



## KOMPARASI ALGORITMA KLASIFIKASI DATAMINING UNTUK PREDIKSI MINAT PENCARI KERJA

Nur Widiastuti<sup>1)</sup>, Arief Hermawan<sup>2)</sup>, Donny Avianto<sup>3)</sup>

<sup>1,2</sup>Magister Teknologi Informasi, Universitas Teknologi Yogyakarta

<sup>3</sup> Program Studi Informatika, Universitas Teknologi Yogyakarta

<sup>1,2,3</sup> Jl. Ring Road Lor, Yogyakarta

Email: <sup>1</sup>widy.jogja@gmail.com, <sup>2</sup>ariefdb@uty.ac.id, <sup>3</sup>donny@uty.ac.id

### Abstract

*The impact of the COVID-19 pandemic has made the unemployment rate in Central Java Province increase. Dinas Tenaga Kerja of Central Java Province created an innovation, namely an application to bring together job seekers and employers, namely the e-Makaryo application or bursakerja.jatengprov.go.id, but the application has not been able to analyze how many numbers there are. job seekers who are interested in working abroad and job seekers who are not interested in working abroad, while this is urgently needed by the government to prepare job vacancies as needed. In connection with this, the author will analyze the data of job seekers who are interested in working abroad using a rapid miner application with the Naive Bayes algorithm classification method and Decision Tree. UK classification error is 0.36%. While the accuracy of performance using the decision tree algorithm is 68.59% with the number of job seekers who are interested in working abroad as many as 2521 and those who are not interested in working abroad as many as 5504. While the results of the classification error are 31.41%.*

**Keyword:** algorithm, classification, job seekers, naïve bayes, decision tree.

### Abstrak

Dampak pandemi covid 19 menjadikan angka pengangguran di Provinsi Jawa Tengah semakin meningkat. Dinas Tenaga Kerja Provinsi Jawa Tengah menciptakan inovasi yakni sebuah aplikasi untuk mempertemukan pencari kerja dan pemberi kerja, yakni aplikasi e-Makaryo atau bursakerja.jatengprov.go.id, namun pada aplikasi tersebut belum dapat menganalisa berapa banyak jumlah pencari kerja yang berminat kerja ke luar negeri dan pencari kerja yang tidak berminat bekerja ke luar negeri, sedangkan hal ini sangat dibutuhkan oleh pemerintah untuk menyiapkan lowongan kerja sesuai kebutuhan. Sehubungan dengan hal tersebut penulis akan melakukan analisa data pencari kerja yang berminat bekerja keluar negeri menggunakan aplikasi rapid miner dengan metode klasifikasi algoritma *Naive Bayes* dan *Decision Tree*. Hasil *accuracy performance* pada klasifikasi data pencari kerja menggunakan algoritma *naïve bayes* adalah sebesar 99,64% dengan hasil yang berminat bekerja keluar negeri sebanyak 2513 orang dan yang tidak berminat bekerja ke luar negeri sebanyak 5483 orang, untuk *classification error* sebesar 0,36 %. Sedangkan *accuracy performance* menggunakan algoritma *decision tree* adalah 68.59 % dengan jumlah pencari kerja yang berminat bekerja ke luar negeri sebanyak 2521 dan yang tidak berminat bekerja keluar negeri sebanyak 5504. Sedangkan hasil *classification error*nya sebesar 31,41%.

**Kata Kunci:** algoritma, klasifikasi, pencari kerja, naïve bayes, decision Tree.

## 1. PENDAHULUAN

Perkembangan teknologi di era digital saat ini semakin canggih dan menuntut sistem pemerintahan juga harus dapat mengikuti perkembangan jaman agar dapat memberikan kemudahan pada pengguna dalam mencari informasi yang lengkap untuk dapat mengolah data-data yang ada. Di era digital ini ada hal yang masih menjadi tugas penting bagi pemerintah yakni angka pengangguran yang masih tinggi, hal tersebut dikarenakan kurangan informasi lowongan pekerjaan [1]. Badan Pusat Statistik (BPS) merilis data jumlah angkatan kerja pada bulan Agustus 2021 adalah sejumlah 18,96 juta orang, angka tersebut sudah bertambah 213 ribu orang dibandingkan dengan jumlah angkatan kerja pada bulan Agustus 2020. Seiring dengan kenaikan jumlah angkatan kerja tersebut, Tingkat Partisipasi Angkatan Kerja (TPAK) ada sedikit peningkatan menjadi 69,58 % (persen). Tingkat pengangguran terbuka (TPT) pada bulan Agustus



2021 sebesar 5,95 persen atau sebanyak 1,13 juta orang, hal itu berarti menurun sebesar 0,53 % (persen) dibanding pada bulan Agustus 2020 [2].

Adapun jumlah penduduk sejak bulan Agustus 2020 yang telah bekerja sejumlah 17,84 juta orang, artinya terjadi peningkatan sebanyak 299 ribu orang. Jumlah tenaga kerja menurun dibanding pada bulan Agustus 2020, diantaranya meliputi sektor bidang Transportasi dan sektor Perdagangan (0,11 persen poin), Pertanian (2,54 persen poin), serta Jasa Lainnya (0,33 persen poin). Dengan melihat data angkatan kerja dari Badan Pusat Statistik diatas maka Dinas Tenaga Kerja Prov. Jateng menciptakan inovasi yakni sebuah aplikasi untuk menekan angka pengangguran dan sebagai sarana untuk memudahkan komunikasi para pencari kerja dan pemberi kerja, yakni aplikasi e-Makaryo atau [bursakerja.jatengprov.go.id](http://bursakerja.jatengprov.go.id). Aplikasi e-makaryo ini adalah pengembangan dari website bursa kerja online (BKO) yang telah diluncurkan sejak tahun 2005. Aplikasi ini dibuat dengan alur yang mudah agar bias dipahami oleh masyarakat pencari kerja. Kepala Dinas Tenaga Kerja Prov. Jateng menjelaskan, E-Makaryo adalah salah satu upaya agar para pencari kerja tetap terhubung dengan pemberi kerja atau perusahaan. Dengan adanya e-Makaryo ini, ia berharap dapat menekan angka pengangguran di Provinsi Jawa Tengah [3].

Beberapa penelitian terdahulu yang pernah dilakukan oleh peneliti lain terkait data mining untuk prediksi antara lain, Penelitian oleh (Hakam, Retno, Wawan, 2020) dengan judul Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan, yakni tentang klasifikasi pelanggan menggunakan metode naive bayes. Data latih dan data uji yang digunakan untuk klasifikasi berjumlah 75 data latih dan 25 data uji. Berdasarkan master pelanggan yang dijadikan data latih, telah berhasil mengklasifikasikan 23 data dari 25 data yang diuji. Sehingga berhasil memprediksi pelanggannya dengan nilai *precision* mencapai 100%, nilai *recall* mencapai 91%, nilai *accuracy* mencapai 92%[4]. Penelitian lain oleh (senna hendrian, 2018) dengan judul Algoritma klasifikasi data mining untuk memprediksi siswa dalam memperoleh bantuan dana pendidikan yakni tentang Analisa Bantuan Dana Pendidikan bagi Siswa yang dianggap kurang mampu secara strata ekonomi menggunakan Algoritma Klasifikasi Datamining yaitu Algoritma C4.5. Sampel data diambil dari SMA Bina Bangsa Mandiri yang beralamat di Kecamatan Gunungputri Kab. Bogor. Dari hasil pengujian digunakan tes Cros Validation dan Confusion Matrix dan Kurva ROC. Hasil yang diperoleh untuk nilai Accuracy Algoritma C4.5 adalah sebesar 98,80%, nilai untuk Precision sebesar 98,02%, dan nilai untuk Sensitivity atau Recall sebesar 99,00%. Dengan demikian Algoritma C4.5 merupakan algoritma dan teknik terbaik untuk Memprediksi Siswa dalam memperoleh Bantuan Dana Pendidikan[5]. Penelitian lain oleh (Riski Annisa, 2019) dengan judul Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Penderita Penyakit Jantung, yakni tentang perbandingan beberapa algoritma klasifikasi yaitu Decision Tree, Naive Bayes, k-Nearest Neighbour, Random Forest, dan Decision Stump dengan menggunakan uji parametrik dengan t-test agar dapat menghasilkan perbandingan metode yang lebih baik untuk prediksi penderita penyakit jantung, adapun data set yang digunakan adalah laki-laki penderita Penyakit jantung. Hasil penelitian mendapatkan nilai akurasi sebesar tertinggi sebesar 80.38%. Hasil penelitian menunjukkan bahwa algoritma random forest dan decision stump melakukan performa terbaik dalam pengklasifikasi di dataset, C4.5 dan Naive bayes juga tampil baik, kemudian k-NN merupakan algoritma yang kurang baik diimplementasikan dalam dataset[6]. Penelitian lainnya oleh (Adhitya, Kurniyatul dan Khadijah, 2021) dengan judul Analisis Perbandingan Algoritma Decision Tree, kNN, dan Naive Bayes untuk Prediksi Kesuksesan Start-up, yakni tentang pengklasifikasian *start-up* sukses dan gagal, sehingga nantinya dapat digunakan untuk melihat faktor-faktor yang paling mempengaruhi keberhasilan *start-up*, dan juga dapat memprediksi keberhasilan suatu *start-up*. Di antara banyaknya metode klasifikasi dalam *data mining*, algoritma Decision Tree, kNN, dan Naive Bayes merupakan algoritma yang penulis pilih untuk melakukan klasifikasi terhadap 923 *record* data *start-up* yang telah didapatkan sebelumnya. Hasil pengujian menggunakan *cross validation* dan *T-test* menunjukkan algoritma Decision Tree merupakan algoritma paling tepat untuk melakukan klasifikasi dalam studi kasus ini. Hal ini dibuktikan dengan nilai akurasi yang diperoleh oleh algoritma Decision Tree lebih besar diantara algoritma lainnya, yaitu sebesar 79,29%, sedangkan algoritma kNN memiliki nilai akurasi 66,69%, dan Naive Bayes sebesar 64,21%[7]. Penelitian selanjutnya oleh (Rado, Laradea dan Irma, 2020) dengan judul Klasifikasi Data Mining Untuk Seleksi Penerimaan Calon Pegawai Negeri Sipil Tahun 2017 Menggunakan Metode Naive Bayes, yakni tentang Klasifikasi seleksi penerimaan Calon Pegawai Negeri sipil (CPNS) 2017. Metode yang digunakan adalah metode Naive Bayes yaitu metode untuk memprediksi berbasis probabilitas. Penelitian ini bertujuan untuk mendapatkan distribusi yang tepat untuk fitur-fitur data kontinu. Dari hasil pengujian yang dilakukan dengan menggunakan data sebanyak 284 dengan 227 sebagai data training dan 57 sebagai data testing baik untuk fitur kontinu menggunakan distribusi normal maupun distribusi gamma, akurasi untuk distribusi normal adalah 81% dengan jumlah data yang tepat sebanyak 46 dan yang tidak tepat sebanyak 11. Sedangkan akurasi distribusi gamma sebesar 70% dengan data yang tepat sebanyak 40 dan data yang tidak tepat sebanyak 17[8].

Pada penelitian ini data pencari kerja yang akan di analisa berjumlah 11.464 baris data sehingga tidak mungkin untuk memilah data tersebut secara manual, tentunya membutuhkan sebuah metode yang tepat untuk mengolah data tersebut. Data mining merupakan proses atau metode yang tepat untuk mengolah dan memprediksi data pencari kerja



dengan jumlah yang sangat banyak. Keuntungan menggunakan data mining, akan lebih mudah untuk memantau dan mengamati kebiasaan perilaku yang dilakukan oleh pencari kerja, sehingga Dinas Tenaga Kerja Provinsi Jawa Tengah bisa menyusun strategi yang lebih efektif untuk menyiapkan lowongan kerja sesuai dengan kondisi pencari kerja.

Namun pada aplikasi bursa kerja online (BKO) atau e-makaryo belum dapat menganalisa berapa banyak jumlah pencari kerja yang berminat kerja ke luar negeri dan pencari kerja yang tidak berminat bekerja ke luar negeri, sedangkan hal ini sangat dibutuhkan oleh pemerintah untuk menyiapkan lowongan kerja sesuai kebutuhan. Sehubungan dengan hal tersebut penulis akan melakukan analisa data pencari kerja yang berminat bekerja keluar negeri menggunakan aplikasi rapid miner dengan metode klasifikasi algoritma Naive Bayes dan Decision Tree. Hasil analisa tersebut nantinya akan digunakan sebagai bahan pertimbangan pemerintah dalam menyiapkan lowongan kerja sesuai kebutuhan.

## 2. METODE PENELITIAN

### 2.1 Tahapan Penelitian

Ada beberapa tahapan dalam penelitian yang akan dilakukan menggunakan metode Datamining Klasifikasi Algoritma Naive Bayes dan Decision Tree dengan software Rapidminer. Jumlah data yang digunakan sebanyak 11.464 baris dan 12 atribut. Adapun alur atau tahapan penelitian sebagai berikut :

1. Menyiapkan *dataset*,
2. *Import* ke dalam aplikasi rapidminer,
3. *Select attribute*,
4. Proses analisa menggunakan operator *naive bayes* dan *decision tree*,
5. *Apply model* dan *performance* sehingga di dapatkan hasil *accuracy performance*.



**Gambar 1.** Tahapan Penelitian

### 2.2 Dataset

Dataset yang digunakan dalam penelitian ini adalah data yang asli, data ini didapatkan langsung dari Bursa Kerja Online Provinsi Jawa Tengah atau e-makaryo. Penelitian ini menggunakan dataset pencari kerja dari website e-Makaryo Jateng, dimana aplikasi berbasis web tersebut adalah sebuah inovasi Dinas Tenaga Kerja Prov. Jateng untuk mempertemukan pencari kerja dan penyedia kerja serta memberikan kemudahan kepada pencari kerja untuk mendapatkan pekerjaan baik di dalam negeri maupun ke luar negeri.



### 2.3 Variabel (Atribut) Penelitian

Pada penelitian ini jumlah variabel yang akan digunakan terdiri dari 12 atribut yang dipilih dari 16 atribut yang ada pada dataset. 12 atribut ini telah dipilih berdasarkan atribut yang paling penting karena ada beberapa atribut yang dapat mewakili beberapa atribut lain seperti nama, jenis kelamin, tanggal lahir dan asal kabupaten. Dari 12 atribut yang dipilih sebagai label atau class yaitu Minat Kerja LN, sedangkan 11 atribut lainnya sebagai atribut biasa. Adapun attribute terlihat pada gambar berikut ini :

**Tabel 1.** Attribute Dataset Pencari Kerja

Atribut	Tipe Data	Class
Minat Kerja Luar Negeri (LN)	Binominal	Class/Label
Tempat Lahir	Polynominal	Attribute
Tanggal Lahir	Date_time	Attribute
Usia (Tahun)	Integer	Attribute
Alamat	Polynominal	Attribute
Kab/Kota	Polynominal	Attribute
Kecamatan	Polynominal	Attribute
Gender	Polynominal	Attribute
Status	Polynominal	Attribute
Pendidikan Terakhir	Polynominal	Attribute
Jurusan	Polynominal	Attribute
Tahun Lulus	Date_time	Attribute

### 2.3 Metode Klasifikasi

- Klasifikasi adalah kegiatan atau pekerjaan mengelompokkan data atau mempelajari sesuatu menurut ciri-ciri yang dimilikinya. Ini dapat dilakukan agar setiap objek diberi kategori tertentu. [9]. Klasifikasi adalah pekerjaan yang melakukan pembuatan sebuah model dengan dasar data latih yang telah disiapkan, selanjutnya dengan model tersebut akan dilakukan klasifikasi data yang baru, tujuannya adalah agar sistem yang dibuat akan dapat melakukan pengklasifikasian data terhadap semua data set dengan baik dan benar, selanjutnya keberhasilan hasilnya akan diukur setelah polanya terbentuk [10]. Adapun tahapan pengklasifikasian pada penelitian ini akan dilakukan menggunakan 2(dua) algoritma, yakni sebagai berikut :
- Naïve bayes, merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Teorema bayes dapat ditulis menggunakan persamaan 1 [11]:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Keterangan :

$P(A|B)$  : Probabilitas A terjadi dengan bukti bahwa B telah terjadi (probabilitas superior)

$P(B|A)$  : Probabilitas B terjadi dengan bukti bahwa A telah terjadi

$P(A)$  : Kesempatan terjadinya A

$P(B)$  : Kesempatan terjadinya B

Teorema Bayes adalah rumus matematika yang digunakan untuk mengetahui probabilitas. Ini dapat digunakan untuk memprediksi hal-hal seperti apakah akan turun hujan besok, berdasarkan faktor lain seperti suhu, kelembapan, dan tekanan. Teorema Bayes mengatakan bahwa kita dapat mengetahui seberapa besar kemungkinan sesuatu didasarkan pada seberapa sering hal itu terjadi di masa lalu [12]. Pada penelitian sebelumnya pernah dilakukan



penelitian oleh M. Y. Putra and D. I. Putri, 2022 [13] yang bertujuan untuk mengklasifikasi jurusan siswa sekolah dengan menggunakan Algoritma Naïve Bayes dan K-Nearest Neighbor. Data siswa digunakan dalam penelitian untuk membandingkan akurasi algoritma Naïve Bayes dan algoritma K-Nearest Neighbor menggunakan Aplikasi Rapidminer. Algoritma Naïve Bayes ternyata lebih akurat dalam mengklasifikasikan kelas dengan jurusan IPA dan IPS, yang dapat memudahkan pengurus sekolah dalam mengolah data kelas XI. Hasil penelitiannya bahwa dengan algoritma Naïve Bayes sebesar 81.82% dengan data sebanyak 55 baris data dari 277 baris data, sedangkan analisa dengan Algoritma K-Nearest Neighbor hasil akurasi lebih tinggi yakni sebesar 92.73% dengan data yang sama. Kedua algoritma menghasilkan hasil yang berbeda, dengan algoritma K-Nearest Neighbor lebih akurat. Hal ini dikarenakan nilai alpha berada di bawah cutoff 0,05 yang berarti tidak terdapat perbedaan yang signifikan antara kedua algoritma tersebut. Dari analisis tersebut dapat disimpulkan bahwa algoritma K-Nearest Neighbor lebih akurat dalam mengklasifikasikan jurusan siswa IPA dan IPS di kelas XI. Ini berarti bahwa jurusan siswa akan didasarkan pada keahlian dan kemampuan mereka, bukan hanya ditentukan oleh kategori mana siswa tersebut masuk. Penelitian ini diharapkan dapat membantu siswa dalam menentukan jurusan sehingga dan memberikan manfaat kepada pihak sekolah memcepat kinerja dalam hal penjurusan siswa dan baik serta tepat dalam pengklasifikasiannya.

- c. Decision tree merupakan suatu struktur yang digunakan untuk membantu proses pengambilan keputusan yang terdiri dari simpul internal, tepi dan simpul daun. Setiap node keputusan disebut tes pada atribut atau subset dari atribut, dan masing-masing sisi mewakili nilai tertentu atau rentang nilai input atribut [14]. Pengklasifikasi Pohon Keputusan lebih akurat daripada metode klasifikasi lainnya dalam hal mengidentifikasi item serupa [15]. Decision tree merupakan teknik prediksi yang dapat digunakan untuk klasifikasi dan prediksi tugas. Decision tree menggunakan teknik “membagi dan menaklukkan” untuk membagi ruang pencarian masalah menjadi himpunan masalah [16] Decision tree adalah algoritma machine learning yang menggunakan seperangkat aturan untuk membuat keputusan dengan struktur seperti pohon yang memodelkan kemungkinan hasil, biaya sumber daya, utilitas dan kemungkinan konsekuensi atau resiko.[17].

### 3. HASIL DAN PEMBAHASAN

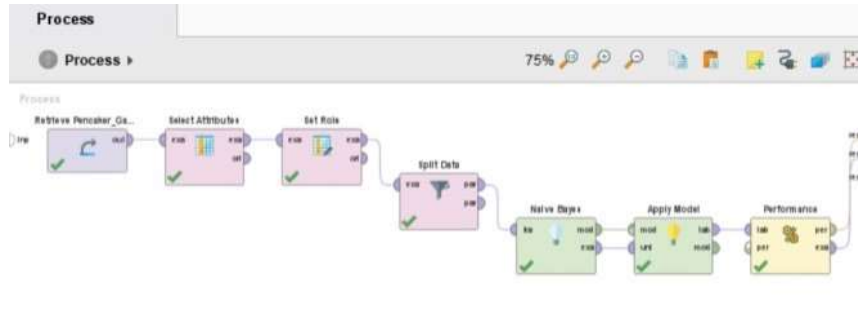
Analisa data pencari kerja akan dilakukan menggunakan 2 algoritma yakni algoritma naïve bayes dan decision tree menggunakan aplikasi rapid miner. Pertama akan dilakukan impor dataset kedalam aplikasi rapidminer kemudian kita tentukan label datanya. Label data yang digunakan adalah minat pencari ke luar negeri.

#### 3.1.1 Data Preparation Phase

Untuk melakukan penelitian ini, Anda perlu mengumpulkan data dari sebanyak mungkin orang. Untuk ini, Anda perlu mengumpulkan informasi dari orang-orang yang telah bepergian ke berbagai negara. *Dataset* yang akan digunakan menggunakan format file csv agar tools Rapid Miner lebih mudah untuk membaca file.

#### 3.1.1 Algoritma Naïve Bayes

Pada penelitian ini proses klasifikasi akan menggunakan algoritma naïve bayes dilakukan dengan proses sebagai berikut, dataset diimpor kedalam rapidminer kemudian dilakukan set attribute yakni untuk memilih atribut mana yang akan digunakan untuk mendukung penelitian. Kemudian ditambahkan operator set role untuk menentukan label attribute. Kemudian ditambahkan operator split data, pada tahap ini untuk menentukan berapa jumlah data yang akan dijadikan data training dan data testing, penulis menggunakan perbandingan 0,7:0,3. Selanjutnya data training akan diuji menggunakan algoritma naïve bayes, hasil pembentukan model ini dapat diketahui faktor apa saja yang sangat mempengaruhi peminat kerja ke luar negeri sebuah start-up. Pemodelan klasifikasi menggunakan Algoritma Naïve Bayes yang dilakukan pada penelitian ini terlihat pada Gambar 2 berikut.



Gambar 2. Proses klasifikasi menggunakan Algoritma Naïve Bayes

Pada gambar proses diatas, dataset diuji menggunakan algoritma naïve bayes kemudian apply model dan dilanjutkan proses performance. Berikut ini adalah hasil klasifikasi menggunakan Algoritma Naïve Bayes, terlihat pada gambar 4 dan gambar 5.

classification\_error: 0.36%

	True Ya	True Tidak	class precision
pred Ya	2513	21	99.17%
pred Tidak	8	5483	99.85%
class recall	99.68%	99.52%	

Gambar 3. Hasil Classification Error menggunakan Algoritma Naïve Bayes

Hasil Classification error sebesar 0.36%, menandakan pengujian ini sangat kecil kemungkinan kesalahan.

```
PerformanceVector
PerformanceVector:
accuracy: 99.64%
ConfusionMatrix:
True: Ya Tidak
Ya: 2513 21
Tidak: 8 5483
classification_error: 0.36%
ConfusionMatrix:
True: Ya Tidak
Ya: 2513 21
Tidak: 8 5483
```

Gambar 4. Hasil Performance Vector

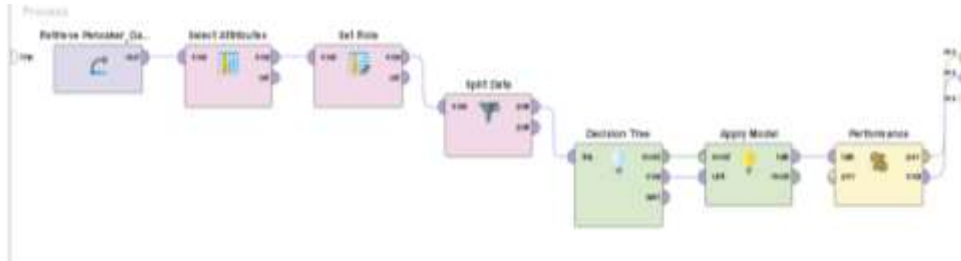
Dari kedua gambar diatas dapat dilihat bahwa nilai accuracy pada klasifikasi data pencari kerja menggunakan algoritma naïve bayes sangat baik dengan hasil sebesar 99,64% dengan hasil yang berminat bekerja keluar negeri sebanyak 2513 orang dan yang tidak berminat bekerja ke luar negeri sebanyak 5483 orang, sedangkan untuk classification error sebesar 0,36 %.

### 3.1.1 Algoritma Dcision Tree

Setelah dilakukan pengujian menggunakan algoritma naïve bayes, selanjutnya dilakukan pengujian menggunakan algoritma dcision tree. Proses Klasifikasi menggunakan algoritma dcision tree dilakukan dengan proses sebagai berikut, dataset diimpor kedalam rapidminer kemudian dilakukan set attribute yakni untuk memilih atribut mana yang akan digunakan untuk mendukung penelitian. Kemudian ditambahkan operator set role untuk menentukan label attribute. Kemudian ditambahkan operator split data, pada tahap ini untuk menentukan berapa jumlah data yang akan



dijadikan data training dan data testing, penulis menggunakan perbandingan 0,7:0,3. Selanjutnya data training akan diuji menggunakan algoritma decision tree, hasil pembentukan model ini dapat diketahui faktor apa saja yang sangat mempengaruhi peminat kerja ke luar negeri. Pemodelan yang dilakukan pada penelitian klasifikasi dataset pencari kerja menggunakan algoritma Decision Tree terlihat pada Gambar 5 berikut.



**Gambar 5.** Proses Klasifikasi menggunakan Algoritma Decision

Pada proses tersebut terlihat bahwa dataset ditentukan atributnya kemudian di lakukan set role untuk menentukan table kemudian diuji menggunakan algoritma Decision Tree dilanjutkan Apply model dan operator performance, sehingga akan muncul hasil performance seperti yang terlihat pada gambar 6, sebagai berikut :

accuracy: 68.59%

	true Ya	true Tidak	class precision
pred. Ya	0	0	0.00%
pred. Tidak	2521	5504	68.59%
class recall	0.00%	100.00%	

**Gambar 6.** Hasil Pengujian Klasifikasi menggunakan Algoritma Decision

Setelah melalui rangkaian proses diatas, maka didapatkan hasil performance seperti terlihat pada gambar 7 dan gambar 8 berikut ini.

classification\_error: 31.41%

	true Ya	true Tidak	class precision
pred. Ya	0	0	0.00%
pred. Tidak	2521	5504	68.59%
class recall	0.00%	100.00%	

**Gambar 7.** Hasil Classification Error menggunakan Algoritma Decision Tree

```

PerformanceVector

PerformanceVector:
accuracy: 68.59%
ConfusionMatrix:
True:  Ya    Tidak
Ya:    0     0
Tidak: 2521  5504
classification_error: 31.41%
ConfusionMatrix:
True:  Ya    Tidak
Ya:    0     0
Tidak: 2521  5504
    
```



**Gambar 7.** Hasil Accuracy Performance menggunakan Algoritma Decision Tree

Hasil Classification error sebesar 31,41%, menandakan pengujian ini kemungkinan banyak terjadi kesalahan. Sedangkan hasil accuracy performance menggunakan algoritma decision tree adalah 68.59 % dengan jumlah pencari kerja yang berminat bekerja ke luar negeri sebanyak 2521 dan yang tidak berminat bekerja keluar negeri sebanyak 5504.

#### 4. KESIMPULAN

Berdasarkan hasil penelitian atau analisa dataset pencari kerja yang berminat untuk bekerja di luar negeri dan pencari kerja yang tidak berminat ke luar negeri menggunakan aplikasi rapidminer dengan metode Klasifikasi dengan algoritma Naïve Bayes dihasilkan accuracy performance tertinggi sebesar 99,64 %, Class recall sebesar 99,68% dan classification error sebesar 0,36%, artinya sangat kecil terjadi error, sedangkan accuracy performance menggunakan algoritma decision tree sebesar 68.59 % dengan jumlah pencari kerja yang berminat bekerja ke luar negeri sebanyak 2521 dan yang tidak berminat bekerja keluar negeri sebanyak 5504 dan hasil classification errornya sebesar 31,41 %.

Hal tersebut membuktikan bahwa naïve bayes sangat sesuai untuk mengklasifikasi dataset pencari kerja, hal ini dapat dibuktikan dari hasil penelitian saat ini yang memiliki tingkat akurasi yang cukup baik. Hasil performace atau penelitian ini dapat digunakan sebagai acuan pemerintah khususnya di Provinsi Jawa Tengah untuk mempersiapkan lowongan kerja baik dalam negeri dan di luar negeri.

#### UCAPAN TERIMAKASIH

Diucapkan terimakasih kepada Allah SWT atas segala karunianya sehingga penulis dapat menyelesaikan jurnal ini, terimakasih kepada Dosen Magister Teknologi Informasi Universitas Teknologi Yogyakarta dan kepada Admin Jurnal Teknoinfo atas diterimanya jurnal ini, semoga bermanfaat.

#### DAFTAR PUSTAKA

- [1] Rian Tinages, "Pahami Metode Decision Tree Sebagai Algoritma Data Science," *Dqlab.Id.* 2022. [Online]. Available: <https://dqlab.id/pahami-metode-decision-tree-sebagai-algoritma-data-science>
- [2] Badan Pusat statistik Jawa Tengah, "BPS Provinsi Jawa Tengah," *Badan Pusat statistik Jawa Tengah.* 2021. <https://jateng.bps.go.id/statictable/2021/04/09/2239/produksi-perkebunan-rakyat-menurut-jenis-tanaman-di-jawa-tengah-tahun-2016-2020-ton-html>
- [3] Dinas Tenaga Kerja Provinsi Jawa Tengah, "E-Makaryo, Cari Kerja dari Rumah," *Disnakertrans Provinsi Jawa Tengah.* 2021. <https://disnakertrans.jatengprov.go.id/publik/berita/NjlyOTMyMDAxZDBjYzE1YzQ3ZTNmMTFhM2VmYzMyMzAwN2ZjMjdIMDM1NDlhNWE0ODdmMjZhMjNIZGNINGY->
- [4] H. F. Putro, R. T. Vlandari, and 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.
- [5] A. L. Maukar, F. Marisa, and A. A. Widodo, "Analisis Data Penerimaan Mahasiswa Baru Berbasis K-Means," *JIKO (Jurnal Inform. dan Komputer)*, vol. 6, no. 2, p. 142, 2022, doi: 10.26798/jiko.v6i2.558.
- [6] R. Annisa, "Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Penderita Penyakit Jantung," *J. Tek. Inform. Kaputama*, vol. 3, no. 1, pp. 22–28, 2019, [Online]. Available: <https://jurnal.kaputama.ac.id/index.php/JTIK/article/view/141/156>
- [5] A. P. Permana, K. Ainiyah, and K. F. H. Holle, "Analisis Perbandingan Algoritma Decision Tree, kNN, dan Naive Bayes untuk Prediksi Kesuksesan Start-up," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 6, no. 3, pp. 178–188, 2021, doi: 10.14421/jiska.2021.6.3.178-188.
- [8] R. Yendra, L. Marifni, and I. Suryani, "Klasifikasi Data Mining Untuk Seleksi Penerimaan Calon Pegawai Negeri Sipil Tahun 2017 Menggunakan Metode Naïve Bayes," *J. Sains Mat. dan Stat.*, vol. 6, no. 1, p. 65, 2020, doi: 10.24014/jsms.v6i1.9254.
- [9] R. E. Putri, Suparti, and R. Rahmawati, "Perbandingan Metode Klasifikasi Naïve Bayes Dan K-Nearest Neighbor Pada Analisis Data Status Kerja Di Kabupaten Demak Tahun 2012," *J. Gaussian*, vol. 3, no. 4, pp. 831–838, 2014.
- [10] B. A. Mukhtar, N. A. Setiawan, and T. B. Adji, "Analisis Perbandingan Tingkat Akurasi Algoritma Naive Bayes Classifier dengan Correlated-Naive Bayes Classifier," *Semin. Nas. Teknol. Inf. dan Multimed.* 2015, pp. 49–54, 2015.
- [11] Pratiwi, R. W., Nugroho, Y. S., 2016, Prediksi Rating Film Menggunakan Metode Naïve Bayes, *Jurnal Teknik Elektro*, No. 2, Vol. 8, Hal. 60 – 63.





- [12] S. Rahayu, J. Jaya Purnama, H. Mahmud Nawawi, F. Septia Nugraha, I. Komputer, and S. Nusa Mandiri Jakarta, “Algoritma Naïve Bayes Classifier Untuk Memprediksi Gejala Autism Spectrum Disorders Pada Anak-Anak,” *Semin. Nas. Rekayasa dan Teknol.*, vol. 27, no. November, 2019.
- [13] M. Y. Putra and D. I. Putri, “Pemanfaatan Algoritma Naïve Bayes dan K-Nearest Neighbor Untuk Klasifikasi Jurusan Siswa Kelas XI,” *J. Tekno Kompak*, vol. 16, no. 2, pp. 176–187, 2022.
- [14] H. M. Nawawi, S. Rahayu, M. J. Shidiq, and J. J. Purnama, “Algoritma C4.5 Untuk Memprediksi Pengambilan Keputusan Memilih Deposito Berjangka,” *J. Techno Nuasa Mandiri*, vol. 16, no. 1, pp. 65–72, 2019.
- [15] J. J. Purnama, H. M. Nawawi, S. Rosyida, Ridwansyah, and Risandar, “Klasifikasi Mahasiswa Her Berbasis Algoritma Svm Dan Decision Classification of Her Students Based on Svm Algorithm and,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 6, pp. 1253–1260, 2020, doi: 10.25126/jtiik.202073080.
- [16] Dunham, Margareth H., “Data Mining Introductory and Advanced Topics”. New Jersey:Prentice Hall. 2003.
- [17] Basuki, Ahmad dan Syarif, Iwan, “Decision Tree”. Surabaya: Politeknik Elektronika Negeri, 2003