

Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Data Penerimaan Pegawai Outsourcing

The application of naïve bayes algorithm to classification of outsourcing employee acceptance

Dahri Yani Hakim Tanjung^{a,1}, Fina Nasari^{b,2}, Rofiqoh Dewi^{c,3}, Wiwi Verina^{d,4}, Khairul Ummi^{e,5}, Muhammad Ihsan^{f,6}

^{a,b} Program Studi Sistem Informasi, Universitas Potensi Utama

Jl.K.L. Yos Sudarso KM 6.5 Tanjung Mulia - Medan

notashapire@gmail.com¹, finanasari@gmail.com², dezie.wie@gmail.com³, wiwiverina.azzahra@gmail.com⁴,
ummigibmie@gmail.com⁵, okm.ihsan@yahoo.co.id⁶

ABSTRAK

Mempekerjakan karyawan dalam ikatan kerja outsourcing nampaknya sedang menjadi trend saat ini, Banyak perusahaan outsourcing yakni perusahaan yang bergerak di bidang penyedia tenaga kerja aktif menawarkan jasa tenaga kerja Dengan kata lain, perusahaan menyetujui untuk melakukan beberapa tugas dengan perusahaan lain. Studi sebelumnya menunjukkan bahwa outsourcing adalah memindahkan beberapa fungsi perusahaan kepada pihak lain, termasuk pengambilan keputusan yang diatur dalam kontrak.. Penelitian ini bertujuan untuk memberikan evaluasi terhadap data penerimaan pegawai outsourcing dengan metode klasifikasi dari data mining. Diharapkan metode klasifikasi data mining mampu mengolah data, membuat analisis penerimaan karyawan outsourcing lebih mudah. Metode klasifikasi data mining Naive Bayes digunakan, dan proses pengujiannya menggunakan alat Rapid Miner untuk mengevaluasi akurasi teknik yang digunakan. Adapun beberapa atribut yang dipakai adalah Usia, wawancara, pengalaman dan akademis. Dimana dari hasil perhitungan terhadap data uji menggunakan algoritma naïve bayes menghasilkan nilai probabilitas sebesar 0,83 untuk klasifikasi Tidak Layak dan dilakukan pengujian dengan tools rapidminer menghasilkan nilai akurasi sebesar 98,18% berada pada class tidak layak, Sedangkan untuk nilai AUC (Area Under Curve) adalah 0.875 yang berada pada class tidak layak, di mana tingkat akurasi meningkat jika nilai AUC mendekati angka 1. Hasil ini menunjukkan bahwa teknik Naive Bayes dapat membantu dalam prediksi. penerimaan pegawai outsourcing dengan nilai probabilitas dan akurasi yang tinggi dengan keputusan yang dihasilkan adalah tidak layak.

Kata Kunci : Data Mining, Klasifikasi, Naïve Bayes, Penerimaan Pegawai Outsourcing.

ABSTRACT

Hiring employees in outsourced work ties seems to be a trend nowadays. Many outsourcing companies, namely companies engaged in the field of labor providers, actively offer labor services. In other words, companies agree to do several tasks with other companies. Previous studies have shown that outsourcing is transferring some of the company's functions to other parties, including decision-making stipulated in the contract. This study aims to provide an evaluation of outsourcing employee recruitment data using the classification method of data mining. It is hoped that the data mining classification method is able to process data, making the analysis of outsourced hiring easier. The Naive Bayes data mining classification method is used, and the testing process uses the Rapid Miner tool to evaluate the accuracy of the techniques used. Some of the attributes used are age, interview, experience and academics. Where from the results of calculations on test data using the naïve Bayes algorithm it produces a probability value of 0.83 for the Ineligible classification and testing with rapidminer tools produces an accuracy value of 98.18% in the ineligible class, As for the AUC value (Area Under Curve) is 0.875 which is in the inappropriate class, where the level of accuracy increases if the AUC value is close to 1. These results indicate that the Naive Bayes technique can help in prediction. hiring outsourced employees with high probability and accuracy with the resulting decision is not feasible.

Keywords: Data Mining, Classification, Naïve Bayes, Outsourcing Employee Acceptance.

Info Artikel :

Disubmit:12 April 2023

Direview:17 Mei 2023

Diterima :18 Juni 2023

1. PENDAHULUAN

Ketenagakerjaan merupakan masalah yang sangat vital karena terkait langsung dengan kesejahteraan hidup masyarakat. Oleh karena itu, ketenagakerjaan tidak dapat dilepaskan dari bidang yang lain seperti sosial, politik, ekonomi, keamanan, dan budaya. Dalam menjalankan bisnis sebuah perusahaan perlu menemukan cara yang kuat untuk menghadapi perkembangan ekonomi global dan kemajuan teknologi yang cepat guna mempertahankan kelangsungan hidupnya. Keberhasilan perusahaan sangat tergantung pada banyak faktor diantaranya adalah produktivitas, efisiensi biaya, dan manajemen strategis perusahaan yang baik. Selain sumber daya alam dan teknologi, tenaga kerja, juga dikenal sebagai sumber daya manusia, memainkan peran penting dalam pertumbuhan ekonomi[1][2]

Pemilik dan pemimpin perusahaan, baik milik negara maupun swasta, tampaknya telah mengadopsi praktik mempekerjakan karyawan dalam ikatan kerja outsourcing. Banyak perusahaan outsourcing, atau penyedia tenaga kerja aktif, menawarkan tenaga kerja kepada pemberi kerja, sehingga bisnis yang memerlukan tenaga kerja tidak perlu susah-susah mencari, memilih, dan melatih karyawan[3] Dengan kata lain, perusahaan menyetujui untuk melakukan beberapa tugas dengan perusahaan lain. Studi sebelumnya menunjukkan bahwa outsourcing adalah memindahkan beberapa fungsi perusahaan kepada pihak lain, termasuk pengambilan keputusan yang diatur dalam kontrak[4]

Karyawan kontrak adalah karyawan yang diberi bantuan untuk menyelesaikan tugas biasa perusahaan, tetapi tidak ada jaminan bahwa mereka akan tetap bekerja. Masa kerja karyawan kontrak ditentukan oleh seberapa baik prestasi mereka; jika prestasi mereka lebih baik, perusahaan akan mempertahankan karyawan tersebut, tetapi jika prestasi mereka tidak meningkat, perusahaan akan memberhentikannya.. Namun permasalahan yang di alami oleh perusahaan dalam penerimaan pegawai outsourcing adalah sistem yang berjalan pada perusahaan masih bersifat walk in interview atau wawancara secara langsung dan banyak nya pegawai outsourcing yang melamar dan menyebabkan Manager bagian SDM sulit menentukan dalam penerimaan pegawai outsourcing. Selain itu, untuk mengambil keputusan penerimaan pegawai outsourcing bukan hanya melihat berdasarkan riwayat pendidikan yang dimiliki tetapi juga harus memiliki pengalaman dan kesiapan dalam menghadapi dunia kerja.

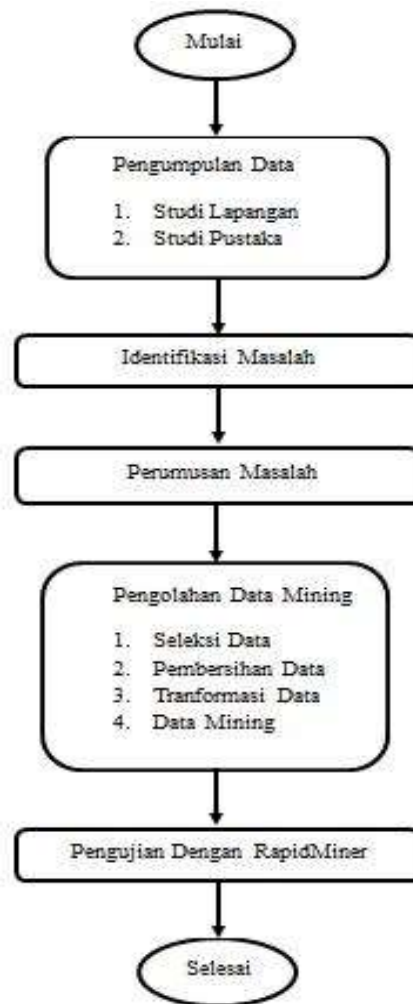
Adapun metode problem solving dari penelitian ini menggunakan model data mining yang menerapkan metode algoritma naïve bayes dan RapidMiner yang akan di gunakan sebagai tools untuk pengujian akurasi Algoritma naïve bayes terhadap data penerimaan pegawai outsourcing. Menurut algoritma naïve bayes, nilai variabel kelas adalah semua variabel bebas. Metode ini hanya memerlukan jumlah data pelatihan untuk menentukan parameter yang diperlukan untuk proses klasifikasi. Selain itu, algoritma Naive Bayes seringkali bekerja lebih baik dalam situasi dunia nyata yang lebih kompleks dari yang diantisipasi. Selain itu, untuk menentukan estimasi parameter proses klasifikasi, algoritma Naive Bayes tidak membutuhkan banyak data pelatihan[5]. Menurut penelitian sebelumnya bahwa algoritma naïve bayes mampu dalam mengklasifikasikan data pemberian bantuan surat keterangan tidak mampu dengan tingkat akurasi 62.86 %, Selanjutnya, algoritma Naive Bayes dapat digunakan untuk memprediksi masa studi siswa. Hasil pengujian menunjukkan bahwa model ini memiliki nilai akurasi rata-rata sebesar 68%, dan dataset yang digunakan untuk membuat model dapat memengaruhi nilai akurasi[6][7].

Selanjutnya akan dilakukan pengujian terhadap data penerimaan karyawan dengan menggunakan model cross validasi. Dengan membagi seluruh data menjadi dua set pelatihan dan tes, cross-validation digunakan untuk menguji keefektifan model pembelajaran mesin [8] kemudian menggunakan kurva ROC (Receiver operating characteristic) yang digunakan untuk mempelajari kemampuan karakteristik tertentu dalam diskriminan, Salah satu pengukuran yang paling populer berdasarkan kurva ROC adalah

Area di bawah kurva (AUC/Area Under Curve). AUC memiliki keuntungan dari menyediakan indeks tunggal yang meringkas kinerja keseluruhan dari tes (atau aturan) berdasarkan pada karakteristik berkelanjutan, bukan seluruh kurva, dan sangat berguna untuk perbandingan di bawah populasi yang berbeda atau tes yang berbeda. Oleh karena itu, jika nilai AUC yang dekat dengan 1 menunjukkan ketepatan tes yang tinggi, sedangkan ketepatannya yang sangat rendah sesuai dengan nilai yang lebih dekat dengan 0,5 [9]

2. METODE

Dalam melakukan penelitian ini, penulis juga menggunakan metode khususnya saat mengumpulkan data dengan menggunakan kerangka kerja penelitian. Adapun kerangka kerja penelitian dapat dilihat pada gambar 1 berikut.



Gambar 1. Kerangka Kerja Penelitian

Tahapan penelitian pada Gambar 1 menggambarkan proses penelitian yang akan di kerjakan penulis serta menggambarkan penelitian secara keseluruhan. Dimana Tahapan keseluruhan dapat dijelaskan sebagai berikut yaitu :

- 1) Pengumpulan Data
Pada tahap ini adalah proses pengumpulan data, dimana data yang diperoleh adalah data penerimaan karyawan outsourcing pada PT. PLN (persero) Up3 serta data yang berhubungan dengan judul berupa artikel atau jurnal-jurnal terkait.
- 2) Identifikasi Masalah
Pada dasarnya semua penelitian akan selalu didahului dengan identifikasi masalah, hal ini berguna untuk peneliti agar fokus pada titik permasalahan. Sebelum menentukan rumusan masalahnya.
- 3) Rumusan masalah
Untuk rumusan masalahnya yang sebelumnya telah dikemukakan pada bagian latar belakang.
- 4) Pengolahan data mining
Setelah sudah dirumuskan masalah tahap penelitian selanjutnya Seleksi Data: Penulis melakukan seleksi data berdasarkan kebutuhan untuk melakukan analisis; ini adalah salah satu proses yang dibutuhkan dalam pengolahan data mining pada tahap ini. Pembersihan Data (Preprocessing): Jika ada data yang tidak valid atau tidak sesuai, penulis membersihkan data. Transformasi Data, penulis melakukan perubahan data atau menggabungkan data ke dalam format yang sesuai untuk di proses dalam bentuk mining. Data mining, pada proses yang terakhir ini penulis melakukan proses analisa dengan menggunakan algoritma naïve bayes untuk memprediksi penerimaan pegawai outsourcing
- 5) Pengujian dengan RapidMiner
Pada tahap ini penulis melakukan ujicoba system yang telah dilakukan pada proses data mining menggunakan Tools RapidMiner dengan model validasi yang digunakan adalah cross validation[10]

3. HASIL DAN PEMBAHASAN

Pada bab ini membahas mengenai penerapan algoritma naïve bayes terhadap data penelitian dan pengujian dari algoritma naïve bayes menggunakan software pengujian rapidminer.

A. Penerapan Algoritma Naïve Bayes

Untuk proses pengolahan data dibutuhkan sejumlah dataset yang digunakan untuk pengujian serta analisis dari algoritma yang digunakan. Dimana dataset yang digunakan diambil dari data calon pegawai outsourcing yang melamar.

Berdasarkan dataset yang ada berikut ini adalah tabel sample data dan nilai dari seluruh atribut :

1) Usia

Usia diklasifikasikan menjadi 3 bagian yaitu untuk Usia “ $> 22 \times > = 28$ Tahun” diklasifikasikan sebagai Usia “Bagus”, untuk usia “ $> 28 \times > = 35$ Tahun” diklasifikasikan sebagai Usia “Kurang Bagus” dan untuk Usia “ > 35 Tahun” dikalsifikasikan sebagai Usia Tidak Bagus”. Hasil klasifikasi Usia terlihat pada tabel 1

Tabel 1. Kriteria Usia

Kriteria Usia	Keterangan
Bagus	$> 22 \times > = 28$ Tahun
Sedang	$> 28 \times > = 35$ Tahun
Tidak Bagus	> 35 Tahun

2) Pengalaman

Pengalaman diklasifikasikan menjadi 3 bagian yaitu untuk Pengalaman “ $x > 2$ Tahun” diklasifikasikan sebagai pengalaman “Berpengalaman”, untuk Pengalaman “ $> 1 \times > = 2$ Tahun” diklasifikasikan sebagai pengalaman “Kurang Berpengalaman” dan untuk Pengalaman “ $= 0$ Tahun”

diklasifikasikan sebagai Pengalaman “Tidak Berpengalaman”. Hasil klasifikasi Pengalaman terlihat pada tabel 2

Tabel 2. Kriteria Pengalaman

Kriteria Pengalaman	Keterangan
Berpengalaman	$x > 2$ Tahun
Kurang Berpengalaman	$> 1 \ x \geq 2$ Tahun
Tidak Berpengalaman	$= 0$ Tahun

3) Tes Akademis

diklasifikasikan menjadi 3 bagian yaitu untuk Tes Akademis “ $>84 \ x \leq 100$ ” diklasifikasikan sebagai Tes Akademis “Bagus”, untuk Tes Akademis $>70 \ x \geq 84$ ” diklasifikasikan sebagai Tes Akademis “Kurang Bagus” dan untuk Tes Akademis “ ≤ 70 ” diklasifikasikan sebagai Tes Akademis “Tidak Bagus”. Hasil klasifikasi Tes Akademis terlihat pada tabel 3

Tabel 3. Kriteria Tes Akademis

Kriteria Tes Akademis	Keterangan
Bagus	$>84 \ x \leq 100$
Kurang Bagus	$>70 \ x \geq 84$
Tidak Bagus	≤ 70

4) Wawancara

Wawancara diklasifikasikan menjadi 3 bagian yaitu untuk Wawancara “ $>84 \ x \leq 100$ ” diklasifikasikan sebagai Wawancara “Bagus”, untuk usia “ $>70 \ x \geq 84$ ” diklasifikasikan sebagai Wawancara “Kurang Bagus” dan untuk Wawancara “ ≤ 70 ” diklasifikasikan sebagai Wawancara “Tidak Bagus”. Hasil klasifikasi wawancara terlihat pada tabel 4

Tabel 4. Kriteria Wawancara

Kriteria Tes Akademis	Keterangan
Bagus	$>84 \ x \leq 100$
Kurang Bagus	$>70 \ x \geq 84$
Tidak Bagus	≤ 70

Berdasarkan dari tabel nilai seluruh atribut, maka untuk memenuhi atribut yang dibutuhkan untuk melakukan perhitungan dengan algoritma naïve bayes, dapat dilihat pada Tabel 5 dan table 6 untuk data testing.

Tabel 5. Data Training

Tes Akademis	Pengalaman	Usia	Wawancara	Prediksi
Bagus	Tidak Berpengalaman	Kurang Bagus	Bagus	Tidak

Tes Akademis	Pengalaman	Usia	Wawancara	Prediksi
Kurang Bagus	Berpengalaman	Bagus	Bagus	Layak
Bagus	Berpengalaman	Kurang Bagus	Kurang Bagus	Tidak
Kurang Bagus	Tidak Berpengalaman	Kurang Bagus	Bagus	Tidak
Bagus	Berpengalaman	Bagus	Bagus	Layak
Kurang Bagus	Tidak Berpengalaman	Bagus	Kurang Bagus	Tidak
Bagus	Berpengalaman	Kurang Bagus	Bagus	Layak
Bagus	Berpengalaman	Bagus	Bagus	Layak
Kurang Bagus	Tidak Berpengalaman	Kurang Bagus	Bagus	Tidak
Bagus	Berpengalaman	Bagus	Kurang Bagus	Layak
Bagus	Berpengalaman	Bagus	Bagus	Layak
Bagus	Berpengalaman	Bagus	Bagus	Layak
Bagus	Berpengalaman	Bagus	Kurang Bagus	Layak
Kurang Bagus	Tidak Berpengalaman	Bagus	Kurang Bagus	Tidak
Bagus	Berpengalaman	Bagus	Kurang Bagus	Layak
Bagus	Berpengalaman	Bagus	Bagus	Layak
Kurang Bagus	Berpengalaman	Bagus	Bagus	Layak
Kurang Bagus	Tidak Berpengalaman	Kurang Bagus	Bagus	Tidak
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
Kurang Bagus	Tidak Berpengalaman	Bagus	Bagus	Tidak
Bagus	Berpengalaman	Bagus	Bagus	Layak
Kurang Bagus	Berpengalaman	Bagus	Bagus	Layak

Dari data training yang ditunjukkan pada tabel 1 di atas, pengujian akan dilakukan dengan data pengujian sebagai berikut:

Tabel 6. Data Testing

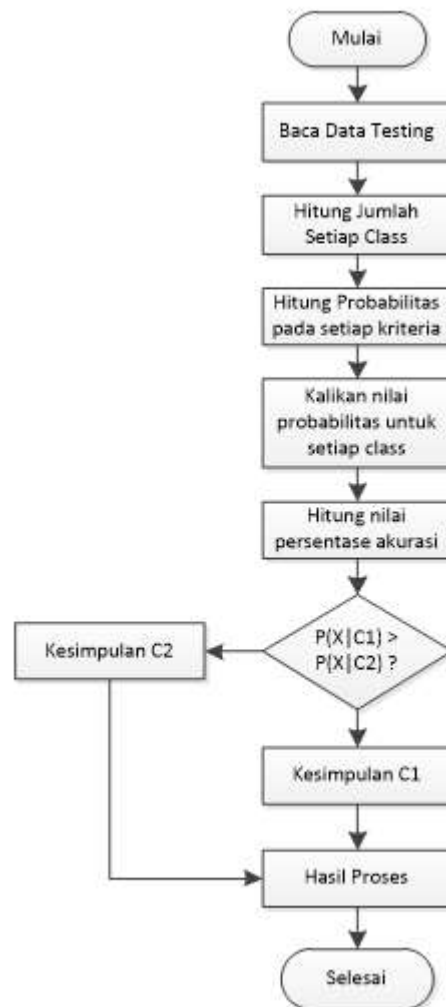
Tes Akademis	Pengalaman	Usia	Wawancara	Prediksi
Kurang Bagus	Tidak Berpengalaman	Bagus	Bagus	?

Ada beberapa kriteria yang pakai untuk pengolahan data dalam penelitian ini yaitu:

- 1) Tes Akademik: variabel tes akademik dikategorikan menjadi Bagus, Kurang Bagus, dan Tidak Bagus

- 2) Pengalaman: variabel pengalaman dikategorikan menjadi Berpengalaman dan Tidak Berpengalaman
- 3) Usia: variabel usia dikategorikan menjadi Bagus, Sedang, dan Tidak Bagus
- 4) Wawancara: variabel wawancara dikategorikan menjadi Bagus, Kurang Bagus, dan Tidak Bagus

Gambar 3 berikut menunjukkan flowchart proses penerapan metode naive bayes.



Gambar 3. Flowchart Algoritma Naïve Bayes

Hasil klasifikasi dapat dilakukan perhitungan sebagai berikut:

Menghitung Probabilitas pada setiap kriteria

- 1) Tes Akademik
 $P(X_1 = \text{Kurang Bagus} \mid \text{Klasifikasi} = \text{Layak})$
 $P(X_1 = 10/43) = 0,23$
 $P(X_1 = \text{Kurang Bagus} \mid \text{Klasifikasi} = \text{Tidak Layak})$
 $P(X_1 = 10/12) = 0,83$

- 2) Pengalaman

$P(X_2 = \text{Tidak Berpengalaman} \mid \text{Klasifikasi} = \text{Layak})$
 $P(X_2 = 5/43) = 0,15$
 $P(X_2 = \text{Tidak Berpengalaman} \mid \text{Klasifikasi} = \text{Tidak Layak})$
 $P(X_2 = 12/12) = 1$

- 3) Usia
- $P(X_3 = \text{Bagus} \mid \text{Klasifikasi} = \text{Layak})$
 $P(X_3 = 43/43) = 1$
 $P(X_3 = \text{Bagus} \mid \text{Klasifikasi} = \text{Tidak Layak})$
 $P(X_3 = 12/12) = 1$
- 4) Wawancara
- $P(X_4 = \text{Bagus} \mid \text{Klasifikasi} = \text{Layak})$
 $P(X_4 = 43/43) = 1$
 $P(X_4 = \text{Bagus} \mid \text{Klasifikasi} = \text{Tidak Layak})$
 $P(X_4 = 12/12) = 1$

Menghitung nilai probabilitas setiap class

- 1) $P(C_1) = \frac{\text{Classprediksi} = \text{Layak}}{\text{Jumlah data}}$
 $P(\text{Layak}) = 43/55 = 0,78$
- 2) $P(C_2) = \frac{\text{Classprediksi} = \text{Tidak Layak}}{\text{Jumlah data}}$
 $P(\text{Tidak Layak}) = 12/55 = 0,22$

Menghitung semua nilai dari setiap klasifikasi Layak dan Tidak Layak

- 1) $P(X \mid \text{Klasifikasi} = \text{Layak})$
 $= 0,23 * 0,15 * 1 * 1$
 $= 0,034$
- 2) $P(X \mid \text{Klasifikasi} = \text{Tidak Layak})$
 $= 0,83 * 1 * 1 * 1 = 0,83$

Perhitungan akhir untuk setiap klasifikasi menunjukkan bahwa probabilitas $P(X \mid \text{Klasifikasi} = \text{Tidak Layak})$ lebih besar dari $P(X \mid \text{Klasifikasi} = \text{Layak})$, yaitu 0,83 banding 0,034. Oleh karena itu, prediksi penerimaan pegawai tersebut adalah Tidak Layak.

B. Pengujian Algoritma Naïve Bayes dengan Rapidminer

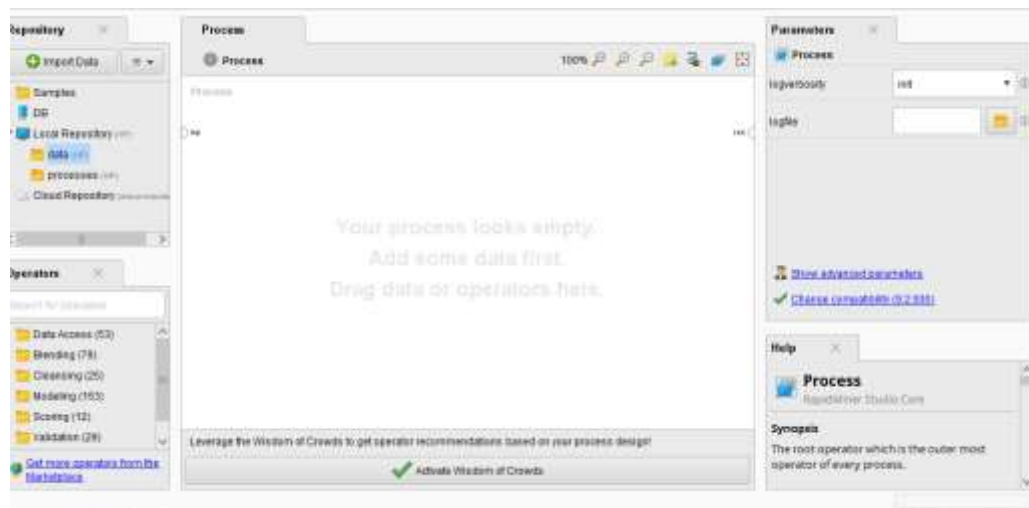
Pada bagian ini akan membahas mengenai pengujian algoritma naïve bayes menggunakan Rapidminer terhadap data calon pegawai outsourcing. Rapidminer adalah software open-source yang menawarkan lingkungan pengembangan aplikasi yang cepat dan visualisasi data yang sesuai, Baik untuk analisis dan simulasi data mining [11]. Adapun tahapan dari tool rapidminer terdiri dari beberapa proses yaitu:

- 1) Membuat data kedalam excel agar dapat terbaca di aplikasi *Rapidminer*, dapat dilihat pada gambar dibawah ini

HIMPUNAN	Tipe	Jumlah	Rasio	Lantai
R1C1	R1C2	R1C3	R1C4	R1C5
R2C1	R2C2	R2C3	R2C4	R2C5
R3C1	R3C2	R3C3	R3C4	R3C5
R4C1	R4C2	R4C3	R4C4	R4C5
R5C1	R5C2	R5C3	R5C4	R5C5
R6C1	R6C2	R6C3	R6C4	R6C5
R7C1	R7C2	R7C3	R7C4	R7C5
R8C1	R8C2	R8C3	R8C4	R8C5
R9C1	R9C2	R9C3	R9C4	R9C5
R10C1	R10C2	R10C3	R10C4	R10C5
R11C1	R11C2	R11C3	R11C4	R11C5
R12C1	R12C2	R12C3	R12C4	R12C5
R13C1	R13C2	R13C3	R13C4	R13C5
R14C1	R14C2	R14C3	R14C4	R14C5
R15C1	R15C2	R15C3	R15C4	R15C5
R16C1	R16C2	R16C3	R16C4	R16C5
R17C1	R17C2	R17C3	R17C4	R17C5
R18C1	R18C2	R18C3	R18C4	R18C5
R19C1	R19C2	R19C3	R19C4	R19C5
R20C1	R20C2	R20C3	R20C4	R20C5
R21C1	R21C2	R21C3	R21C4	R21C5
R22C1	R22C2	R22C3	R22C4	R22C5
R23C1	R23C2	R23C3	R23C4	R23C5
R24C1	R24C2	R24C3	R24C4	R24C5
R25C1	R25C2	R25C3	R25C4	R25C5
R26C1	R26C2	R26C3	R26C4	R26C5
R27C1	R27C2	R27C3	R27C4	R27C5
R28C1	R28C2	R28C3	R28C4	R28C5
R29C1	R29C2	R29C3	R29C4	R29C5
R30C1	R30C2	R30C3	R30C4	R30C5
R31C1	R31C2	R31C3	R31C4	R31C5
R32C1	R32C2	R32C3	R32C4	R32C5
R33C1	R33C2	R33C3	R33C4	R33C5
R34C1	R34C2	R34C3	R34C4	R34C5
R35C1	R35C2	R35C3	R35C4	R35C5
R36C1	R36C2	R36C3	R36C4	R36C5
R37C1	R37C2	R37C3	R37C4	R37C5
R38C1	R38C2	R38C3	R38C4	R38C5
R39C1	R39C2	R39C3	R39C4	R39C5
R40C1	R40C2	R40C3	R40C4	R40C5
R41C1	R41C2	R41C3	R41C4	R41C5
R42C1	R42C2	R42C3	R42C4	R42C5
R43C1	R43C2	R43C3	R43C4	R43C5
R44C1	R44C2	R44C3	R44C4	R44C5
R45C1	R45C2	R45C3	R45C4	R45C5
R46C1	R46C2	R46C3	R46C4	R46C5
R47C1	R47C2	R47C3	R47C4	R47C5
R48C1	R48C2	R48C3	R48C4	R48C5
R49C1	R49C2	R49C3	R49C4	R49C5
R50C1	R50C2	R50C3	R50C4	R50C5
R51C1	R51C2	R51C3	R51C4	R51C5
R52C1	R52C2	R52C3	R52C4	R52C5
R53C1	R53C2	R53C3	R53C4	R53C5
R54C1	R54C2	R54C3	R54C4	R54C5
R55C1	R55C2	R55C3	R55C4	R55C5
R56C1	R56C2	R56C3	R56C4	R56C5
R57C1	R57C2	R57C3	R57C4	R57C5
R58C1	R58C2	R58C3	R58C4	R58C5
R59C1	R59C2	R59C3	R59C4	R59C5
R60C1	R60C2	R60C3	R60C4	R60C5
R61C1	R61C2	R61C3	R61C4	R61C5
R62C1	R62C2	R62C3	R62C4	R62C5
R63C1	R63C2	R63C3	R63C4	R63C5
R64C1	R64C2	R64C3	R64C4	R64C5
R65C1	R65C2	R65C3	R65C4	R65C5
R66C1	R66C2	R66C3	R66C4	R66C5
R67C1	R67C2	R67C3	R67C4	R67C5
R68C1	R68C2	R68C3	R68C4	R68C5
R69C1	R69C2	R69C3	R69C4	R69C5
R70C1	R70C2	R70C3	R70C4	R70C5
R71C1	R71C2	R71C3	R71C4	R71C5
R72C1	R72C2	R72C3	R72C4	R72C5
R73C1	R73C2	R73C3	R73C4	R73C5
R74C1	R74C2	R74C3	R74C4	R74C5
R75C1	R75C2	R75C3	R75C4	R75C5
R76C1	R76C2	R76C3	R76C4	R76C5
R77C1	R77C2	R77C3	R77C4	R77C5
R78C1	R78C2	R78C3	R78C4	R78C5
R79C1	R79C2	R79C3	R79C4	R79C5
R80C1	R80C2	R80C3	R80C4	R80C5
R81C1	R81C2	R81C3	R81C4	R81C5
R82C1	R82C2	R82C3	R82C4	R82C5
R83C1	R83C2	R83C3	R83C4	R83C5
R84C1	R84C2	R84C3	R84C4	R84C5
R85C1	R85C2	R85C3	R85C4	R85C5
R86C1	R86C2	R86C3	R86C4	R86C5
R87C1	R87C2	R87C3	R87C4	R87C5
R88C1	R88C2	R88C3	R88C4	R88C5
R89C1	R89C2	R89C3	R89C4	R89C5
R90C1	R90C2	R90C3	R90C4	R90C5
R91C1	R91C2	R91C3	R91C4	R91C5
R92C1	R92C2	R92C3	R92C4	R92C5
R93C1	R93C2	R93C3	R93C4	R93C5
R94C1	R94C2	R94C3	R94C4	R94C5
R95C1	R95C2	R95C3	R95C4	R95C5
R96C1	R96C2	R96C3	R96C4	R96C5
R97C1	R97C2	R97C3	R97C4	R97C5
R98C1	R98C2	R98C3	R98C4	R98C5
R99C1	R99C2	R99C3	R99C4	R99C5
R100C1	R100C2	R100C3	R100C4	R100C5

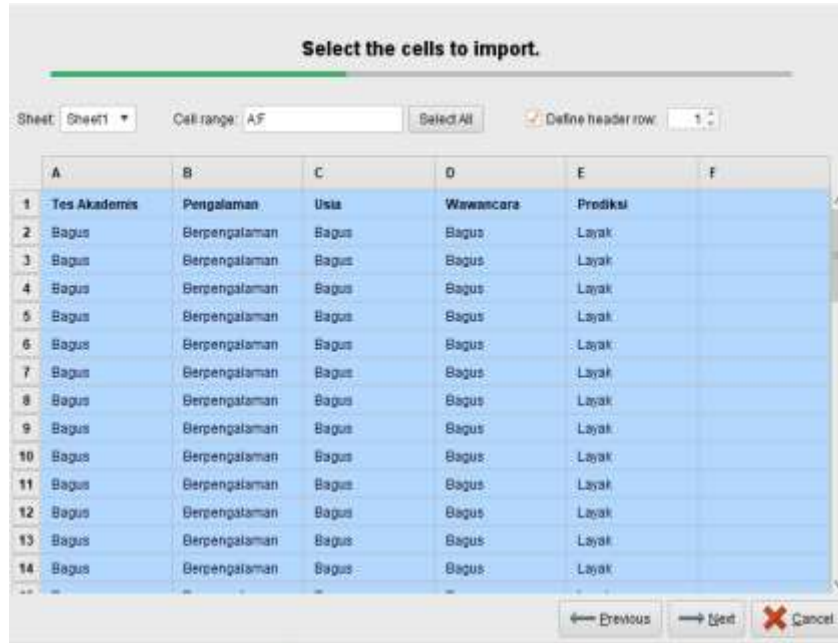
Gambar 4. DataSet Bentuk Excel Sheet

2) Jalankan aplikasi Rapidminer



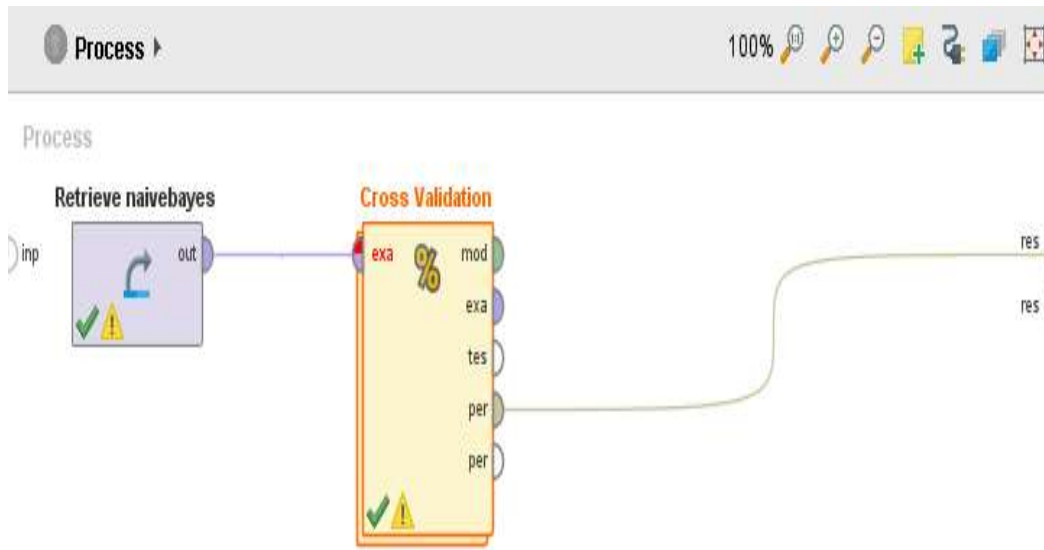
Gambar 5. Tampilan Awal Rapidminer

- 3) Pilih file yang akan di *import* , kemudian atur tiap atribut dari data excel yang telah di import



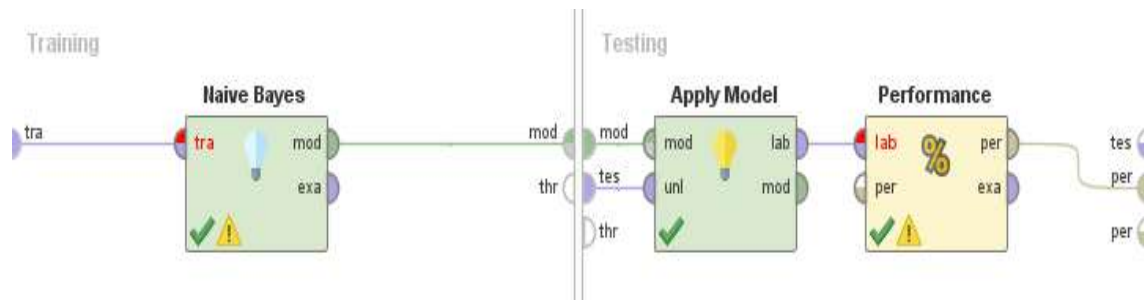
Gambar 6. Tampilan *Setting Attribute*

- 4) Menampilkan tahap-tahap individual operator di dalam proses analisa dan juga interkoneksi di antara mereka, dengan memilih *option modeling*



Gambar 7. Tampilan *Koneksi Individual*

- 5) Gambar 8 menunjukkan hasil uji algoritma Naive Bayes pada mesin pemecah batu cepat. Model dan akurasi yang dihasilkan akan dipengaruhi oleh operator dan parameter yang digunakan.



Gambar 8. Pengujian Validasi Naïve Bayes

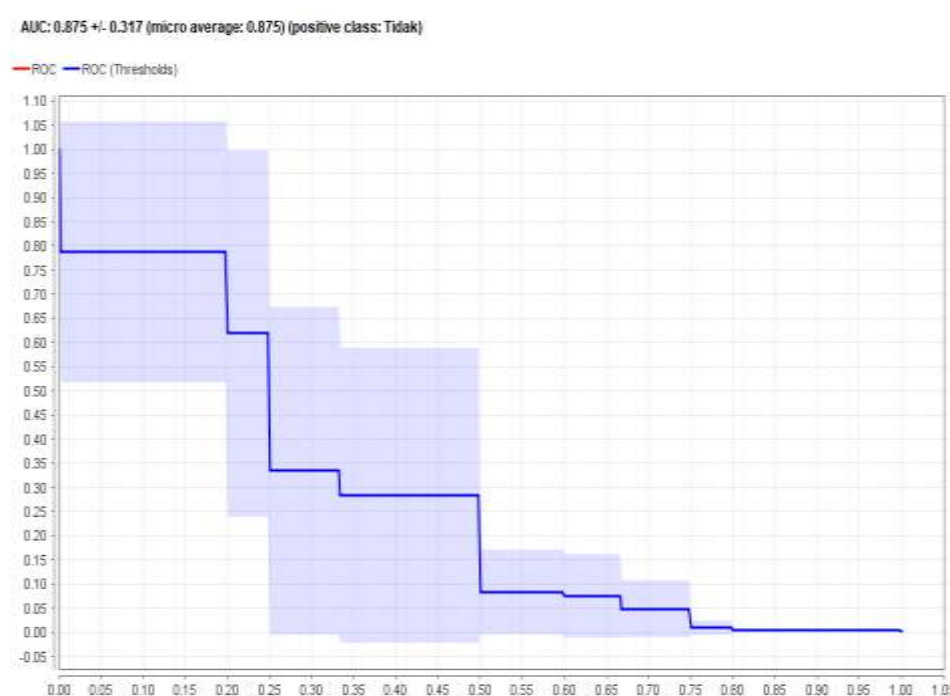
Dari hasil pengujian yang telah dilakukan dengan menggunakan model cross validation dapat dilihat bahwa nilai akurasi yang dihasilkan dengan algoritma naïve bayes cukup baik yaitu 98,18% yang dapat dilihat pada gambar 9 berikut

accuracy: 98.33% +/- 5.27% (micro average: 98.18%)

	true Layak	true Tidak	class precision
pred. Layak	42	1	97.67%
pred. Tidak	0	12	100.00%
class recall	100.00%	92.31%	

Gambar 9. Perfomance Vector Algoritma Naïve Bayes

Sedangkan hasil dari *Area Under Curve* (AUC) dapat dilihat pada gambar 9 berikut ini



Gambar 10. Hasil AUC (*Area Under Curve*) dengan Model Cross Validasi

Nilai AUC yang dihasilkan adalah 0,875 berada pada class tidak layak, dimana hasil akurasi untuk grafik AUC akan semakin bagus apabila nilai AUC mendekati angka 1. Selain nilai akurasi dan AUC yang dihasilkan dapat diketahui juga bahwa nilai recall dan precision yang dihasilkan adalah 100% untuk nilai precision dan 92,31% untuk nilai recall berada pada class tidak layak.

4. KESIMPULAN

Berdasarkan pembahasan dari bab-bab sebelumnya yang telah dilakukan maka dapat diambil beberapa kesimpulan sebagai berikut :

- 1) Algoritma Naïve Bayes mampu mengklasifikasikan data penerimaan karyawan dimana dari 55 data, 54 data berhasil diklasifikasikan dengan nilai probabilitas yaitu 0,83 untuk klasifikasi Tidak Layak dimana persentase akurasi adalah 98,18%
- 2) Pengujian yang dilakukan menggunakan model cross validasi menghasilkan nilai akurasi 98,18% , nilai recall 92,31%, nilai presisi 100% dan nilai AUC sebesar 0,875

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Potensi Utama yang telah banyak memberikan dukungan dalam penelitian ini.

REFERENSI

- [1] J. Fitriyaningrum, "Implementasi Sistem Alih Daya atau Outsourcing Dalam Mencapai Kesejahteraan Pekerja Indonesia Ditinjau Dari Undang-Undang Nomor 13 Tahun 2003," *Indones. State Law Rev.*, vol. 2, no. 1, pp. 88–102, 2019.
- [2] Triyono, "Outsourcing dalam Perspektif Pekerja dan Pengusaha," *J. Kependud. Indones.*, vol. 6, no. 1, pp. 45–62, 2011, [Online]. Available: <https://ejurnal.kependudukan.lipi.go.id/index.php/jki/article/download/88/159>
- [3] I. Widayastuti, "2490-Article Text-5879-1-10-20200831," *J. Manaj. Bisnis*, vol. 17, no. 3, pp. 326–341, 2020.
- [4] N. M. B. Sintya Devi, B. A. Noer, and Y. Rahmawati, "Analisis Perbandingan Pegawai Tetap dan Pegawai Outsourcing Ditinjau dari Pemberdayaan, Kepuasan Kerja, dan Komitmen Organisasional," *J. Sains dan Seni ITS*, vol. 6, no. 2, pp. 2–6, 2017, doi: 10.12962/j23373520.v6i2.26119.
- [5] D. Y. Hakim Tanjung, "Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Data Pengisian ATM," *Infosys (Information Syst. J.)*, vol. 7, no. 1, p. 12, 2022, doi: 10.22303/infosys.7.1.2022.12-24.
- [6] N. Riyanah, "Penerapan Algoritma Naive Bayes Untuk Klasifikasi Penerima Bantuan Surat Keterangan Tidak Mampu (Implementation of Algorithms Naïve Bayes for Classification Recipients Help Letter Description Not Able)," vol. 2, no. 4, pp. 206–213, 2021.
- [7] I. W. Saputro and B. W. Sari, "Uji Performa Algoritma Naïve Bayes untuk Prediksi Masa Studi Mahasiswa," *Creat. Inf. Technol. J.*, vol. 6, no. 1, p. 1, 2020, doi: 10.24076/citec.2019v6i1.178.
- [8] D. Y. H. Tanjung and Fujiati, "Analisis Perbandingan Algoritma ID3 Dan C4.5 Terhadap Data Pengisian Uang ATM," pp. 231–242, 2021, [Online]. Available: <http://dx.doi.org/10.22303/csrid.13.3a.2021.231-242>
- [9] G. Cortese, "Estimating the Area under the ROC Curve with Modified Profile Likelihoods," *Int. J. Stat. Probab.*, vol. 6, no. 1, p. 1, 2016, doi: 10.5539/ijsp.v6n1p1.
- [10] K. and S. A. Rajesh, "Analysis of SEER Dataset for Breast Cancer Diagnosis using C4.5 Classification

Algorithm,” *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 1, no. 2, pp. 72–77, 2012, [Online]. Available: www.ijarcce.com

- [11] F. Syahputra, H. Hartono, and R. Rosnelly, “Penerapan Algoritma C4.5 Dalam Memprediksi Ketersediaan Uang Pada Mesin ATM,” *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 556, 2021, doi: 10.30865/mib.v5i2.2933.