

Optimasi Klasifikasi Naive Bayes Menggunakan Analisis Kemiripan Data untuk Mengukur Kinerja Dosen

Dandy Pramana Hostiadi¹, Ni Luh Putri Srinadi², I Made Darma Susila³

Institut Teknologi dan Bisnis STIKOM Bali

e-mail: ¹ dandy@stikom-bali.ac.id, ²putri@stikom-bali.ac.id, ³darma_s@stikom-bali.ac.id

Diajukan: 6 Desember 2022; Direvisi: 9 Januari 2023; Diterima: 11 Januari 2023

Abstrak

Dosen memiliki kewajiban Tridharma yaitu wajib melaksanakan pendidikan, penelitian dan pengabdian kepada masyarakat. Terdapat beberapa variabel yang dapat digunakan dalam mengukur kualitas kinerja dosen, seperti motivasi, kepemimpinan, budaya organisasi, dan pelatihan. Beberapa penelitian sebelumnya memodelkan pengukuran dengan pendekatan klasifikasi. Namun model klasifikasi perlu dioptimalkan untuk mendapatkan hasil yang optimal, misalnya dalam dengan pengukuran kemiripan data kinerja dosen sehingga mampu mendapatkan model yang optimal. Metode pengukuran kemiripan data dapat digunakan agar dapat mengelompokkan berdasarkan kelompok pola data dan mengurangi data yang bersifat redundan. Semakin banyak data yang memiliki kemiripan maka model akan mengklasifikasikan lebih cepat dan tepat. Penelitian ini melakukan optimasi klasifikasi menggunakan pengukuran kemiripan yaitu euclidean distance untuk mengelompokkan data sehingga didapatkan kelompok data yang serupa dan digunakan dalam pengukuran klasifikasi. Metode klasifikasi yang digunakan adalah naive bayes, dan optimasi menunjukkan model memiliki performa yang lebih baik dibandingkan tanpa menggunakan metode analisis kemiripan data dengan akurasi sebesar 0.757, precision 0.763, recall 0.757 dan F1-score adalah 0.755. Pendekatan yang diusulkan dapat digunakan di perguruan tinggi untuk mengukur kinerja dosen yang menggunakan metode klasifikasi dengan hasil yang tepat dan cepat.

Kata kunci: Kemiripan, Naive-Bayes, Klasifikasi, Kinerja dosen.

Abstract

Lecturers have Tridharma's responsibilities: teaching, research, and community service. The institution can use several variables to measure the quality of lecturer performance, such as motivation, leadership, organizational culture, and training. Previous studies introduced models to measure the lecturer's work performance using a classification approach. However, the classification model needs to be optimized to get optimal results. To obtain optimal performance, the classification model can use the similarity analysis technique on lecturer performance data to group data patterns based on groups and reduce redundant data. The more similar data there is, the more quickly and precisely the model will classify it. This research performs classification optimization using similarity measurements, namely Euclidean distance, to group data so that similar data groups are obtained and used in classification measurements. The classification method used is naive Bayes, and optimization shows that the model performs better than without using the data similarity analysis method with an accuracy of 0.757, precision of 0.763, recall of 0.757 and F1-score of 0.755. This model can be used in higher education to measure lecturer performance using classification methods with precise and fast results.

Keywords: similarity, Naive Bayes, classification, lecturer work performance.

1. Pendahuluan

Undang Undang no 12 tahun 2012 menyatakan bahwa perguruan tinggi adalah sebagai penyelenggara pendidikan tinggi dengan fungsi untuk pengembangan pendidikan atau civitas pendidikan yang inovatif, kreatif, terampil, berdaya saing dan kooperatif [1]. Dosen dalam hal tersebut memiliki peran vital yaitu melaksanakan Pendidikan, Penelitian dan Pengabdian [2]. Sehingga dibutuhkan suatu pengukuran yang objektif dalam menentukan kinerja dosen yang tepat.

Kinerja dosen dapat diukur dari beberapa parameter diantaranya adalah pelatihan[3], budaya organisasi[4], lingkungan kerja[4], kepuasan[5], perilaku kepemimpinan [4, 6–8], motivasi[9] dan komitmen organisasi[7]. Namun, tidak semua parameter dapat ditemukan dalam sebuah institusi

pendidikan dan tidak serta merta dapat dimodelkan dalam model pengukuran kinerja dosen. Misalnya, dalam pengukuran model hanya terdapat 1 hingga 2 variabel yang relevan untuk digunakan dalam pemodelan pengukuran kualitas kinerja dosen, misalnya penggunaan variabel kepemimpinan dan motivasi kerja [10–12].

Pada penelitian [13, 14], menyebutkan bahwa budaya organisasi diperlukan dalam sebuah perusahaan. Budaya organisasi dikatakan dapat mengukur kualitas tenaga pengajar dan menciptakan suasana lingkungan kerja yang optimal dan menunjang kualitas pendidikan. Teknik pengukuran budaya organisasi dapat dilihat berdasarkan bagaimana perilaku dan tindakan dosen dalam lingkungannya. Sedangkan menurut [15], keberadaan budaya di dalam organisasi atau disebut dengan budaya organisasi tidak bisa dilihat oleh mata, tapi bisa dirasakan. Budaya organisasi itu bisa dirasakan keberadaannya melalui perilaku anggota karyawan di dalam organisasi itu sendiri. Kebudayaan tersebut memberikan pola, cara-cara berpikir, merasa menanggapi dan menuntun para anggota dalam organisasi. Oleh karena itu, budaya organisasi akan berpengaruh juga terhadap efektif atau tidaknya suatu organisasi. Budaya organisasi biasanya juga dijelaskan sebagai keyakinan inti yang dimiliki secara bersama oleh anggota organisasi. Keyakinan ini mencerminkan pemahaman tentang nilai-nilai kebersamaan yang diekspresikan dalam perilaku anggota organisasi.

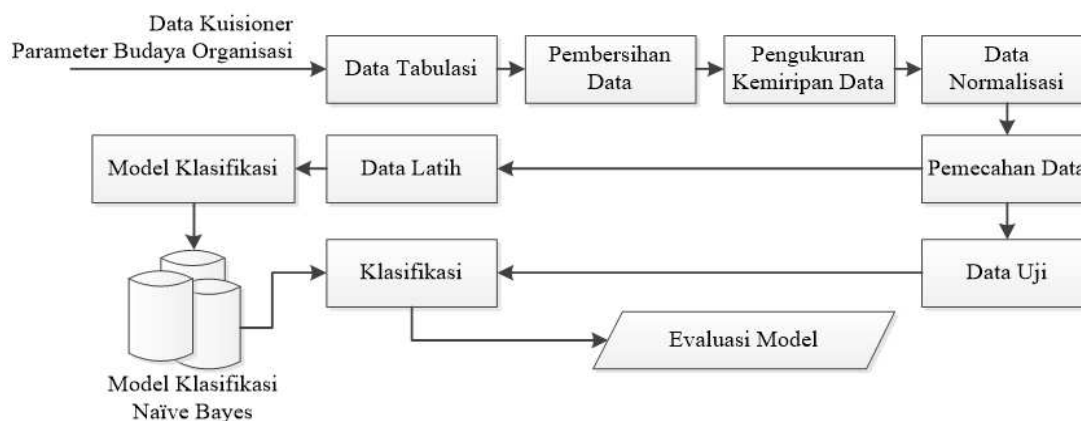
Beberapa penelitian sebelumnya, mengenalkan proses analisis kinerja dosen yang melibatkan penggunaan model analisis data secara statistik. Penelitian terdahulu [4, 10, 11] menganalisis secara statistik mencari kausalitas dengan menggunakan SEM (*Structural Equation Modelling*) dan secara sekuensial. Penelitian ini mampu menganalisis hubungan setiap parameter pengukuran dengan hasil nilai yang optimal. Namun, pengukuran yang dilakukan hanya menunjukkan pengaruh satu arah berdasarkan parameter yang diukur seperti melibatkan dua parameter yaitu kepemimpinan dan budaya organisasi.

Seiring dengan perkembangan teknologi pengolahan data, maka untuk mengukur kinerja dosen dapat menggunakan pendekatan *artificial intelligent* [16]. Beberapa metode klasifikasi yang umum digunakan adalah Decision Tree, Naive Bayes, k-NN, SVM dan Random Forest [13, 16–20]. Model ini mampu bekerja dengan baik dalam melakukan klasifikasi, namun memerlukan optimasi dalam penanganan data yang besar dan beberapa data yang bersifat ganda. Penanganan data awal dalam tahap pra pemrosesan sering menggunakan pendekatan data normalisasi [20, 21] dan pemetaan kemiripan data. Teknik ini mampu menunjukkan pengaruh terhadap peningkatan akurasi model prediksi.

Penelitian ini mengenalkan teknik optimasi metode klasifikasi dengan penggunaan pendekatan analisis kemiripan yaitu metode euclidean distance untuk mengukur kemiripan data kinerja dosen. Tujuannya adalah untuk menekan kemungkinan data yang bersifat redundan sehingga dapat meningkatkan efisiensi klasifikasi dan akurasi metode klasifikasi Naive Bayes. Keterbaruan dari penelitian ini adalah untuk mendapatkan model kerja dari metode euclidean terhadap model klasifikasi Naive Bayes. Manfaat dari penelitian ini adalah agar dapat membantu analisis dari setiap perguruan tinggi dalam menganalisis kinerja dosen yang tepat, efisien dan akurat dengan menggunakan metode klasifikasi Naive Bayes.

2. Metode Penelitian

Dalam penelitian ini, dataset yang digunakan adalah pengukuran kinerja dosen yang menggunakan satu variabel yaitu budaya organisasi yang memiliki *sub-variable* yaitu Keyakinan, Perilaku, Sistem Kontrol, Norma-norma, Kepercayaan, Nilai Kebersamaan, Pembagian Kerja. Alur Model yang diusulkan dalam penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Alur Model

2.1. Data Tabulasi

Dalam penelitian ini pengambilan data menggunakan kuesioner, dimana mengambil 53 data penilaian kinerja dosen dari variabel budaya organisasi. Jika budaya organisasi dinotasikan sebagai $\mathcal{G}$ dan memiliki 7 sub-variabel $\varphi_{i,7}$ adalah Keyakinan φ_1 , Perilaku φ_2 , Sistem Kontrol φ_3 , Norma-norma φ_4 , Kepercayaan φ_5 , Nilai Kebersamaan φ_6 , Pembagian Kerja φ_7 , maka dalam tabulasi data, nilai budaya data dituliskan menjadi $\varphi \in \mathcal{G}$, $\mathcal{G} = \{\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6, \varphi_7\}$.

2.2. Pembersihan Data

Di tahap ini dilakukan pembersihan data yaitu penghapusan nilai *record* data yang tidak memiliki penilaian sub-variabel penilaian. Dalam penelitian ini, jika data data *record* dinotasikan sebagai r dan dataset merupakan kumpulan data (R), maka berlaku persamaan (1).

$$Data\ r_i \begin{cases} hapus, jika\ nilai\ \varphi = null \\ data\ r_i = bagian\ data\ R, jika\ nilai\ \varphi \neq null \end{cases} \quad (1)$$

2.3. Pengukuran Kemiripan

Pengukuran kemiripan dilakukan dengan menggunakan pengukuran euclidean distance yang ditunjukkan pada persamaan (2) yaitu:

$$d = \sqrt{[(x_2 - x_1)^2 + (y_2 - y_1)^2]}, \quad (2)$$

dimana d adalah hasil pengukuran jarak, x adalah kordinat selisih data x pertama dan kedua dan y adalah koordinat selisih data y pertama dan kedua. Sehingga dalam penelitian ini, jika x_2 adalah data $\varphi_{i,r2}$, y_2 adalah data $\varphi_{j,r2}$, x_1 adalah data $\varphi_{i,r1}$, y_1 adalah data $\varphi_{j,r1}$ maka perhitungan euclidean distance untuk mengukur kemiripan data menjadi persamaan (3)

$$d = \sqrt{\left[(\varphi_{1,r2} - \varphi_{1,r1})^2 + (\varphi_{2,r2} - \varphi_{2,r1})^2 + (\varphi_{\dots,r2} - \varphi_{\dots,r1})^2 + (\varphi_{7,r2} - \varphi_{7,r1})^2\right]}, \quad (3)$$

Hasil dari kemiripan menghasilkan kumpulan data r yang memiliki kemiripan. Jika terdapat data yang memiliki kemiripan tinggi, maka dinyatakan data tersebut digabungkan ke dalam data yang sama.

2.4. Data Normalisasi

Pada tahap ini dilakukan proses normalisasi data, yaitu menormalkan data menjadi skala nilai hingga 1. Tujuan dari proses ini adalah menstandarisasi skala nilai dari setiap fitur, ditunjukkan pada persamaan (4).

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}, \quad (4)$$

dimana x' adalah nilai data yang telah dinormalisasi, x adalah nilai asli dari data, x_{min} adalah nilai minimum dari dataset, dan x_{max} adalah nilai maksimum dari dataset.

2.5. Pemecahan Data

Di Tahap ini dilakukan pemecahan data ke dalam dua data, yaitu data latih dan data uji. Data latih digunakan dalam training data dan data uji digunakan untuk proses klasifikasi. Data latih adalah data yang digunakan untuk proses pelatihan dari model klasifikasi. Data latih ini terbentuk pada proses pemecahan data. Data uji adalah data yang akan diuji dalam model klasifikasi. Dalam penelitian ini pemecahan data menggunakan komposisi 70% sebagai data latih dan 30% sebagai data uji.

2.6. Model Klasifikasi

Pada tahap ini dilakukan pelatihan dengan model klasifikasi naive bayes. Model akan melakukan pelatihan terhadap data, dan memprediksi data uji dalam tahap klasifikasi. Persamaan naive bayes ditunjukkan pada persamaan (5). Parameter yang diestimasi adalah mean dan variansi untuk setiap fitur dalam setiap kelas.

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}, \quad (5)$$

2.7. Proses Klasifikasi

Pada tahap ini dilakukan klasifikasi terhadap data uji. Terdapat 2 kelas prediksi yaitu performa bagus dan performa buruk terhadap kinerja dosen.

2.8. Proses Evaluasi Model

Dalam proses evaluasi metode klasifikasi, dilakukan pengukuran nilai akurasi, precision dan recall. Pengukuran dilakukan dengan menganalisis nilai yang terdapat pada confusion matrix. Dalam confusion matrix terdapat True Positive (TP): Jumlah kasus di mana model memprediksi positif dan sebenarnya positif. True Negative (TN): Jumlah kasus di mana model memprediksi negatif dan sebenarnya negatif. False Positive (FP): Jumlah kasus di mana model memprediksi positif tetapi sebenarnya negatif. Ini juga dikenal sebagai tipe I error. False Negative (FN): Jumlah kasus di mana model memprediksi negatif tetapi sebenarnya positif. Ini juga dikenal sebagai tipe II error. Persamaan akurasi ditunjukkan pada persamaan (6), precision pada persamaan (7), recall pada persamaan (8), dan F1-score pada persamaan (9).

$$accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (6)$$

$$precision = \frac{TP}{TP+FP} \quad (7)$$

$$recall = \frac{TP}{TP+FN} \quad (8)$$

$$F1 - score = 2 \times \frac{precision \times recall}{precision + recall} \quad (9)$$

3. Hasil dan Pembahasan

Di penelitian ini, model diolah dengan spesifikasi komputer dengan processor core i5, RAM 8 GB dan storage SSD 128GB. Bahasa pemrograman yang digunakan adalah Python versi 3. Di tahap awal, data yang diolah dalam bentuk data tabular ditunjukkan pada Tabel 1.

Tabel 1. Contoh Data Tabular dari Kuesioner.

Data r_i	Keyakinan	Perilaku	Sistem Kontrol	Norma-norma	Kepercayaan	Nilai Kebersamaan	Pembagian Kerja
r1	90	75	82	92	86	70	88
r2	77	93	84	69	87	67	71
r3	65	92	65	85	93	86	67
r4	93	94	89	65	87	76	70
r5	72	93	75	68	82	71	81
r6	84	91	73	70	68	72	84
r7	77	73	70	70	72	80	74
r8	72	81	82	84	82	89	73
r9	80	92	68	83	70	71	85
r10	90	72	89	80	92	70	74
r11	71	72	82	84	67	75	83
r12	83	86	81	91	78	89	71
r13	85	77	91	77	93	82	89
r14	89	66	66	91	81	94	70
r15	93	65	80	80	65	91	72
r16	84	74	74	79	87	86	76
r17	79	69	68	88	92	69	66
r18	72	83	86	94	93	83	72
r19	68	81	94	87	71	72	67
...	...	...	...	...	...	...	...

Pada tahap pembersihan data, tidak ditemukan data yang bersifat kosong. Sehingga data yang digunakan tetap berjumlah 53 data. Kemudian dilanjutkan dengan proses pengukuran data dengan metode

euclidean distance. Dalam pengukuran Euclidean Distance, penggunaan data yang diukur adalah terhadap 53 data dengan 7 sub-parameter, dan menghasilkan hasil pengukuran matirk dalam 2 pasangan data sejumlah 1378 data Matrix. Hasil pengukuran ditunjukkan pada Gambar 2, dan hasil dari pengukuran data ditunjukkan pada Tabel 2.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
1		36.401	44.699	38.846	36.592	34.655	35.958	32.031	30.529	20.833	29.069	29.766	23.000	35.553	37.350	25.904	29.833	29.698	36.865	21.166	22.293	13.115	33.211	28.983	24.115
2	36.401		34.278	19.494	15.748	27.019	31.623	30.183	31.032	28.267	36.249	33.808	31.906	49.467	47.308	31.559	35.369	32.357	30.757	24.920	38.210	36.647	18.921	43.623	37.115
3	44.699	34.278		43.692	31.321	41.964	35.791	25.020	36.277	43.955	42.178	30.033	43.186	38.807	50.961	30.315	32.156	26.192	40.682	36.332	41.097	47.159	35.143	49.477	40.115
4	38.846	19.494	43.692		28.601	30.725	36.606	34.971	38.432	28.196	44.136	33.985	30.561	48.631	43.128	32.218	43.370	38.743	39.799	27.258	39.522	39.179	19.287	39.205	47.115
5	36.592	15.748	31.321	28.601		19.026	26.153	28.931	22.338	35.426	31.496	34.278	32.680	47.739	46.519	30.033	37.457	35.256	34.641	27.037	36.332	36.263	26.344	40.534	33.115
6	34.655	27.019	41.964	30.725	19.026		23.707	33.571	14.731	37.855	28.601	33.000	36.332	44.598	37.683	31.496	42.036	42.930	37.202	26.363	29.462	34.453	35.412	31.064	38.115
7	35.958	31.623	35.791	36.606	26.153	23.707		24.718	27.368	33.630	22.539	29.883	35.214	30.708	26.344	20.298	30.545	37.563	33.526	35.930	24.331	34.511	36.442	26.173	37.115
8	32.031	30.183	25.020	34.971	28.931	33.571	24.718		31.528	30.529	24.556	14.697	27.221	29.223	31.796	18.028	30.887	16.673	24.799	29.189	31.890	37.947	24.393	30.364	27.115
9	30.529	31.032	36.277	38.432	22.338	14.731	27.368	31.528		39.446	26.589	29.360	38.484	41.231	40.559	31.417	37.470	37.842	35.819	22.361	26.476	32.650	38.301	34.117	32.115
10	20.833	28.267	43.955	28.196	35.426	37.855	33.630	30.529	39.446		34.015	31.559	20.809	37.417	36.235	23.473	26.458	28.705	33.660	26.721	27.911	21.260	24.104	32.680	34.115
11	29.069	36.249	42.178	44.136	31.496	28.601	22.539	24.556	26.589	34.015		29.172	33.347	37.296	30.561	28.844	35.143	33.151	22.891	33.749	29.155	33.808	38.000	27.767	22.115
12	29.766	33.808	30.033	33.985	34.278	33.000	29.883	14.697	29.360	31.559	29.172		31.289	26.382	28.931	21.284	33.226	20.640	28.089	23.452	27.911	37.363	28.478	27.964	33.115
13	23.000	31.906	43.186	30.561	32.680	36.332	35.214	27.221	38.484	20.809	33.347	31.289		40.087	38.626	22.891	38.066	28.443	38.497	28.443	34.785	26.665	20.199	29.138	27.115
14	35.553	49.467	38.807	48.631	47.739	44.598	30.708	29.223	41.231	37.417	37.296	26.382	40.087		24.556	20.809	29.732	35.440	45.376	40.694	23.302	37.417	45.442	27.423	45.115
15	37.350	47.308	50.961	43.128	46.519	37.683	26.344	31.796	40.559	36.235	30.561	28.931	38.626	24.556		26.907	40.853	42.953	39.345	41.797	26.683	39.812	44.385	15.000	46.115
16	25.904	31.559	30.315	32.218	30.033	31.496	20.298	18.028	31.417	23.473	28.844	21.284	22.891	20.809	26.907		24.104	25.593	36.083	28.373	20.688	26.552	27.276	23.043	33.115
17	29.833	35.369	32.156	43.370	37.457	42.036	30.545	30.887	37.470	26.458	35.143	33.226	38.066	29.732	40.853	24.104		28.948	37.323	34.029	24.104	27.821	38.328	41.183	37.115
18	29.698	32.357	26.192	38.743	35.256	42.930	37.563	16.673	37.842	28.705	33.151	20.640	28.443	35.440	42.953	25.593	28.948		27.622	27.568	36.729	37.443	24.104	40.596	26.115
19	36.865	30.757	40.682	39.799	34.641	37.202	33.526	24.799	35.819	33.660	22.891	28.089	38.497	45.376	39.345	36.083	37.323	27.622		33.451	39.699	43.692	32.435	40.878	30.115
20	21.166	24.920	36.332	27.258	27.037	26.363	35.930	29.189	22.361	26.721	33.749	23.452	28.443	40.694	41.797	28.373	34.029	27.568	33.451		27.695	26.870	25.357	35.043	31.115
21	22.293	38.210	41.097	39.522	36.332	29.462	24.331	31.890	26.476	27.911	29.155	27.911	34.785	23.302	26.683	20.688	24.104	36.729	39.699	27.695		20.372	40.522	23.685	37.115
22	13.115	36.647	47.159	39.179	36.263	34.453	34.511	37.947	32.650	21.260	33.808	37.363	26.665	37.417	39.812	26.552	27.821	37.443	43.692	26.870	20.372		37.054	31.591	32.115
23	28.983	18.921	35.143	19.287	26.344	35.412	36.442	24.393	38.301	24.104	38.000	28.478	20.199	45.442	44.385	27.276	38.328	24.104	32.435	25.357	40.522	37.054		39.332	34.115
24	28.983	43.623	49.477	39.205	40.534	31.064	26.173	30.364	34.117	32.680	37.767	27.964	33.115	27.423	15.000	33.043	40.596	40.878	43.692	32.435	23.685	31.591	39.332		34.115

Gambar 2. Contoh Hasil Pengukuran Kemiripan dengan Euclidean Distance

Tabel 2. Statistik Nilai Hasil Pengukuran Euclidean Distance

Evaluasi Kemiripan	Nilai Maksimum	Nilai Minimum	Nilai rata-rata	Nilai Median
Nilai	58.57	8.72	31.59	31.57

Dari hasil pengukuran kemiripan, didapatkan nilai statistik seperti nilai maksimal kemiripan diantara dua pasangan data, nilai minimal, nilai rata-rata dan nilai median. Pengukuran nilai statistik digunakan untuk mendapatkan pengukuran nilai ambang kemiripan. Nilai ambang diperlukan untuk mendapatkan seberapa kuat kemiripan diantara data kinerja dosen satu dengan lainnya. Untuk mendapatkan nilai ambang, dilakukan pemetaan dari seluruh total pengukuran data. Selain itu, nilai ambang yang didapatkan dalam penelitian ini digunakan untuk menggabungkan mana data yang serupa dan terkelompok menjadi satu data dengan diwakili oleh nilai tertinggi. Dalam penelitian ini, hasil yang dari Tabel 2, digunakan pengukuran nilai ambang kemiripan yang dihitung menggunakan menggunakan persamaan (5).

$$\text{Kemiripan} = \frac{\text{maksimal nilai kemiripan}_{r_i r_{53}} + \text{minimum nilai kemiripan}_{r_i r_{53}}}{2}, \quad (5)$$

sehingga dari nilai yang didapatkan, nilai ambang kemiripan yang digunakan dalam reduksi data adalah sebesar 33.645. Data yang berhasil direduksi adalah sebesar 30.19% atau terdapat 16 data yang digabung dan tercatat sebagai data redundan. Sisa data sejumlah 37 data digunakan dalam model klasifikasi. Hasil klasifikasi metode naive bayes dalam bentuk confusion matrix ditunjukkan pada Gambar 3, dan evaluasi klasifikasi ditunjukkan pada Tabel 3.

		Predicted		
		Kualitas Baik	Kualitas Buruk	Σ
Actual	Kualitas Baik	80.0 %	27.3 %	18
	Kualitas Buruk	20.0 %	72.7 %	19
Σ		15	22	37

Gambar 3. Confusion Matrix Hasil Klasifikasi Naïve Bayes

Tabel 3. Hasil Klasifikasi Naive Bayes

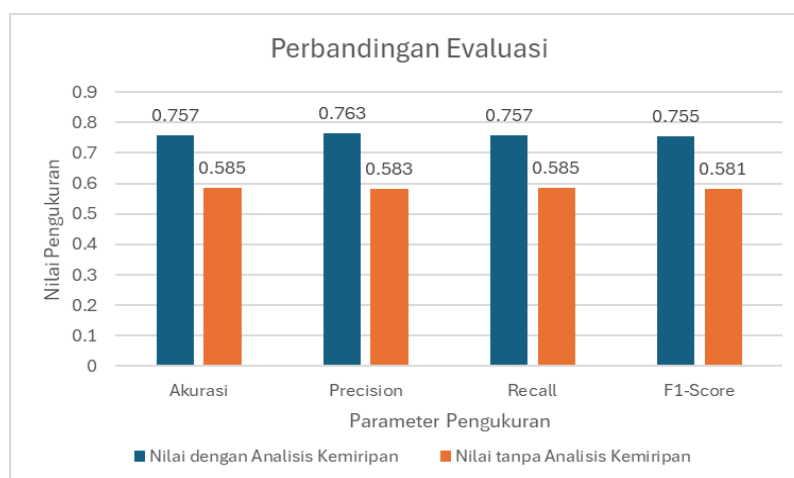
Parameter Evaluasi	Akurasi	Precision	Recall	F1-Score
Nilai	0.757	0.763	0.757	0.755

Hasil yang didapatkan dibandingkan dengan pengolahan data yang tidak menggunakan analisis kemiripan. Perbandingan antara hasil klasifikasi dengan penggunaan analisis kemiripan dan tanpa analisis kemiripan ditunjukkan pada Tabel 4.

Tabel 4. Detail Perbandingan Performa Klasifikasi

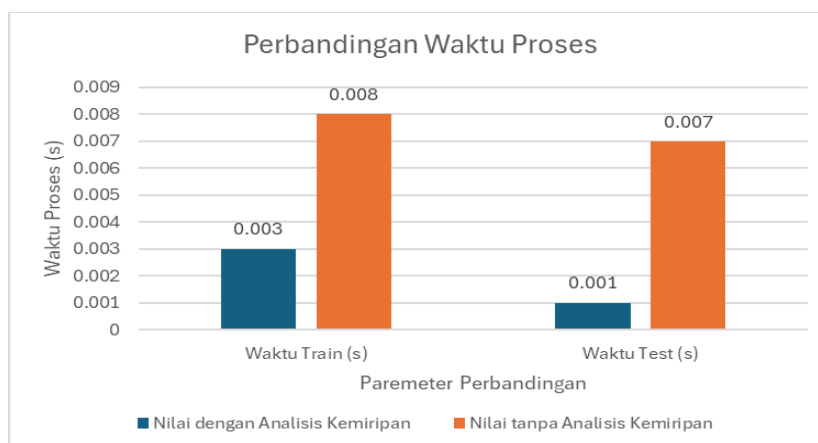
Parameter Evaluasi	Akurasi	Precision	Recall	F1-Score	Waktu Train (s)	Waktu Test (s)
Nilai dengan Analisis Kemiripan	0.757	0.763	0.757	0.755	0.003	0.001
Nilai tanpa Analisis Kemiripan	0.585	0.583	0.585	0.581	0.008	0.007

Hasil dari perbandingan menunjukkan bahwa model yang menggunakan analisis kemiripan memiliki performa yang lebih baik dibandingkan tanpa menggunakan analisis kemiripan. Nilai akurasi menunjukkan 22.72% lebih baik dimana nilai dari model tanpa analisis kemiripan hanya mencapai nilai 0.585, sedangkan model yang menggunakan analisis kemiripan mampu mencapai nilai 0.757. Perbandingan nilai precision menunjukkan bahwa model yang menggunakan analisis kemiripan memiliki nilai lebih tinggi yaitu 0.763 dibandingkan tanpa menggunakan analisis kemiripan dengan nilai 0.583 dengan selisih sebesar 23.59%. Perbandingan nilai recall, model dengan analisis kemiripan memiliki performa lebih baik sebesar 0.757 dibandingkan dengan tanpa analisis kemiripan sebesar 0.585 atau berbeda sebesar 22.72%. evaluasi terakhir pada parameter F1-score, model tanpa analisis kemiripan jauh lebih rendah yaitu dengan nilai 0.581 dibandingkan dengan model yang menggunakan analisis kemiripan yaitu 0.755 dengan perbedaan sebesar 23.05%. Hasil perbandingan dari evaluasi model klasifikasi ditunjukkan pada Gambar 4.



Gambar 4. Perbandingan Evaluasi Akurasi, Precision, Recall dan F1-score

Selain membandingkan analisis akurasi, precision, recall dan F1-score, penelitian ini melakukan perbandingan dari sisi waktu komputasi pada proses pelatihan dan pengujian. Hasil model dengan menggunakan analisis kemiripan memiliki performa komputasi yang lebih cepat yaitu 0.003 detik pada proses pelatihan dan 0.001 detik pada proses pengujian. Sedangkan pada model yang tidak menggunakan analisis kemiripan memiliki waktu komputasi lebih lama yaitu 0.008 detik pada proses pelatihan dan 0.007 detik pada proses pengujian. Hal ini menunjukkan bahwa model dengan analisis kemiripan mampu lebih cepat melakukan proses pelatihan dan pengujian data berdasarkan data yang digunakan. Hasil perbandingan dari komputasi pelatihan model klasifikasi dan pengujian ditunjukkan pada Gambar 5.



Gambar 5. Perbandingan Waktu Komputasi Pelatihan dan Pengujian Data Uji

4. Kesimpulan

Pada penelitian ini mengusulkan model optimasi metode klasifikasi Naive Bayes dengan pendekatan analisis kemiripan menggunakan pengukuran jarak yaitu Euclidean Distance pada studi kasus data pengukuran kinerja dosen. Pengukuran kemiripan bertujuan untuk mereduksi data ganda dan meningkatkan akurasi metode klasifikasi. Hasil yang didapatkan bahwa model memiliki performa klasifikasi yang baik, yaitu mampu memiliki performa di atas 0.75 dengan nilai akurasi 0.757, precision 0.763, recall 0.757 dan F1-score adalah 0.755. Walaupun hasil evaluasi menunjukkan model belum mampu menghasilkan nilai di atas 90%, model ini memiliki hasil yang lebih baik dibandingkan dengan model klasifikasi yang tanpa menggunakan metode analisis kemiripan. Penelitian selanjutnya akan dilakukan pengembangan terhadap analisis optimasi agar model mampu menghasilkan performa rata-rata di atas 90%. Selain itu, dapat dilakukan analisis terhadap perbandingan dari setiap parameter kinerja dosen dan menambah sub-parameter yang diukur dalam pengukuran kinerja dosen.

Daftar Pustaka

- [1] (Presiden Republik Indonesia), *Republic of Indonesia Law No.12 of 2012 Concerning Higher Education*, vol. 66. 2012, pp. 37–39.
- [2] R. S. Sumarsih, D. P. Sari, A. M. Yusuf, E. Setiawati, F. Hidayati, and A. T. Baedawi, “College student work as a form of community service and the implementation of the tri dharma of higher education,” vol. 2, no. 3, pp. 222–230, 2022.
- [3] T. Alsubaie, “The influence of participative leadership on employee performance: a case of the public sector in Saudi Arabia performance: A case of the public sector in Saudi Arabia,” pp. 1–140, 2021.
- [4] M. Bakare and Y. C. Ojeleye, “Participative Leadership Style and Employee Commitment in Federal College of Education (Technical) Gusau: Moderating role of Organizational Culture,” *Int. J. Intellect. Discourse*, vol. 3, no. 1, pp. 17–31, 2020.
- [5] D. S. Gunawan, Suwardi; Sirajang, Ratna Sari; Utomo, “Pengaruh Motivasi, Kepuasan Kerja, dan Komitmen Organisasional Terhadap Kinerja Karyawan (Studi Kasus PT. PLN Sektor Mahakam Samarinda),” *J. Din. Tek.*, vol. 13, no. 1, pp. 14–23, 2020.
- [6] M. H. Bhatti, Y. Ju, U. Akram, M. H. Bhatti, Z. Akram, and M. Bilal, “Impact of participative leadership on organizational citizenship behavior: Mediating role of trust and moderating role of continuance commitment: Evidence from the pakistan hotel industry,” *Sustain.*, vol. 11, no. 4, 2019,

- doi: 10.3390/su11041170.
- [7] M. Sağnak, "Participative Leadership and Change-Oriented Organizational Citizenship: The Mediating Effect of Intrinsic Motivation," *Eurasian J. Educ. Res.*, vol. 16, no. 62, pp. 181–194, 2016, doi: 10.14689/ejer.2016.62.11.
 - [8] F. Sinani, "The Effects of Participative Leadership Practices on Job Satisfaction for Highly Skilled Virtual Teams," *ProQuest Diss. Theses*, p. 173, 2016.
 - [9] D. Al Rahbi, A. Dhabbi University Khalizani Khalid, and M. Khan, "the Effects of Leadership Styles on Team Motivation," *Acad. Strateg. Manag. J.*, vol. 16, no. 2, pp. 1939–6104, 2017.
 - [10] M. Indrasari, "The Effect Of Organizational Culture, Environmental Work, Leadership Style On The Job Satisfaction And Its Impact On The Performance Of Teaching In State Community Academy Bojonegoro," *Sinergi J. Ilm. Ilmu Manaj.*, vol. 7, no. 1, pp. 58–73, 2017, doi: 10.25139/sng.v7i1.30.
 - [11] S. Langgeng Ratnasari, . R., G. Sutjahjo, and D. Yana, "Lecturer's Performance: Leadership, Organizational Culture, Work Motivation, and Work Behavior," *KnE Soc. Sci.*, vol. 3, no. 10, pp. 1–7, 2018, doi: 10.18502/kss.v3i10.3416.
 - [12] S. Langgeng Ratnasari, . R., G. Sutjahjo, and D. Yana, "Lecturer's Performance: Leadership, Organizational Culture, Work Motivation, and Work Behavior," *KnE Soc. Sci.*, vol. 3, no. 10, 2018, doi: 10.18502/kss.v3i10.3416.
 - [13] A. Tahnia, H. Fitria, and A. Wahidy, "The influence of organization culture on teacher performance of elementary school," *JPGI (Jurnal Penelit. Guru Indones.)*, vol. 6, no. 2, p. 460, 2021, doi: 10.29210/021071jpgi0005.
 - [14] H. Fitria, "The influence of organizational culture and trust through the teacher performance in the private secondary school in Palembang," *Int. J. Sci. Technol. Res.*, vol. 7, no. 7, pp. 82–86, 2018.
 - [15] B. Schneider and K. M. Barbera, *The Oxford Handbook of Organizational Climate and Culture*. Oxford University Press, 2014.
 - [16] I. Khan, A. R. Ahmad, N. Jabeur, and M. N. Mahdi, "An artificial intelligence approach to monitor student performance and devise preventive measures," *Smart Learn. Environ.*, vol. 8, no. 1, 2021, doi: 10.1186/s40561-021-00161-y.
 - [17] Kasmiaty, Baharuddin, M. N. Fattah, H. Nasaruddin, Y. Yusriadi, M. I. Usman, and Suherman, "The influence of leadership and work motivation on work effectiveness through discipline," *Proc. Int. Conf. Ind. Eng. Oper. Manag.*, vol. 565, no. INCoEPP, pp. 3648–3655, 2021.
 - [18] A. Sharma and I. Alvi, "Evaluating pre and post COVID 19 learning: An empirical study of learners' perception in higher education," *Educ. Inf. Technol.*, vol. 26, no. 6, pp. 7015–7032, 2021, doi: 10.1007/s10639-021-10521-3.
 - [19] E. KARAAHMETOĞLU, S. ERSÖZ, A. K. TÜRKER, V. ATEŞ, and A. F. İNAL, "Evaluation of Profession Predictions for Today and the Future with Machine Learning Methods : Emperical Evidence From Turkey," *Politek. Derg.*, vol. 26, no. 1, pp. 107–124, 2021, doi: 10.2339/politeknik.985534.
 - [20] J. Straub, "Machine learning performance validation and training using a 'perfect' expert system," *MethodsX*, vol. 8, no. August, p. 101477, 2021, doi: 10.1016/j.mex.2021.101477.
 - [21] M. M. E. I. Bali, M. P. Kumalasani, and D. Yunilasari, "Artificial Intelligence in Higher Education: Perspicacity Relation between Educators and Students," *J. Innov. Educ. Cult. Res.*, vol. 3, no. 2, pp. 146–152, 2022, doi: 10.46843/jiecr.v3i2.88.