

Implementasi Data Mining Dalam Klasifikasi Tingkat Kesenjangan Kompetensi PNS Menggunakan Metode Naive Bayes

Putra Kurniawan, Wasilah*, Sutedi, Handoyo Widi Nugroho

Ilmu Komputer, Magister Teknik Informatika, IIB Darmajaya, Bandar Lampung, Indonesia

Email: putra.kr18@gmail.com, wasilah@darmajaya.ac.id, sutedi@darmajaya.ac.id, handoyo.wn@darmajaya.ac.id

Email Penulis Korespondensi: wasilah@darmajaya.ac.id

Submitted: 20/07/2024; Accepted: 08/09/2024; Published: 09/09/2024

Abstrak—Aparatur Sipil Negara (ASN) berperan sebagai pelaksana kebijakan publik, pelayan masyarakat, dan pemersatu bangsa, dengan peningkatan kualitas dan efisiensi layanan publik menjadi prioritas utama pemerintah. Pada studi kasus Pemerintah Provinsi Lampung, perencanaan peningkatan kompetensi Aparatur Sipil Negara (ASN) menjadi salah satu prioritas kegiatan saat ini, dikarenakan belum adanya data acuan dalam penentuan pengembangan kompetensi pada masing-masing ASN. *Assesment Center* merupakan salah satu cara dalam penentuan tingkat kompetensi ASN, namun dalam pelaksanaannya terkendala beberapa faktor seperti keterbatasan dana, keterbatasan waktu hingga keterbatasan jumlah tenaga *assesor*. Berdasarkan hasil penilaian Indeks Sistem Merit tahun 2023 oleh Komite Aparatur Sipil Negara (KASN), didapatkan rekomendasi bahwa dalam pemetaan dan evaluasi kesenjangan kompetensi pegawai dapat dilaksanakan melalui *Human Capital Development Plan* (HCDP). Pada pelaksanaannya digunakan metode *self-assessment* dengan kuisioner berbasis kamus kompetensi dari Peraturan Menteri PANRB Nomor 38 Tahun 2017 digunakan untuk mengatasi kendala *assessment centre*. Kuisioner di khususkan kepada Pegawai Negeri Sipil (PNS) tenaga teknis di lingkungan Pemerintah Provinsi Lampung. Analisis data kuisioner ini menghasilkan klasifikasi ASN berdasarkan tingkat kesenjangan kompetensi (tidak ada, rendah, sedang, tinggi). Pada penelitian ini dilakukan pengujian hasil kasifikasi menggunakan salah satu teknik pengklasifikasian data mining, yaitu metode Naive Bayes. Tujuan dari penelitian ini, yaitu melakukan pengujian terhadap performa algoritma Naive Bayes dalam melakukan klasifikasi tingkat kesenjangan kompetensi PNS. Dari hasil penelitian yang dilakukan didapatkan kesimpulan bahwa, sistem klasifikasi tingkat kesenjangan kompetensi PNS Pemerintah Provinsi Lampung dapat dimodelkan dan berdasarkan hasil pengujian model dengan mengimplementasikan metode klasifikasi Naive Bayes menggunakan tools rapidminer terhadap dataset pada objek penelitian diperoleh tingkat akurasi sebesar 98,02%, dengan kesimpulan bahwa algoritma Naive Bayes dapat berjalan dengan baik dalam proses klasifikasi tingkat kesenjangan kompetensi PNS. Dengan tingkat akurasi yang didapatkan, hasil klasifikasi yang terbentuk dapat dimanfaatkan oleh Pemerintah Provinsi Lampung dalam perencanaan kebutuhan pengembangan kompetensi PNS.

Kata Kunci: Kesenjangan Kompetensi; Klasifikasi; Naive Bayes; *Data Mining*; *Assesment Centre*

Abstract—Civil Servants (Aparatur Sipil Negara or ASN) play crucial roles as implementers of public policy, community service providers, and national unifiers. The government's primary focus is on enhancing the quality and efficiency of public services. In the Provincial Government of Lampung, planning for the enhancement of the competencies of Civil Servants (Aparatur Sipil Negara or ASN) has become a current priority activity. This emphasis is due to the absence of reference data for determining competency development for each ASN. The Assessment Center is one method for determining the competency level of Civil Servants (ASN). However, its implementation faces several challenges such as budget constraints, time limitations, and a shortage of assessors. Based on the results of the 2023 Merit System Index assessment by the Civil Service Commission (KASN), it was recommended that mapping and evaluating employee competency gaps can be carried out through the Human Capital Development Plan (HCDP). In its implementation, a self-assessment method using a questionnaire based on the competency dictionary from the Regulation of the Minister of Administrative and Bureaucratic Reform No. 38 of 2017 is used to address the constraints of the assessment center. The questionnaire is specifically targeted at technical civil servants (PNS) in the Lampung Provincial Government. The analysis of this questionnaire data produces a classification of civil servants based on the level of competency gaps (none, low, medium, high). In this study, the classification results are tested using one of the data mining classification techniques, namely the Naive Bayes method. The objective of this research is to evaluate the performance of the Naive Bayes algorithm in classifying the levels of competency gaps among civil servants. Based on the research findings, it can be concluded that the classification system for competency gap levels among civil servants in the Lampung Province Government can be modeled. The testing of the model, which implemented the Naive Bayes classification method using RapidMiner tools on the research dataset, achieved an accuracy rate of 98.02%. The conclusion is that the Naive Bayes algorithm performs well in classifying the competency gap levels among civil servants. With the achieved accuracy level, the resulting classifications can be utilized by the Lampung Provincial Government in planning the development needs of civil servant competencies.

Keywords: Competency gap; Classification; Naive Bayes; Data Mining; Assesment Centre

1. PENDAHULUAN

ASN memiliki fungsi sebagai pelaksana kebijakan publik, pelayan publik, dan pemererat dan pemersatu bangsa. Peningkatan kualitas dan efisiensi layanan publik menjadi fokus utama bagi pemerintah dalam menjalankan tugasnya. Dalam mencapai upaya tersebut, peningkatan kompetensi Aparatur Sipil Negara (ASN) menjadi hal yang mendesak. Pada Pemerintah Provinsi Lampung, pemetaan dan evaluasi kesenjangan (*Gap*) kompetensi ASN merupakan langkah utama dalam menentukan rencana pengembangan kompetensi pada masing-masing ASN, setiap ASN diharapkan dapat ditempatkan tingkat kesenjangan kompetensinya dalam kelas kesenjangan rendah, kesenjangan sedang, kesenjangan tinggi atau tidak ada kesenjangan.

Pada study kasus Pemerintah Provinsi Lampung, isu strategis yang ditemukan yaitu belum adanya dokumen perencanaan pengembangan kompetensi ASN. Dengan keadaan tersebut ditemukan suatu permasalahan dalam pengembangan kompetensi ASN yaitu tidak tepatnya jenis pengembangan kompetensi sesuai dengan kinerja dan jabatan pada masing-masing ASN. Oleh sebab itu, untuk menentukan perencanaan pengembangan kompetensi pegawai dibutuhkan data terkait Tingkat Kesenjangan Kompetensi ASN sehingga dapat direncanakan jenis pengembangan kompetensi yang tepat bagi masing-masing ASN. Berdasarkan hasil penilaian Indeks Sistem Merit dalam Manajemen Aparatur Sipil Negara tahun 2023, didapat rekomendasi oleh Komite Aparatur Sipil Negara (KASN) untuk membentuk strategi pengembangan kompetensi berupa *Human Capital Development Plan* (HCDP)[1].

Untuk mendapatkan data kesenjangan kompetensi ASN, dapat dilakukan analisis dengan metode *assessment centre* pada seluruh ASN yang melibatkan multi asesi, multi asesor dan multi simulasi[2] Terdapat beberapa kendala yang melatar-belakangi ketidak-terlaksanakannya kegiatan *assessment* pegawai, mulai dari ketersediaan tenaga Fungsional Assesor Kepegawaian pada Pemerintah Provinsi Lampung, biaya yang besar untuk kegiatan *assessment* pada seluruh pegawai hingga keterbatasan waktu pelaksanaan kegiatan *assessment*. Berkaitan dengan hal tersebut, dengan berjalannya HCDP dapat membantu dalam penentuan Tingkat kesenjangan kompetensi ASN.

Dalam penelitian ini HCDP membantu dalam mempersiapkan data perilaku kerja ASN, Dimana data tersebut dapat menentukan tingkat kesenjangan kompetensi ASN. Metode *self-assessment* dalam pengumpulan data perilaku kerja pegawai menjadi cara yang lebih sederhana untuk mengatasi kendala pelaksanaan kegiatan *assessment center*, proses *self-assessment* dilaksanakan dengan cara mendistribusikan kuisioner kepada seluruh pegawai yang bertugas di lingkungan Pemerintah Provinsi Lampung dengan materi kuisioner yang disesuaikan dengan kamus kompetensi manajerial dan sosial kultural yang tercantum dalam Peraturan Menteri PANRB Nomor 38 Tahun 2017 tentang Standar Kompetensi Jabatan ASN[3], Dari data kuisioner yang didapatkan, akan tersedia dataset terkait perilaku kerja ASN yang dapat menentukan Tingkat Kesenjangan Kompetensi ASN. ASN terdiri dari Pegawai Negeri Sipil (PNS) dan Pegawai Pemerintah dengan Perjanjian Kerja (PPPK), pada penelitian ini akan dibatasi terhadap jenis kompetensi Manajerial dan Sosial Kultural pada PNS.

Dalam konteks ini, pemetaan tingkat kesenjangan kompetensi PNS memerlukan pendekatan yang sistematis dan efektif. Setiap PNS akan dikelompokkan dalam jenis kesenjangan kompetensi antara lain: tidak ada kesenjangan, kesenjangan rendah, kesenjangan sedang dan kesenjangan tinggi. Salah satu metode yang dapat digunakan adalah metode klasifikasi, yang bertujuan untuk mengelompokkan entitas berdasarkan atribut-atribut tertentu. Dalam hal ini, Berdasarkan penelitian-penelitian sebelumnya, untuk melakukan klasifikasi terhadap suatu objek untuk mempermudah pengambilan Keputusan pada perencanaan suatu kegiatan, metode Naïve Bayes merupakan teknik yang populer dan efisien, Naïve Bayes menerapkan teknik klasifikasi dengan menggunakan probabilitas bersyarat [4][5][6][7].

Penggunaan metode Naïve Bayes terbukti dapat memecahkan permasalahan dalam hal pengklasifikasian sebuah objek. Dalam penelitian yang dilakukan oleh Angga Saputra Dinata, Nisar di tahun 2023, penerapan Metode Naïve bayes dalam pengadaan jenis buku referensi, dapat dengan baik memberikan rekomendasi kepada petugas pengadaan dengan mengelompokkan jenis buku berdasarkan kriteria “ya” atau “tidak” untuk dilakukan pengadaan[8]. Pada penelitian-penelitian lainnya, algoritma Naïve Bayes juga digunakan dalam mengklasifikasikan tingkat kelancaran pembayaran sewa teras UMKM, pada penelitian tersebut beberapa probabilitas dapat menghitung keakuratan prediksinya, diperoleh persentasi akurasi sebesar 81,81%[9]. Algoritma Naïve Bayes melakukan proses klasifikasi dengan metode probabilitas dan statistik, yaitu menghitung peluang di masa depan dengan cara membandingkan pengalaman di masa sebelumnya[10]. Penelitian ini akan menggunakan metode yang sama dengan penelitian-penelitian sebelumnya, namun dengan variabel data yang berbeda, akan diuji apakah algoritma Naïve Bayes akan melakukan klasifikasi pada dataset kesenjangan kompetensi PNS dengan baik dan menghasilkan performa dengan tingkat akurasi yang lebih baik dari penelitian sebelumnya.

Pada penelitian ini data statistik yang menjadi acuan sebagai dataset pelatihan diambil dari statistik perilaku kerja pegawai dalam menentukan tingkat kesenjangan kompetensi manajerial dan sosiokultural. Penyiapan dataset ini berdasarkan kamus kompetensi yang tercantum dalam Peraturan Menteri PANRB Nomor 38 Tahun 2017 tentang Standar Kompetensi Jabatan ASN[3]. Penerapan kamus kompetensi manajerial dan sosial kultural ini pernah dilakukan pada penelitian oleh lusius aman pada tahun 2022, yang menyatakan kamus kompetensi yang tercantum dalam Peraturan Menteri PANRB Nomor 38 Tahun 2017 dapat digunakan pada dalam mengukur tingkat kompetensi pegawai, namun masih ada ketidak-cocokan pada beberapa jenis jabatan fungsional tertentu[11]. Dalam penelitian ini, dataset yang digunakan merupakan data PNS dengan Jabatan Teknis, untuk Jabatan Fungsional Guru tidak disertakan dalam data, karena memiliki tingkat pengembangan kompetensi tersendiri yang berbeda dengan Jabatan Teknis sehingga data dapat mengikuti sesuai kamus kompetensi manajerial dan sosial kultural.

Penelitian ini bertujuan untuk menguji penerapan metode Naive Bayes dalam klasifikasi tingkat kesenjangan (*Gap*) kompetensi PNS di Pemerintah Provinsi Lampung. Dengan menggunakan pendekatan ini, diharapkan akan diperoleh keakurasian tingkat Kesenjangan (*Gap*) kompetensi PNS. Hal ini menjadi langkah awal dalam meningkatkan efektivitas manajemen Sumber Daya Manusia Aparatur di Lingkungan Pemerintah Daerah dan secara luas.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Berdasar permasalahan yang telah dijabarkan di dalam pendahuluan, peneliti mencari referensi berupa berbagai literatur terkait. Selanjutnya, dapat ditemukan analisis permasalahan yang ada dan ditetapkan metode Naïve Bayes sebagai algoritma yang akan digunakan dalam pemecahan masalah klasifikasi Tingkat kesenjangan kompetensi PNS pada Pemerintah Provinsi Lampung[12]. Pemodelan klasifikasi tingkat kesenjangan kompetensi PNS dengan memanfaatkan algoritma Naive Bayes dijalankan dengan menggunakan tools rapidminer, pemodelan ini bertujuan untuk mendapatkan hasil klasifikasi tingkat kesenjangan kompetensi PNS yang akurat sesuai dengan perilaku kerja masing-masing PNS.

Dalam penelitian ini diterapkan 2 tahapan dalam metode penelitian, yaitu tahap pengumpulan data dan tahap pengolahan data[13]. Tahap pengelolaan data dilanjutkan dengan tahap pengujian implementasi algoritma naïve bayes pada proses klasifikasi menggunakan tools rapidminer, dari proses pengujian akan didapatkan hasil analisis dan Kesimpulan. Secara umum, langkah-langkah pada metodologi penelitian dapat tecermin dalam gambar 1 berikut ini.



Gambar 1. Tahapan Penelitian

2.2 Tahap Pengumpulan Data

Tahap ini merupakan tahap penyiapan *dataset* yang akan dijadikan data pelatihan dan data pengujian dengan algoritma Naïve Bayes. Data yang dikumpulkan, yaitu data perilaku kerja PNS yang sesuai kamus kompetensi dapat menentukan tingkat kesenjangan kompetensi PNS.

Data dihimpun dari hasil kuisioner berupa *form self-assessment* dari Seluruh PNS Provinsi Lampung. Instrumen kuisioner disusun berdasarkan kamus kompetensi sesuai Peraturan Menteri PANRB Nomor 38 Tahun 2017[3], untuk ketepatan data instrument, daftar pertanyaan disusun oleh pejabat fungsional asesor kepegawaian pada Badan Kepegawaian Daerah Provinsi Lampung. Penyebaran formulir *self-assessment* dilakukan melalui *link google form* yang akan di isi oleh masing-masing PNS. Data responses yang masuk akan menjadi dataset penelitian

2.3 Tahap Pengelolaan Data

Sebelum dilakukan pelatihan dan pengujian menggunakan metode Naïve Bayes, data *responses* dari kuisioner sebelumnya dilakukan pengelolaan data terlebih dahulu, untuk memilih atribut-atribut yang paling berpengaruh dalam penentuan klasifikasi kesenjangan kompetensi ASN. Maka dilakukan proses *selection* untuk menghilangkan atribut yang tidak diperlukan pada proses klasifikasi.

Selanjutnya, dilakukan *preprocessing* data, pada tahap ini bertujuan untuk membuang data duplikat, memeriksa inkonsisten data serta memperbaiki data yang salah. Proses *selection* dan *preprocessing* dijalankan menggunakan aplikasi Microsoft excel 2021. Data yang tersedia, yaitu data PNS Provinsi Lampung sebanyak 6.818 data.

2.4 Tahap Pengujian Naïve Bayes Dengan Rapidminer

Pada tahap ini akan dilakukan proses klasifikasi menggunakan metode Naïve Bayes. Perlengkapan yang digunakan, yaitu Laptop Asus Tuff A15, dengan spesifikasi *processor* AMD Ryzen 7 dengan RAM 16GB. *Tools* yang digunakan dalam penelitian ini menggunakan *software rapidminer*. *Rapidminer* merupakan salah satu *software* yang sering digunakan untuk melakukan pengolahan *data mining*[10].

Aturan Bayes merupakan sebuah teorema dalam teori probabilitas untuk menghitung probabilitas suatu hipotesis atau peristiwa (H) berdasarkan pada bukti atau data yang telah diamati (X). Aturan Bayes menggabungkan probabilitas awal dari hipotesis (sebelum bukti diamati) dengan probabilitas bukti di bawah hipotesis tersebut[14]. Teorema Naive Bayes menggunakan alat matematika syang digunakan dalam konsep probabilitas dalam mendeteksi kemungkinan menjadi sebuah hasil klasifikasi, konsep tersebut tergambar dalam rumus berikut ini[15]:

$$P(h|e) = \frac{P(e|h).P(h)}{P(e)} \quad (1)$$

Teorema Bayes adalah prinsip fundamental dalam teori probabilitas dan inferensi statistik yang digunakan untuk memperbarui keyakinan suatu hipotesis berdasarkan data baru. Dalam konteks ini, $P(h|e)$ mewakili probabilitas posterior, yaitu Probabilitas akhir dimana hipotesis h terjadi jika diberikan bukti e . $P(e|h)$ menunjukkan Probabilitas bukti e terjadi akan mempengaruhi hipotesis h . $P(h)$ adalah probabilitas prior, yang merupakan probabilitas awal dari hipotesis h tanpa memandang bukti e , hal ini merepresentasikan pengetahuan awal terhadap hipotesis h . Selanjutnya, $P(e)$ adalah probabilitas marginal, yaitu total probabilitas dari bukti e terjadi tanpa memandang hipotesis. Teorema Bayes menyatakan bahwa probabilitas posterior $P(h|e)$ dapat dihitung dengan mengalikan probabilitas *likelihood* $P(e|h)$ dengan probabilitas prior $P(h)$, kemudian membaginya dengan probabilitas marginal $P(e)$. Formula ini memungkinkan untuk memperbarui keyakinan tentang hipotesis h secara rasional berdasarkan bukti baru yang diperoleh.

Dataset yang telah dipersiapkan pada tahap sebelumnya akan dilakukan proses klasifikasi dengan menguji performa algoritma Naïve Bayes dengan *tools rapidminer*. *Dataset* terdiri dari 6.818 data 7 atribut termasuk atribut label kelas (atribut *output*), yaitu atribut kesenjangan[16]. Berdasarkan *dataset* yang tersedia, dibagi menjadi 80% data *training* dan 20% data *testing*. Penelitian ini bertujuan untuk mendapatkan hasil akurasi yang diberikan algoritma Naïve Bayes untuk sebuah permasalahan klasifikasi multi-class dalam memprediksi kesenjangan kompetensi ASN apakah masuk kedalam kelas kesenjangan tinggi, kesenjangan sedang, kesenjangan rendah, atau tidak ada kesenjangan.

2.5 Analisis

Setelah dilakukan pemodelan dengan menjalankan *training* terhadap 80% data, lalu dilakukan proses *testing* pada 20% data lainnya, maka akan didapatkan hasil klasifikasi terhadap *data testing* tersebut. Untuk melakukan validasi terhadap hasil klasifikasi tersebut, digunakan *confusion matrix*. *Confusion matrix* akan memberikan tingkat performa dari model berdasarkan objek benar atau salah[13].

Untuk menguji hasil *confusion matrix* yang ditampilkan, dilakukan perhitungan manual terhadap *Accuracy*, *Precision* dan *Recall* dengan perhitungan sebagai berikut.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+FP+TN+FN)} \times 100\% \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\% \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (4)$$

Pada konteks evaluasi kinerja model prediksi atau klasifikasi, ada empat jenis hasil yang dapat terjadi, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). *True Positive* (TP) adalah jumlah kasus di mana model memprediksi positif dan kenyataannya memang positif, menunjukkan bahwa model berhasil mendeteksi kasus positif dengan benar. *True Negative* (TN) adalah jumlah kasus di mana model memprediksi negatif dan kenyataannya memang negatif, yang berarti model berhasil mengidentifikasi kasus negatif dengan benar. *False Positive* (FP) terjadi ketika model memprediksi positif tetapi kenyataannya negatif, mengindikasikan bahwa model menghasilkan alarm palsu dengan mengidentifikasi kasus negatif sebagai positif. *False Negative* (FN) terjadi ketika model memprediksi negatif tetapi kenyataannya positif, menunjukkan bahwa model gagal mendeteksi kasus positif dan salah mengidentifikasinya sebagai negatif.

Akurasi adalah nilai seberapa sering suatu model memprediksi benar secara keseluruhan. Rumus (2) menghitung proporsi dari semua prediksi (baik positif maupun negatif) yang benar (TP + TN) dari total keseluruhan prediksi yang dibuat (TP + FP + TN + FN). *Precision* adalah nilai seberapa banyak dari prediksi positif yang benar-benar positif. Rumus (3) menunjukkan perhitungan proporsi dari prediksi positif yang benar (TP) terhadap total prediksi positif yang dibuat (TP + FP). Sedangkan, *Recall* adalah ukuran seberapa baik model dapat mendeteksi semua kasus positif yang sebenarnya. Rumus (4) menunjukkan perhitungan proporsi dari kasus positif yang benar-benar terdeteksi (TP) terhadap total kasus positif yang sebenarnya (TP + FN).

3. HASIL DAN PEMBAHASAN

Berdasarkan hasil analisis masalah, yaitu pada Pemerintah Provinsi Lampung dibutuhkan data kesenjangan kompetensi PNS, guna dijadikan dasar dalam penentuan pengembangan kompetensi PNS. Data Kesenjangan kompetensi yang diharapkan, yaitu hasil klasifikasi terhadap data perilaku kerja PNS. Bentuk klasifikasi, yaitu mengelompokkan PNS dalam kategori: Kesenjangan Tinggi, Kesenjangan Sedang, Kesenjangan Rendah, dan Tidak ada Kesenjangan.

Data Kesenjangan didapatkan dari perilaku kerja PNS sesuai dengan jabatan yang diduduki. Untuk mendapatkan data perilaku PNS, dibuat sebuah kuisioner dengan metode *self-assesment* di mana nantinya responden akan melakukan *assessment* terhadap diri sendiri terkait perilaku kerja yang telah dijalankan selama ini. Daftar pertanyaan dalam kuisioner ini didasarkan pada kamus kompetensi Manajerial dan Sosial Kultural yang terdapat pada Lampiran II dan Lampiran III Peraturan Menteri PANRB Republik Indonesia Nomor 38 Tahun 2017 tentang Standar

Kompetensi Jabatan Aparatur Sipil Negara. Sesuai dengan target penelitian, yaitu mendapatkan data kesenjangan kompetensi PNS, maka responden pada kuisioner ini adalah PNS Tenaga Teknis pada Pemerintah Provinsi Lampung.

Untuk memudahkan dalam pengumpulan data hasil *self-assessment*, maka kuisioner disajikan menggunakan aplikasi *google form*. Daftar pertanyaan akan disesuaikan dengan Jabatan pada masing-masing responden. Jenis pertanyaan akan dikelompokkan sesuai dengan jenis kompetensi yang di ujikan dan level jabatan PNS, ketentuan isi kuisioner dapat dilihat pada Tabel 1.

Tabel 1. Distribusi Pertanyaan Kuisioner

LEVEL JABATAN	JUMLAH PERTANYAAN								K2 S1	TOTAL
	K1									
	M1	M2	M3	M4	M5	M6	M7	M8		
LEVEL 1	3	3	3	3	3	3	3	3	3	27
LEVEL 2	3	3	3	3	3	3	3	3	3	27
LEVEL 2+	3	3	3	3	3	3	3	3	3	27
LEVEL 3	3	3	3	3	3	3	3	3	3	27
LEVEL 4	3	3	3	3	3	3	3	3	3	27
LEVEL 5	3	3	3	3	3	3	3	3	3	27
LEVEL 5+	3	3	3	3	3	3	3	3	3	27
TOTAL	21	21	21	21	21	21	21	21	21	189

Sesuai dengan kamus kompetensi, akan dicari Kesenjangan Kompetensi Manajerial dan Sosial Kultural yang secara berurutan disebutkan dalam tabel 1, yaitu K1 dan K2. Untuk Kompetensi Manajerial (K1) terbagi dalam 8 Kompetensi dan untuk Kompetensi Sosial Kultural (K2) terdapat 1 Kompetensi, yaitu:

- a. Integritas (M1),
- b. Kerja sama (M2),
- c. Komunikasi (M3)
- d. Orientasi pada hasil (M4),
- e. Pelayanan publik (M5),
- f. Pengembangan diri dan orang lain (M6),
- g. Mengelola perubahan (M7),
- h. Pengambilan Keputusan (M8),
- i. Perekat Bangsa (S1).

Untuk Level Jabatan terbagi dalam 5 level jabatan ditambah 2 level jabatan tambahan di mana materi kompetensi kedua jabatan tersebut beririsan antara 2 level jabatan yang berbeda. Jenis level jabatan tersebut diantaranya, Level 1 (*Awareness*), Level 2 (*Basic*), Level 3 (*Intermediate*), Level 4 (*Advance*), dan Level 5 (*Expert*)[2]. Jenis jabatan untuk masing-masing level antara lain:

- a. Level 1, yaitu untuk jabatan:
 1. Pelaksana
 2. Fungsional Pemula
- b. Level 2, yaitu untuk jabatan:
 1. Pengawas
 2. Fungsional Ahli Pertama
 3. Fungsional Mahir
- c. Level 3, yaitu untuk jabatan:
 1. Administrator
 2. Fungsional Ahli Muda
 3. Fungsional Penyelia
- d. Level 4, yaitu untuk jabatan:
 1. Pimpinan Tinggi Pratama
 2. Fungsional Ahli Madya
- e. Level 5, yaitu untuk jabatan:
 1. Pimpinan Tinggi Utama
 2. Pimpinan Tinggi Madya
- f. Level 2+, yaitu untuk jabatan Fungsional Terampil
- g. Level 5+, yaitu untuk jabatan Fungsional Ahli Utama

Untuk setiap kuisioner, responden akan memilih perilaku kerja dengan pilihan jawaban, yaitu Selalu, Sering, Kadang-kadang, Jarang, atau Tidak Pernah. Contoh pengisian kuisioner tertera pada Gambar 2.

**STANDAR KOMPETENSI MANAJERIAL DAN SOSIAL KULTURAL -
JABATAN PENGAWAS/JF AHLI PERTAMA/JF MAHIR**

Integritas (Level 2): *
Mampu mengingatkan, mengajak rekan kerja untuk bertindak sesuai nilai, norma, dan etika organisasi.
Indikator penilaian:

	Selalu	Sering	Kadang-kadang	Jarang	Tidak Pernah
Mengingatkan rekan kerja untuk bertindak sesuai dengan nilai, norma, dan etika organisasi dalam segala situasi dan kondisi; Mengajak orang lain untuk bertindak sesuai etika dan kode etik.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Menerapkan norma-norma secara konsisten dalam setiap situasi, pada unit kerja terkecil/kelompok kerjanya.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Memberikan informasi yang dapat dipercaya sesuai dengan etika organisasi.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Gambar 2. Kuisisioner Self-Assesment

Data *Responses* yang masuk akan terkumpul pada *spreadsheet* dan dapat di unduh berbentuk *file excel*. Data perilaku kerja diolah dalam Microsoft excel untuk menemukan data kesenjangan kompetensi ASN. Seluruh data yang masuk kemudian disandingkan dengan *database* kepegawaian Provinsi Lampung, untuk melakukan pengecekan pegawai berdasarkan NIP yang di inputkan dalam memastikan responden tersebut benar merupakan pegawai di Lingkungan Pemerintah Provinsi Lampung dengan melihat nama satuan Kerja Perangkat Daerah tempat pegawai tersebut bekerja. Data kesenjangan kompetensi yang didapatkan selanjutnya akan dijadikan *dataset* yang di uji pada penelitian ini.

3.1 Pengelolaan Data

Sesuai dengan tujuan penelitian ini, yaitu seberapa akurat metode Naïve Bayes dalam melakukan klasifikasi Tingkat kesenjangan kompetensi PNS. Untuk itu, maka dari *dataset* yang terbentuk, akan di lakukan proses *selection* untuk menentukan variabel mana saja yang berpengaruh pada proses klasifikasi dan membuang beberapa variabel yang tidak berkaitan sehingga akan menghasilkan sekumpulan data yang sudah bersih[17].

Dari *dataset* yang terbentuk, akan diberikan batasan pengujian, yaitu pada kompetensi manajerial pada aspek integritas, maka pada pengujian akan terfokus pada Kesenjangan Kompetensi Manajerial pada aspek Integritas. Sehingga pada tahap ini terdapat variable yang dikeluarkan dari *dataset*, yaitu variable Nama Perangkat Daerah asal, nama jabatan, dan Riwayat Pendidikan karena dianggap tidak memberikan informasi yang berkontribusi dengan tujuan penelitian[18].

Begitu juga pada variabel aspek kompetensi yang lainnya selain aspek integritas sementara dikeluarkan karena pengujian akan terfokus pada kompetensi manajerial pada aspek integritas. Untuk kerahasiaan data pegawai, nama pegawai ditampilkan dalam bentuk kode *employee*, dari proses *selection* ini didapatkan variabel yang terkait, yaitu:

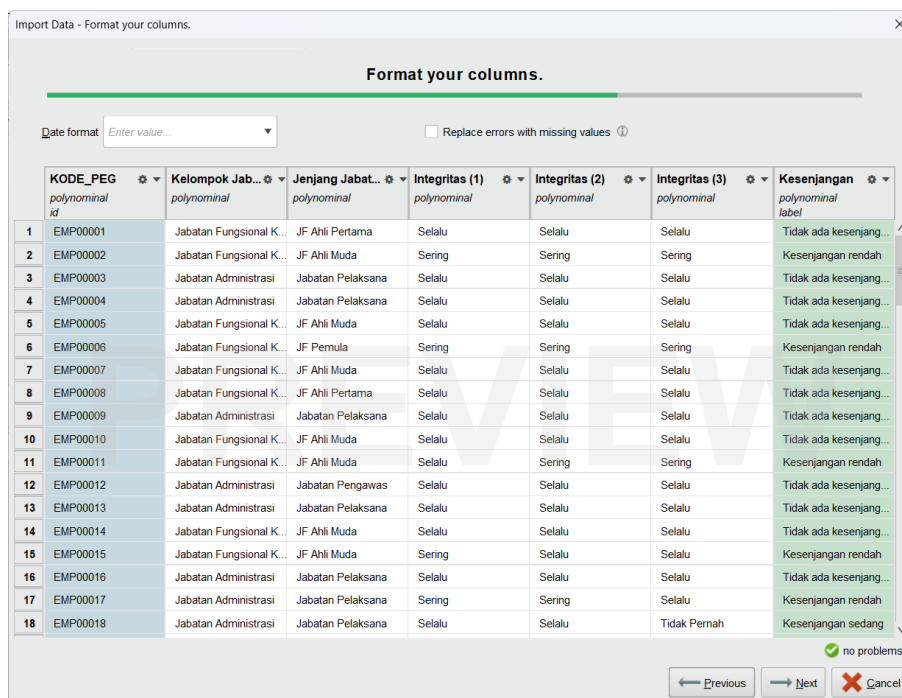
- kode *employee*,
- kelompok jabatan PNS,
- jenjang jabatan PNS,
- integritas (1): Hasil *Responses* Pertanyaan Integritas 1,
- integritas (2): Hasil *Responses* Pertanyaan Integritas 2,
- integritas (3): Hasil *Responses* Pertanyaan Integritas 3.

Label pada *dataset* ini adalah Kesenjangan, dengan atribut target kelas adalah "Tidak ada kesenjangan", "Kesenjangan rendah", "Kesenjangan sedang", dan "Kesenjangan tinggi."

3.2 Pengujian Naïve Bayes dengan menggunakan Rapidminer

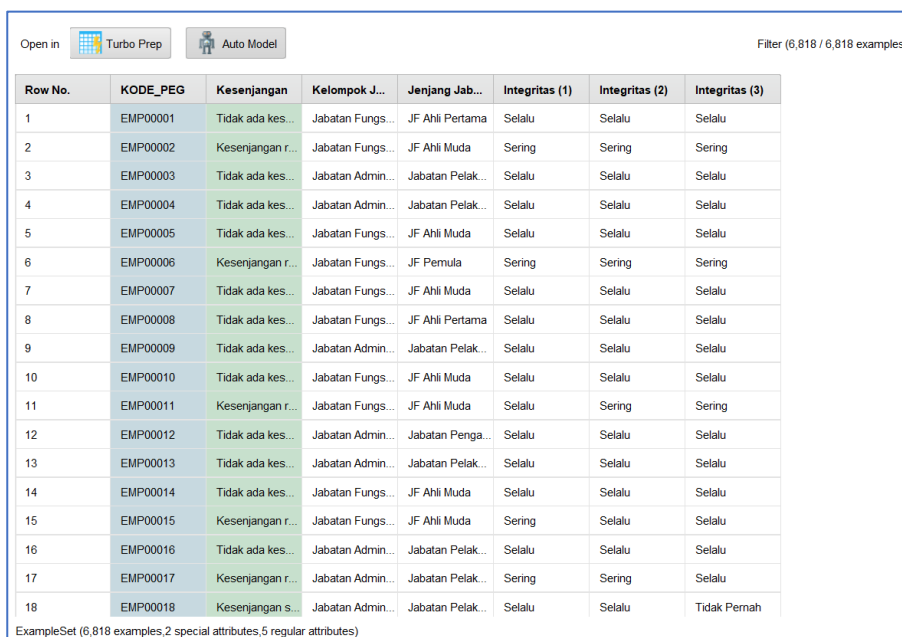
Pada tahap ini dilakukan pengujian terhadap seberapa besar akurasi yang didapatkan dalam pemodelan menggunakan metode Naïve Bayes dalam melakukan klasifikasi terhadap data kesenjangan kompetensi PNS. Pemodelan ini menggunakan *tools rapidminer studio*. Penelitian ini menggunakan 6.818 baris data pada *dataset* Kesenjangan

Kompetensi PNS, *dataset* ini disimpan dalam format *file Comma Separated Values (csv)* agar dapat dibaca oleh *tools rapidminer*[19]. Gambar 3 menunjukkan langkah dalam *import data csv* kedalam *tools rapidminer*, dalam langkah ini ditentukan atribut-atribut dalam *dataset* yang merupakan *id* dan atribut *label* (target klasifikasi)[20].



Gambar 3. Import file csv

Dari data yang telah di-*import*, dapat dilihat bahwa *dataset* terdiri dari 6.818 baris data, 2 *special attributes*, dan 5 *regular attributes*, yaitu Kelompok Jabatan, Jenjang Jabatan, serta Integritas(1), Integritas(2) dan Integritas(3) yang merupakan data perilaku kerja integritas pada setiap PNS yang didapatkan dari hasil *self-assessment*. Data yang ditampilkan dapat dilihat pada gambar 4.



Row No.	KODE_PEG	Kesenjangan	Kelompok J...	Jenjang Jab...	Integritas (1)	Integritas (2)	Integritas (3)
1	EMP00001	Tidak ada kes...	Jabatan Fungs...	JF Ahli Pertama	Selalu	Selalu	Selalu
2	EMP00002	Kesenjangan r...	Jabatan Fungs...	JF Ahli Muda	Sering	Sering	Sering
3	EMP00003	Tidak ada kes...	Jabatan Admin...	Jabatan Pelak...	Selalu	Selalu	Selalu
4	EMP00004	Tidak ada kes...	Jabatan Admin...	Jabatan Pelak...	Selalu	Selalu	Selalu
5	EMP00005	Tidak ada kes...	Jabatan Fungs...	JF Ahli Muda	Selalu	Selalu	Selalu
6	EMP00006	Kesenjangan r...	Jabatan Fungs...	JF Pemula	Sering	Sering	Sering
7	EMP00007	Tidak ada kes...	Jabatan Fungs...	JF Ahli Muda	Selalu	Selalu	Selalu
8	EMP00008	Tidak ada kes...	Jabatan Fungs...	JF Ahli Pertama	Selalu	Selalu	Selalu
9	EMP00009	Tidak ada kes...	Jabatan Admin...	Jabatan Pelak...	Selalu	Selalu	Selalu
10	EMP00010	Tidak ada kes...	Jabatan Fungs...	JF Ahli Muda	Selalu	Selalu	Selalu
11	EMP00011	Kesenjangan r...	Jabatan Fungs...	JF Ahli Muda	Selalu	Sering	Sering
12	EMP00012	Tidak ada kes...	Jabatan Admin...	Jabatan Penga...	Selalu	Selalu	Selalu
13	EMP00013	Tidak ada kes...	Jabatan Admin...	Jabatan Pelak...	Selalu	Selalu	Selalu
14	EMP00014	Tidak ada kes...	Jabatan Fungs...	JF Ahli Muda	Selalu	Selalu	Selalu
15	EMP00015	Kesenjangan r...	Jabatan Fungs...	JF Ahli Muda	Sering	Selalu	Selalu
16	EMP00016	Tidak ada kes...	Jabatan Admin...	Jabatan Pelak...	Selalu	Selalu	Selalu
17	EMP00017	Kesenjangan r...	Jabatan Admin...	Jabatan Pelak...	Sering	Sering	Selalu
18	EMP00018	Kesenjangan s...	Jabatan Admin...	Jabatan Pelak...	Selalu	Selalu	Tidak Pernah

Gambar 4. Attributes dataset

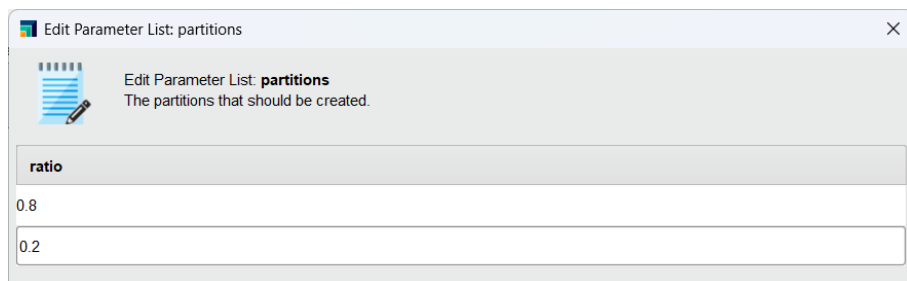
Dari *dataset* yang tersedia, atribut label menunjukkan data paling sedikit yaitu pada kelas kesenjangan tinggi, yaitu sebanyak 30. Sedangkan data paling banyak yaitu pada kelas tidak ada kesenjangan yaitu sebanyak 4.258 data. Data statistik kesenjangan dapat dilihat pada gambar 5.



Gambar 5. Statistik data kesenjangan

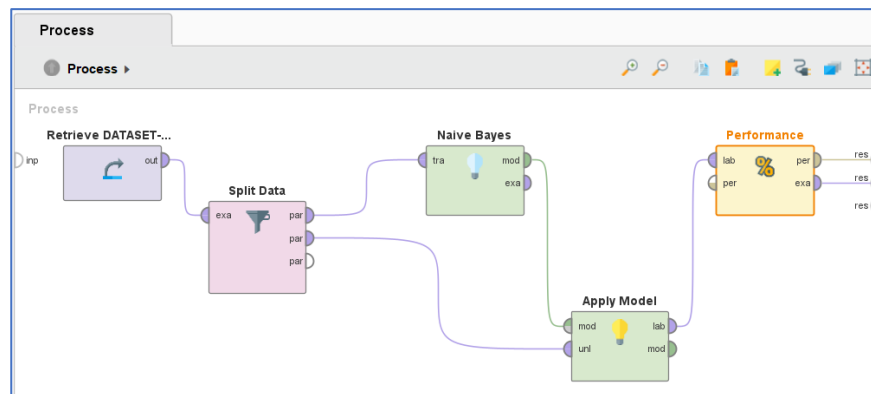
Dalam pemodelan algoritma Naïve Bayes terkait pengklasifikasian kesenjangan kompetensi PNS, terlebih dahulu dilakukan *split data* dikarenakan *dataset* tergabung menjadi satu tidak terpisah antara *data training* dan *data testing*. Dalam proses *split* ini ditentukan *data training* 80% dan *data testing* 20%.

Gambar 6 menunjukkan proses *split data*, *ratio* data dibuat dari angka 0.0 – 1.0, dalam penelitian ini *data training* dibuat 0.8 yang berarti 80% data untuk proses *training model*. *Data training* dibuat lebih besar daripada *data testing* yang hanya 0.2 (20% data), dikarenakan jika semakin banyak referensi belajar yang diterima oleh model, nantinya akan menghasilkan *knowledge* yang banyak pula pada model, sehingga berpengaruh pada performa yang lebih baik.



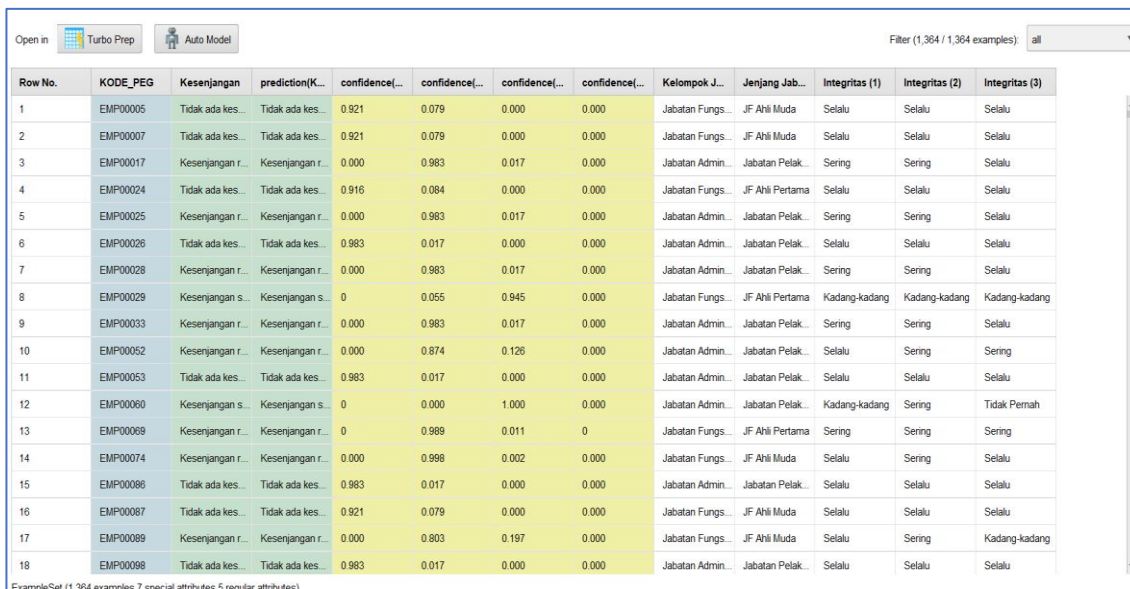
Gambar 6. Split data training dan testing

Setelah *dataset* dilakukan *split*, dilakukan training pada *tools rapidminer*, *data training* dihubungkan pada operator Naïve Bayes untuk dilakukan pembelajaran dan hasil pemodelan yang terbentuk dihubungkan pada operator *apply model*. Sedangkan *data testing* langsung dihubungkan pada operator *apply model*. Setelah model berhasil dijalankan, selanjutnya dihubungkan *apply model* kepada operator *performance* untuk dapat mengukur performa dari model klasifikasi dengan algoritma Naïve Bayes yang telah dijalankan. Proses dapat dilihat pada gambar 7.



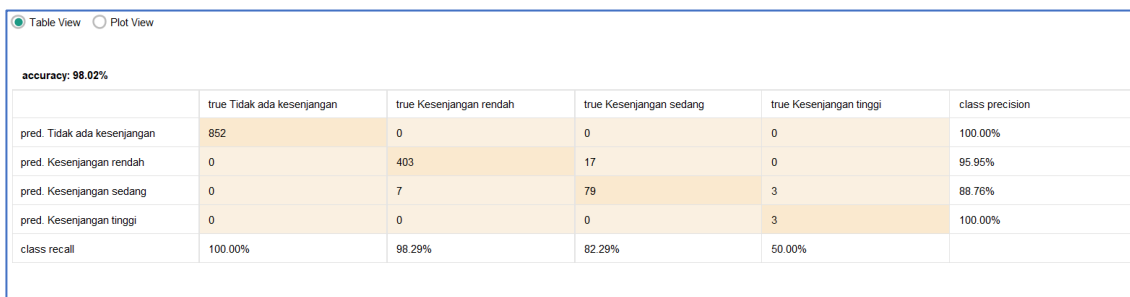
Gambar 7. Proses Pemodelan Naïve Bayes

Dari hasil pemodelan, dapat dilihat hasil performa algoritma Naïve Bayes dalam melakukan klasifikasi *multi-class* pada *dataset* kompetensi manajerial aspek integritas PNS pada Pemerintah Provinsi Lampung. Data prediksi yang diberikan dari algoritma Naïve Bayes tertera pada gambar 8. Proses klasifikasi dengan metode Naïve Bayes menggunakan *tools rapidminer* menghasilkan performa model dengan menampilkan nilai *accuracy*, *class recall*, dan *class precision* yang dapat dilihat pada gambar 9 [21][22].



Row No.	KODE_PEG	Kesenjangan	prediction(K...	confidence(...	confidence(...	confidence(...	confidence(...	Kelompok J...	Jenjang Jab...	Integritas (1)	Integritas (2)	Integritas (3)
1	EMP00005	Tidak ada kes...	Tidak ada kes...	0.921	0.079	0.000	0.000	Jabatan Fungs...	JF Ahli Muda	Selalu	Selalu	Selalu
2	EMP00007	Tidak ada kes...	Tidak ada kes...	0.921	0.079	0.000	0.000	Jabatan Fungs...	JF Ahli Muda	Selalu	Selalu	Selalu
3	EMP00017	Kesenjangan r...	Kesenjangan r...	0.000	0.983	0.017	0.000	Jabatan Admin...	Jabatan Pelak...	Sering	Sering	Selalu
4	EMP00024	Tidak ada kes...	Tidak ada kes...	0.916	0.084	0.000	0.000	Jabatan Fungs...	JF Ahli Pertama	Selalu	Selalu	Selalu
5	EMP00025	Kesenjangan r...	Kesenjangan r...	0.000	0.983	0.017	0.000	Jabatan Admin...	Jabatan Pelak...	Sering	Sering	Selalu
6	EMP00026	Tidak ada kes...	Tidak ada kes...	0.983	0.017	0.000	0.000	Jabatan Admin...	Jabatan Pelak...	Selalu	Selalu	Selalu
7	EMP00028	Kesenjangan r...	Kesenjangan r...	0.000	0.983	0.017	0.000	Jabatan Admin...	Jabatan Pelak...	Sering	Sering	Selalu
8	EMP00029	Kesenjangan s...	Kesenjangan s...	0	0.055	0.945	0.000	Jabatan Fungs...	JF Ahli Pertama	Kadang-kadang	Kadang-kadang	Kadang-kadang
9	EMP00033	Kesenjangan r...	Kesenjangan r...	0.000	0.983	0.017	0.000	Jabatan Admin...	Jabatan Pelak...	Sering	Sering	Selalu
10	EMP00052	Kesenjangan r...	Kesenjangan r...	0.000	0.874	0.126	0.000	Jabatan Admin...	Jabatan Pelak...	Selalu	Sering	Sering
11	EMP00053	Tidak ada kes...	Tidak ada kes...	0.983	0.017	0.000	0.000	Jabatan Admin...	Jabatan Pelak...	Selalu	Selalu	Selalu
12	EMP00060	Kesenjangan s...	Kesenjangan s...	0	0.000	1.000	0.000	Jabatan Admin...	Jabatan Pelak...	Kadang-kadang	Sering	Tidak Pernah
13	EMP00069	Kesenjangan r...	Kesenjangan r...	0	0.989	0.011	0	Jabatan Fungs...	JF Ahli Pertama	Sering	Sering	Sering
14	EMP00074	Kesenjangan r...	Kesenjangan r...	0.000	0.998	0.002	0.000	Jabatan Fungs...	JF Ahli Muda	Selalu	Sering	Selalu
15	EMP00086	Tidak ada kes...	Tidak ada kes...	0.983	0.017	0.000	0.000	Jabatan Admin...	Jabatan Pelak...	Selalu	Selalu	Selalu
16	EMP00087	Tidak ada kes...	Tidak ada kes...	0.921	0.079	0.000	0.000	Jabatan Fungs...	JF Ahli Muda	Selalu	Selalu	Selalu
17	EMP00089	Kesenjangan r...	Kesenjangan r...	0.000	0.803	0.197	0.000	Jabatan Fungs...	JF Ahli Muda	Selalu	Sering	Kadang-kadang
18	EMP00098	Tidak ada kes...	Tidak ada kes...	0.983	0.017	0.000	0.000	Jabatan Admin...	Jabatan Pelak...	Selalu	Selalu	Selalu

Gambar 8. Hasil prediksi yang didapatkan dengan algoritma naïve bayes



accuracy: 98.02%

	true Tidak ada kesenjangan	true Kesenjangan rendah	true Kesenjangan sedang	true Kesenjangan tinggi	class precision
pred. Tidak ada kesenjangan	852	0	0	0	100.00%
pred. Kesenjangan rendah	0	403	17	0	95.95%
pred. Kesenjangan sedang	0	7	79	3	88.76%
pred. Kesenjangan tinggi	0	0	0	3	100.00%
class recall	100.00%	98.29%	82.29%	50.00%	

Gambar 9. Performa algoritma naïve bayes

Hasil pemodelan dengan mengklasifikasikan data testing berjumlah 1.364 data menunjukkan hasil *accuracy* yang didapat adalah sebesar 98,02%, dengan masing-masing *class recall*, yaitu:

- a. Tidak ada kesenjangan (TK) : 100%
- b. Kesenjangan rendah (KR) : 98,29%
- c. Kesenjangan sedang (KS) : 82,29%
- d. Kesenjangan tinggi (KT) : 50%

Sementara masing-masing *class precision*, yaitu:

- a. Tidak ada kesenjangan (TK) : 100,00%
- b. Kesenjangan rendah (KR) : 95,95%
- c. Kesenjangan sedang (KS) : 88,76%
- d. Kesenjangan tinggi (KT) : 100,00%

3.3 Analisis

Dari hasil pemodelan klasifikasi yang dilakukan menggunakan algoritma Naïve Bayes, data performa model yang dihasilkan tergambar dalam *confusion matrix* [23] yang dapat dilihat pada tabel 2.

Table 2. Confusion matrix

		TRUE			
		TK	KR	KS	KT
PREDICTION	TK	852	0	0	0
	KR	0	403	17	0
	KS	0	7	79	3
	KT	0	0	0	3



Dari *confusion matrix* dapat dihitung nilai akurasi. Akurasi merupakan proporsi dari total prediksi yang benar dari seluruh hasil prediksi yang didapatkan. Secara sederhana, pada klasifikasi *multi-class*, nilai akurasi bisa didapat dengan menghitung total semua class yang dikenali benar (total *True Positives*) dibagi dengan jumlah total data secara keseluruhan[24]. Maka secara manual dapat dihitung nilai akurasi yaitu :

$$\begin{aligned}
 \text{Accuracy} &= \frac{\text{Jumlah prediksi benar}}{\text{Jumlah keseluruhan data}} \times 100\% & (5) \\
 &= \frac{852+403+79+3}{1364} \times 100\% \\
 &= \frac{1337}{1364} \times 100\% \\
 &= \mathbf{98.02\%}
 \end{aligned}$$

Sedangkan untuk *precision* dan *recall* pada setiap kelas dapat dihitung dengan rumus (2) dan rumus (3), Secara manual dapat dilihat pada perhitungan *class precision* dan *class recall* pada setiap kelasnya berikut ini:

a. Kelas Tidak ada kesenjangan (TK)

$$1. \text{Precision} = \frac{852}{852+0} \times 100\% = 100\%$$

$$2. \text{ReCall} = \frac{852}{852+0} \times 100\% = 100\%$$

b. Kelas Kesenjangan rendah (KR)

$$1. \text{Precision} = \frac{403}{403+17} \times 100\% = 95,95\%$$

$$2. \text{ReCall} = \frac{403}{403+7} \times 100\% = 98,29\%$$

c. Kelas Kesenjangan sedang (KS)

$$1. \text{Precision} = \frac{79}{79+3} \times 100\% = 88,76\%$$

$$2. \text{ReCall} = \frac{79}{79+17} \times 100\% = 82,29\%$$

d. Kelas Kesenjangan tinggi (KT)

$$1. \text{Precision} = \frac{3}{3+0} \times 100\% = 100\%$$

$$2. \text{ReCall} = \frac{3}{3+3} \times 100\% = 50\%$$

Untuk menghitung harmonisasi antara *class precision* dan *class recall*, dapat dicari dengan perhitungan *F1 Score* per-kelas[23], yaitu:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

a. Kelas Tidak ada kesenjangan (TK)

$$\text{F1 Score} = 2 \times \frac{1 \times 1}{1 + 1} = 100\%$$

b. Kelas Kesenjangan rendah (KR)

$$\text{F1 Score} = 2 \times \frac{0.9595 \times 0.9829}{0.9595 + 0.9829} = 97,10\%$$

c. Kelas Kesenjangan sedang (KS)

$$\text{F1 Score} = 2 \times \frac{0.8876 \times 0.8229}{0.8876 + 0.8229} = 85,40\%$$

d. Kelas Kesenjangan tinggi (KT)

$$\text{F1 Score} = 2 \times \frac{1 \times 0.5}{1 + 0.5} = 66,67\%$$

Rekapitulasi performa dari model klasifikasi dengan menggunakan metode Naïve Bayes pada *dataset* kesenjangan kompetensi PNS Pemerintah Provinsi Lampung dapat dilihat pada Tabel 3.

Tabel 3. Rekapitulasi performa per-kelas

	TK	KR	KS	KT
Accuracy	98,02%			
Precision	100%	95,95%	88,76%	100%
Recall	100%	98,29%	82,29%	50%
F1 Score	100%	97,10%	85,40%	66,67%

Untuk menghitung *precision*, *recall*, dan *F1 Score* secara keseluruhan, digunakan dua metode umum yaitu *macro-average* dan *weighted-average*, berikut ini perhitungannya:

a. *Macro Average*

$$1. \text{ Macro Precision} = \frac{1}{N} \sum_{i=1}^N \text{precision}_i \quad (7)$$

$$= \frac{1}{4} \times (100\% + 95,95\% + 88,76\% + 100\%)$$

$$= \mathbf{96,18\%}$$

$$2. \text{ Macro Recall} = \frac{1}{N} \sum_{i=1}^N \text{recall}_i \quad (8)$$

$$= \frac{1}{4} \times (100\% + 98,29\% + 82,29\% + 50\%)$$

$$= \mathbf{82,64\%}$$

$$3. \text{ Macro F1 Score} = \frac{2 \times \text{Macro Precision} \times \text{Macro Recall}}{\text{Macro Precision} + \text{macro Recall}} \quad (9)$$

$$= \frac{2 \times 96,18\% \times 82,64\%}{96,18\% + 82,64\%}$$

$$= \mathbf{88,88\%}$$

b. *Weighted Average*

$$1. \text{ Weighted Precision} = \frac{1}{T} \sum_{i=1}^N (\text{precision}_i \times \text{jumlah aktual}_i) \quad (10)$$

$$= \frac{1}{1364} \times (100\% \times 852 + 95,95\% \times 410 + 88,76\% \times 96 + 100\% \times 6)$$

$$= \mathbf{97,99\%}$$

$$2. \text{ Weighted Recall} = \frac{1}{T} \sum_{i=1}^N (\text{recall}_i \times \text{jumlah aktual}_i) \quad (11)$$

$$= \frac{1}{1364} \times (100\% \times 852 + 98,29\% \times 410 + 82,29\% \times 96 + 50\% \times 6)$$

$$= \mathbf{98,02\%}$$

$$3. \text{ Weighted F1 Score} = \frac{2 \times \text{Macro Precision} \times \text{Macro Recall}}{\text{Macro Precision} + \text{macro Recall}} \quad (12)$$

$$= \frac{2 \times 97,99\% \times 98,02\%}{97,99\% + 98,02\%}$$

$$= \mathbf{98,00\%}$$

Hasil perhitungan metrik evaluasi untuk sebuah model klasifikasi menggunakan dua pendekatan, yaitu *Macro Average* dan *Weighted Average*. Didapatkan bahwa *Macro Average* menghitung rata-rata metrik (*Precision*, *Recall*, dan *F1 Score*) dengan memperlakukan setiap kelas secara setara, tanpa memperhatikan distribusi kelas, sehingga didapatkan *Macro Precision* sebesar 96,18%, *Macro Recall* sebesar 82,64%, dan *Macro F1 Score* sebesar 88,88%. Sementara itu, *Weighted Average* menghitung rata-rata metrik dengan memberikan bobot sesuai dengan jumlah sampel di setiap kelas, menghasilkan *Weighted Precision* sebesar 97,99%, *Weighted Recall* sebesar 98,02%, dan *Weighted F1 Score* sebesar 98,00%. Hasil ini menunjukkan bahwa model yang dibangun memiliki performa yang sangat baik, baik secara keseluruhan maupun ketika mempertimbangkan distribusi kelas.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa metode Naïve Bayes dapat diterapkan untuk mengklasifikasikan kesenjangan kompetensi PNS di Pemerintah Provinsi Lampung pada kelas kesenjangan tinggi, kesenjangan sedang, kesenjangan rendah atau tidak ada kesenjangan. Algoritma Naïve Bayes diterapkan pada data latih untuk membentuk model klasifikasi. Proses ini melibatkan perhitungan probabilitas setiap kelas kesenjangan kompetensi PNS berdasarkan fitur-fitur yang ada dalam data latih dan diimplementasikan pada data *testing*. Dari hasil pengujian model, diperoleh tingkat akurasi sebesar 98,02%, yang menunjukkan bahwa algoritma Naïve Bayes dapat berjalan dengan baik dalam proses klasifikasi tingkat kesenjangan kompetensi PNS. Implementasi dari hasil penelitian ini dapat digunakan sebagai dasar perencanaan pengembangan kompetensi ASN pada Instansi Pemerintah Provinsi Lampung yang lebih efektif dan tepat sasaran sesuai kebutuhan pengembangan kompetensi masing-masing ASN. Namun perlu diperhatikan bahwa dalam penelitian ini hanya terbatas pada PNS dengan data kompetensi manajerial dan sosial kultural. Untuk penelitian selanjutnya disarankan agar dapat memperluas cakupan pengujian seperti pengujian pada kompetensi teknis ASN yang memiliki variabel penentu tingkat kompetensi ASN lebih beragam seperti data tingkat pendidikan, pelatihan, unit kerja penempatan pegawai atau faktor individu lainnya. Selain itu, dapat dilakukan pengujian dengan memanfaatkan algoritma klasifikasi lainnya seperti *Decision Tree*, *Random Forest*, atau *Support Vector Machine* (SVM) untuk mengevaluasi kinerja relatif dari metode-metode tersebut.

REFERENCES

- [1] Komisi Aparatur Sipil Negara, *Keputusan Komisi Aparatur Sipil Negara Nomor 90/KEP.KASN/C/XI/2023 Tentang Penetapan Kategori, Penilaian dan Indeks Penerapan Sistem Merit Dalam Manajemen ASN di Lingkungan Pemerintah Provinsi Lampung*. Indonesia, 2023, pp. 1–8.
- [2] A. Fitriani and P. Halik, “Analisis Kebutuhan Pengembangan Kompetensi Manajerial Melalui Penilaian Kompetensi Pada Jabatan Fungsional Tertentu Provinsi Sulawesi Selatan,” *PESHUM : Jurnal Pendidikan, Sosial dan Humaniora*, vol. 2, no. 2, pp. 318–326, 2023.
- [3] Menteri Pendayagunaan Aparatur Negara dan Reformasi Birokrasi Republik Indonesia, *Peraturan Menteri Pendayagunaan Aparatur Negara dan Reformasi Birokrasi Republik Indonesia Nomor 38 Tahun 2017 tentang Standar Kompetensi Jabatan Aparatur Sipil Negara*. Indonesia, 2017, pp. 1–108. Accessed: May 03, 2024. [Online]. Available: <https://peraturan.bpk.go.id/Download/123418/PERMENPAN%20NOMOR%2038%20TAHUN%202017.pdf>
- [4] Y. Irfayanti, “Penerapan Metode Naive Bayes untuk Klasifikasi Status Pegawai pada Perusahaan Swasta,” *Syntax Literate: Jurnal Ilmiah Indonesia*, vol. 7, no. 10, pp. 17934–17941, 2022, doi: 10.36418/syntax-literate.v7i10.13230.
- [5] H. Annur, “KLASIFIKASI MASYARAKAT MISKIN MENGGUNAKAN METODE NAÏVE BAYES,” *ILKOM Jurnal Ilmiah*, vol. 10, no. 2, pp. 160–165, 2018.
- [6] Sadimin and H. Widi Nugroho, “PERBANDINGAN KINERJA ALGORITMA DATAMINING UNTUK PREDIKSI KELULUSAN MAHASIWA,” vol. 17, no. 2, Jul. 2023, [Online]. Available: <https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index>
- [7] M. Agarina and Sutedi, “Teknika 14 (02): 165-174 Penerapan Data Mining dalam Perancangan Sistem Pendukung Keputusan Seleksi Penerimaan Beasiswa Menggunakan Naive Bayes Classifier (Studi Kasus: IIB Darmajaya),” *Teknika*, vol. 14, no. 2, pp. 165–174, Dec. 2020.
- [8] A. Saputra Dinata and nisar, “Penerapan Algoritma Naive Bayes Dalam Pengadaan Buku Referensi Pada Perpustakaan SMA Negeri 1 Trimurjo Berbasis Web,” *Indonesian Journal of Science, Technology and Humanities*, vol. 1, no. 2, pp. 80–90, 2023.
- [9] R. Rachman, R. N. Handayani, and I. Artikel, “Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM,” *JURNAL INFORMATIKA*, vol. 8, no. 2, 2021, [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/ji>
- [10] D. Florencia, “Prediksi Jenis Kesehatan Kejiwaan Berdasarkan Usia Menggunakan Metode Naive Bayes Berbasis Website,” *Jurnal Pendidikan Tambusai*, vol. 8, no. 1, pp. 15030–15040, 2024.
- [11] L. Aman, “Permasalahan Penerapan Kamus Kompetensi Manajerial dan Sosial Kultural Permenpan Rb Nomor 38 Tahun 2017 dalam Menilai Kompetensi Pejabat Fungsional Tertentu,” *Civil Service*, vol. 16, no. 2, pp. 61–88, 2022.
- [12] I. T. Monowati and R. Setyadi, “Penerapan Algoritma Naive Bayes Dalam Memprediksi Pengusulan Penghapusan Peralatan dan Mesin Kantor,” *Journal of Information System Research (JOSH)*, vol. 4, no. 2, pp. 483–491, Jan. 2023, doi: 10.47065/josh.v4i2.2674.
- [13] M. Anggraini, R. A. Tyas, I. A. Sulasyah, and Q. Aini, “Implementasi Algoritma Naive Bayes Dalam Penentuan Rating Buku,” *SISTEMASI*, vol. 9, no. 3, pp. 557–566, Sep. 2020, doi: 10.32520/stmsi.v9i3.915.
- [14] M. Asfi and N. Fitrianiingsih, “Implementasi Algoritma Naive Bayes Classifier sebagai Sistem Rekomendasi Pembimbing Skripsi,” *InfoTekJar : Jurnal Nasional Informatika dan Teknologi Jaringan*, vol. 5, no. 1, pp. 44–50, 2020, doi: 10.30743/infotekjar.v5i1.2536.
- [15] D. R. Andriyani, M. Afdal, and S. Monalisa, “Analisis Sentimen Masyarakat Terhadap Penghapusan Honorer Berdasarkan Opini Dari Twitter Menggunakan Naive Bayes Classifier,” *Building of Informatics, Technology and Science (BITS)*, vol. 5, no. 1, pp. 49–58, Jun. 2023, doi: 10.47065/bits.v5i1.3541.
- [16] J. P. Tanjung, F. C. Tampubolon, A. W. Panggabean, and M. A. A. Nandrawan, “Customer Classification Using Naive Bayes Classifier With Genetic Algorithm Feature Selection,” *Sinkron*, vol. 8, no. 1, pp. 584–589, Feb. 2023, doi: 10.33395/sinkron.v8i1.12182.
- [17] H. Derjad Wijaya and S. Dwiasnati, “Implementasi Data Mining dengan Algoritma Naive Bayes pada Penjualan Obat,” *JURNAL INFORMATIKA*, vol. 7, no. 1, pp. 1–7, 2020, [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/ji>
- [18] P. Rahmawati, A. Larasati, and Marsono, “Pengembangan Model Persetujuan Kredit Nasabah Bank Dengan Algoritma Klasifikasi Naive Bayes, Decision Tree, Dan Artificial Neural Network,” *J@ti Undip: Jurnal Teknik Industri*, vol. 17, no. 1, pp. 1–12, 2022.
- [19] D. Ismiyana Putri and M. Yudhi Putra, “Komparasi Algoritma Dalam Memprediksi Perubahan Harga Saham GOTO Menggunakan Rapidminer,” *JURNAL KHATULISTIWA INFORMATIKA*, vol. 11, no. 1, pp. 14–20, 2023.
- [20] F. Ariani, Amir, N. Alam, and K. Rizal, “Klasifikasi Penetapan Status Karyawan Dengan Menggunakan Metode Naive Bayes,” *Paradigma*, vol. XX, no. 2, pp. 33–38, 2018, doi: 10.31294/p.v%vi%i.4021.
- [21] R. A. Anggraini, G. Widagdo, A. Setya Budi, and M. Qomaruddin, “Penerapan Data Mining Classification untuk Data Blogger Menggunakan Metode Naive Bayes,” *Jurnal sistem dan teknologi informasi*, vol. 7, no. 1, pp. 47–51, 2019.
- [22] S. Tri Utami, Sriyanto, S. Lestari, H. Widi Nugroho, and Zarnelly, “PREDICTION OF ANEMIA USING THE PARTICLE SWARM OPTIMIZATION (PSO) AND NAÏVE BAYES ALGORITHM,” *Jurnal CoreIT*, vol. X, No.X, pp. 1–8, Jun. 2024, doi: 10.24014/coreit.v10i1.28428.
- [23] Nosiell, S. Andriyanto, and M. Said Hasibuan, “Application of Nave Bayes Algorithm for SMS Spam Classification Using Orange,” *International Journal of Advanced Science and Computer Applications*, vol. 1, no. 1, pp. 16–24, 2022, doi: 10.47679/ijasca.v1i1.3.
- [24] G. Fun, “Performansi Multiclass Classification,” <https://golchafun.medium.com/performansi-multiclass-classification-83dd4cba8d2>. Accessed: Jul. 03, 2024. [Online]. Available: <https://golchafun.medium.com/performansi-multiclass-classification-83dd4cba8d2>