books?id=Y_wYCwAAQBAJ)
- [17] U. M. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "Knowledge Discovery and Data Mining: Towards a Unifying Framework.," in *KDD*, 1996, vol. 96, pp. 82–88.
- [18] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*, 3rd ed. USA: Morgan Kaufmann Publishers, 2011. [Online]. Available: <https://books.google.co.id/books?id=pQws07tdpjoC>