



Implementasi Algoritma Multinomial Naïve Bayes, TF-IDF dan Confusion Matrix dalam Pengklasifikasian Saran Monitoring dan Evaluasi Mahasiswa Terhadap Dosen Teknik Informatika Universitas Dayanu Ikhsanuddin

Nurul Hidayah ^{*1}, Dodiman ²

^{1,2} Teknik Informatika, Universitas Dayanu Ikhsanuddin, Baubau
e-mail: nurul@unidayan.ac.id ¹, dodiman@unidayan.ac.id ²

* Corresponding Author

Received: 22 Maret 2024

Revised: 1 April 2024

Accepted: 5 Mei 2024

Abstrak

Monitoring dan Evaluasi (monev) Pembelajaran merupakan salah satu bentuk pengukuran kinerja dosen yang diselenggarakan Universitas Dayanu Ikhsanuddin yang melibatkan mahasiswa terhadap dosen. Saran dan opini yang diberikan mahasiswa pada saat monev cukup beragam sehingga perlu dilakukan analisis sentimen terhadap saran tersebut. Tujuan dari penelitian ini adalah untuk melakukan pengklasifikasian data saran monev pembelajaran oleh mahasiswa kepada dosen pengampu mata kuliah. Sebanyak 1037 data set digunakan dalam penelitian ini yang bersumber dari formulir *online* yang dibagikan kepada mahasiswa. Algoritma yang digunakan dalam penelitian ini adalah Multinomial Naïve Bayes untuk pengklasifikasian, TF-IDF untuk pembobotan teks dan *Confusion Matrix* untuk mengukur akurasi algoritma. Klasifikasi teks dilakukan dengan memberikan tiga label yaitu positif, negatif dan netral. Tahap penelitian ini dimulai dengan text processing, kemudian pembobotan dengan TF-IDF, lalu hasil klasifikasi ditampilkan dan diuji dengan *Confusion Matrix* yang terdiri dari unsur *precision*, *recall* dan akurasi. Hasil pengujian menunjukkan nilai *precision* yang didapatkan adalah 89%, *recall* adalah 85% dan akurasi sebanyak 85%. Hal ini menunjukkan bahwa algoritma memberikan hasil klasifikasi yang sangat baik.

Kata kunci: klasifikasi, analisis sentimen, multinomial naïve bayes

Abstract

Monitoring and Evaluation (M&E) of Learning is one form of assessing the performance of lecturers conducted by Dayanu Ikhsanuddin University involving students towards lecturers. The feedback and opinions provided by students during M&E are quite diverse, thus requiring sentiment analysis of such feedback. The aim of this research is to classify the data of M&E feedback from students to course instructors. A total of 1037 datasets were used in this study, sourced from online forms distributed to students. The algorithm employed in this research is Multinomial Naïve Bayes for classification, TF-IDF for text weighting, and Confusion Matrix for measuring algorithm accuracy. Text classification is conducted by assigning three labels: positive, negative, and neutral. The research process begins with text processing, followed by TF-IDF weighting, and then the classification results are displayed and tested with a Confusion Matrix consisting of precision, recall, and accuracy elements. The test results show a precision value of 89%, recall of 85%, and an accuracy of 85%. This indicates that the algorithm provides excellent classification results.

Keywords: classification, sentiment analysis, multinomial naïve bayes

PENDAHULUAN

Dosen merupakan sumber daya manusia utama yang berperan sebagai tenaga pendidik profesional serta ilmuwan yang memiliki tugas pokok dan bertujuan untuk dapat mengukusisi, mentransformasikan, menyebarluaskan, mengembangkan, menerapkan ilmu pengetahuan, seni dan teknologi melalui pendidikan, penelitian dan pengabdian atau pelayanan terhadap masyarakat (BAN-PT, 2011). Oleh karena begitu pentingnya peran dosen dalam sebuah program studi maka perlu diketahui tentang kinerja dosen dalam program studi tersebut.

Evaluasi Kinerja Dosen merupakan salah satu bentuk pengukuran kinerja dosen yang sesuai

dengan tugas pokok perguruan tinggi dalam menyelenggarakan pendidikan nasional yaitu Tridharma Perguruan Tinggi yang terdiri dari Pendidikan dan Pengajaran, Penelitian dan Pengembangan, serta Pengabdian kepada Masyarakat (UUD, 2012).

Dalam rangka mengetahui kinerja dosen yang melakukan pengajaran di program studi Teknik Informatika, Universitas Dayanu Ikhsanuddin, maka diselenggarakan monitoring dan evaluasi (monev) pembelajaran setiap akhir semester. Monev pembelajaran dilakukan dengan menyebarkan kuisioner secara *online* kepada mahasiswa, kemudian diolah oleh Unit Penjaminan Mutu lingkup universitas hingga lingkup program studi. Terdapat kolom saran pada kuisioner *online* yang

dibagikan kepada mahasiswa tersebut. Kolom saran ditujukan kepada dosen pengampu mata kuliah terkait proses pembelajaran.

Opini dalam kolom saran inilah yang dapat dimanfaatkan untuk melihat sentimen mahasiswa terhadap kinerja dosen selama proses pembelajaran berlangsung di semester tersebut. Opini-opini itu dapat diklasifikasikan ke dalam tiga kelas yaitu kelas positif, negatif dan netral dengan menggunakan *text mining*. *Text mining* dapat memberikan solusi dari permasalahan seperti pemrosesan, pengorganisasian/pengelompokan dan analisis *unstructured data* dalam jumlah besar. (Nurhuda et al., 2016). *Text mining* juga bertujuan untuk menemukan informasi berharga yang tersembunyi baik dari sumber informasi terstruktur dan tidak terstruktur (Firdaus & Firdaus, 2021).

Sebelumnya, telah dilakukan penelitian terkait di antaranya oleh (Salam et al., 2019) dengan judul Analisis Sentimen Opini Mahasiswa Terhadap Saran Kuesioner Penilaian Kinerja Dosen dengan Menggunakan TF-IDF dan K-Nearest Neighbor. Dalam penelitian tersebut dilakukan analisis sentimen terhadap kinerja dosen pada Fakultas Ilmu Komputer Universitas Brawijaya dengan menggunakan algoritma KNN, di mana hasil penelitian ini diperoleh rata-rata hasil dari tiga semester yaitu *Accuracy* sebesar 81%, *Precision* sebesar 81%, *Recall* sebesar 81%, dan *F1-Score* sebesar 81%. Kemudian terdapat pula penelitian yang dilakukan oleh (Yuyun et al., 2021) dengan judul Algoritma Multinomial Naïve Bayes untuk Klasifikasi Sentimen Pemerintah Terhadap Penanganan Covid-19 Menggunakan Data Twitter, yang didapatkan bahwa kategorisasi yang dilakukan menghasilkan perolehan akurasi sebesar 74%, *precision* sebesar 74% dan *recall* sebesar 74%, yang menandakan bahwa algoritma Multinomial Naïve Bayes memiliki nilai *fair classification* atau nilai diagnostik yang cukup baik.

Namun, penelitian yang ingin penulis angkat berfokus kepada analisis sentimen yang dilakukan terhadap opini mahasiswa, dengan kata lain penulis fokus melihat bagaimana respon mahasiswa terhadap kinerja dosen dalam kolom saran pada penilaian monev pembelajaran Teknik Informatika Unidayan. Analisis sentimen ini akan menggunakan algoritma Multinomial Naïve Bayes untuk memprediksi klasifikasi sentimen ke dalam tiga label yaitu positif, negatif dan netral. Algoritma ini penulis pilih karena mampu melakukan komputasi yang cepat dalam proses klasifikasi dokumen berupa teks dan mampu menyelesaikan kasus *multiple class* (Widyawati & Sutanto, 2020).

METODE PENELITIAN

Jenis Penelitian

Dalam penelitian ini penulis menggunakan jenis penelitian eksperimen dan studi kasus. Adapun penelitian eksperimen merupakan jenis penelitian yang dilakukan dengan mengimplementasikan serangkaian tindakan dengan tujuan membuktikan suatu konsep (Sugiyono, 2015). Pada penelitian ini, penulis melakukan eksperimen yang berupa proses pengambilan data saran dari mahasiswa selama

monev, kemudian mengklasifikasikan data tersebut menggunakan algoritma Multinomial Naïve Bayes.

Dalam hal ini terdapat beberapa proses eksperimental yang menggunakan pendekatan *text mining*. Langkah awal dari *text mining* mencakup kategorisasi teks, kemudian melakukan ekstraksi informasi, dan mengekstraksi kata. Metode *text mining* ini bertujuan untuk mengambil informasi dari sumber data dengan mengidentifikasi pola-pola menarik dalam teks. *Text mining* digunakan dalam berbagai konteks seperti klasifikasi, pengelompokan, ekstraksi informasi, dan pencarian informasi. Secara umum, *text mining* melibatkan tiga langkah utama: pra-pemrosesan teks, operasi penambahan teks, dan pascapemrosesan. Pra-pemrosesan teks melibatkan pemilihan data, pengklasifikasian, dan ekstraksi fitur untuk mengubah dokumen menjadi format yang lebih mudah untuk diproses, sesuai dengan tujuan pencarian yang berbeda (Firdaus & Firdaus, 2021).

Kemudian untuk studi kasus dilakukan pada data saran monev mahasiswa Program Studi Teknik Informatika Unidayan. Penulis juga menggunakan jenis penelitian yang bersifat deskriptif untuk menjelaskan data dalam deskriptif angka dan tabel. Kemudian, data yang telah diolah sebelumnya, dianalisis dengan menggunakan pendekatan kuantitatif yang dipaparkan dengan hasil perhitungan angka.

Waktu dan Tempat Penelitian

Waktu penelitian di lakukan pada bulan Februari-Maret 2024 dengan menggunakan data monev semester genap 2022-2023 dalam periode monev pembelajaran selama bulan Juni-Juli 2023. Tempat penelitian di laksanakan adalah pada Program Studi Teknik Informatika Unidayan.

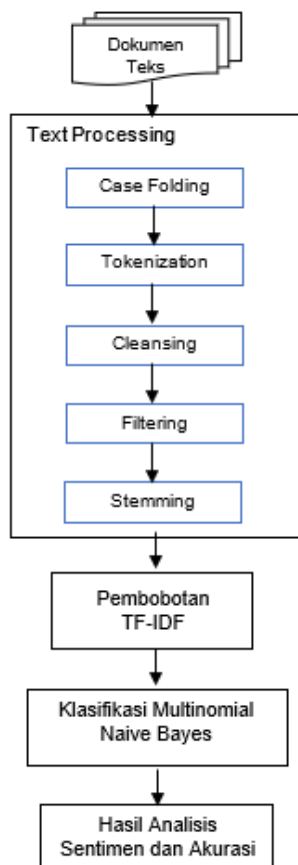
Target/Subjek Penelitian

Target penelitian adalah mendapatkan hasil analisis sentimen berupa pengklasifikasian data saran dari formulir monev yang diisi oleh mahasiswa program studi Teknik Informatika terhadap kinerja dosen selama pembelajaran genap 2022-2023 yang dilakukan rutin setiap semester. Data tersebut akan diklasifikasi menjadi tiga label yaitu label positif, negatif dan netral.

Adapun teknik *sampling* yang digunakan dalam penelitian ini adalah penggabungan *Simple Random Sampling* dan *Stratified Sampling*. *Simple Random Sampling*, atau yang sering disingkat sebagai *Random Sampling*, adalah suatu metode pengambilan sampel di mana setiap anggota populasi memiliki kesempatan yang sama untuk dipilih menjadi *sample*, sementara *Stratified Sampling* adalah kondisi di mana tingkatan dalam populasi cenderung memiliki variasi sehingga perlu dibedakan dalam tingkatannya, misal kelompok berat badan *underweight*, *normal* dan *overweight*, setelah itu pengambilan sample dilakukan masing-masing. Dalam kasus ini adalah data positif, negatif dan netral (Arieska & Herdiani, 2018).

Prosedur

Berikut ini merupakan alur atau prosedur penelitian yang dilakukan.



Gambar 1. Alur Prosedur Penelitian

Penelitian ini dibagi menjadi lima tahapan, tahap pertama adalah pengambilan atau pengumpulan dokumen teks, tahap kedua yaitu berupa *text processing*. *Text processing* merupakan tahapan penting dalam analisis teks yang bertujuan untuk menyiapkan dokumen teks yang tadinya berupa data tidak terstruktur menjadi data terstruktur dan siap digunakan. Tahapan ini melibatkan beberapa proses seperti *Cleansing*, *tokenizing*, *Case Folding*, *Filtering*, dan *Stemming*.

Setelah proses *text processing*, tahap berikutnya adalah pembobotan TF-IDF. Pembobotan TF-IDF digunakan untuk memberikan bobot pada masing-masing kata dalam dokumen berdasarkan frekuensi kemunculan kata tersebut dalam dokumen dan keunikannya dalam korpus teks secara keseluruhan.

Tahap selanjutnya adalah tahap klasifikasi menggunakan algoritma Multinomial Naive Bayes. Algoritma Multinomial Naive Bayes digunakan untuk melakukan klasifikasi teks berdasarkan probabilitas kemunculan kata-kata dalam kategori yang telah ditentukan, dalam konteks ini adalah sentimen negatif, positif, dan netral.

Tahap terakhir adalah pengujian data *testing* klasifikasi menggunakan *confusion matrix*.

Data, Instrumen, dan Teknik Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari data saran mahasiswa hasil penyebaran kuesioner *monitoring* dan evaluasi terhadap pembelajaran semester genap 2022-2023 pada bulan Juni-Juli 2023 yang memang dilakukan rutin oleh universitas setiap akhir semester. Dalam penelitian ini sejumlah 1038 saran diambil sebagai

data set yang digunakan untuk bahan analisis, di mana 990 data saran digunakan sebagai data latih dan sisanya 48 data digunakan untuk data *testing*. Penulis menggunakan data yang proporsional untuk memperoleh keakuratan yang baik. Sebab dalam penelitian *text mining*, keseimbangan atau proporsionalnya data akan mempengaruhi hasil klasifikasi, baik itu yang jumlah kelasnya kecil maupun besar (Afshoh, 2017) (Rozi et al., 2018).

Teknik Analisis Data

Labeling

Penelitian ini menggunakan pelabelan atau anotasi manual terhadap data set untuk menentukan pengklasifikasian teks saran, baik label positif, negatif dan netral. Proses labeling manual ini penting dalam pengembangan model klasifikasi karena menyediakan data pelatihan yang diberi label dengan benar untuk melatih model. Dengan memiliki data yang diberi label dengan benar, model klasifikasi dapat belajar untuk mengenali pola dan karakteristik dari teks yang terkait dengan setiap kategori atau label yang ditentukan. Contoh hasil pelabelan kelas dinyatakan pada tabel berikut:

Tabel 1. Contoh Pelabelan Positif, Negatif dan Netral

Id	Teks Saran	Label
1	Cara pemberian materinya sangat baik sekali	Positif
2	Agar kedepannya dosen tidak terlambat agar pembelaan bisa berlangsung lama	Negatif
3	Tidak ada komentar	Netral

Multinomial Naive Bayes

Multinomial Naive Bayes adalah sebuah metode *supervised learning*, sehingga sebelum melakukan *training*, data perlu dilabeli terlebih dahulu (Sabrani et al., 2020). Algoritma Multinomial Naive Bayes merupakan salah satu metode pembelajaran probabilistik didasarkan pada teorema Bayes yang digunakan dalam *Natural Language Processing* (NLP). Algoritma ini bekerja pada konsep *term frequency* yang berarti berapa kali kata tersebut muncul dalam sebuah dokumen. Model ini menjelaskan dua fakta yaitu apakah kata tersebut muncul dalam sebuah dokumen atau tidak serta frekuensinya kemunculan dalam dokumen (Singh et al., 2019). Probabilitas suatu dokumen *d* berada di kelas *c* dapat dihitung menggunakan Persamaan (1) (Sabrani et al., 2020).

$$P(p|n) \propto P(p) \prod_{1 \leq k \leq n} P(tk|p) \quad (1)$$

dimana:

$P(tk|p)$ adalah probabilitas munculnya *document text* (*tk*), *n* adalah jumlah dokumen dan *p* adalah polaritas. Kemudian untuk menghitung polaritasnya atau dokumen yang mempunyai kemiripan dirumuskan sebagai berikut, rumus 2.

$$P(tk|p) = \frac{\text{count}(tk|p)+1}{\text{count}(tp)+|V|} \quad (2)$$

dimana:

(tk | p) adalah jumlah tk muncul di dokumen *text* yang memiliki polaritas p dan jumlah (tp) berarti jumlah token yang ada di artikel berita dengan polaritas p.

Pembobotan Data Teks dengan TF-IDF

Teknik pembobotan data teks pada penelitian ini menggunakan metode TF-IDF. Metode TF-IDF adalah metode yang digunakan untuk menentukan bobot dari hubungan antara sebuah kata (term) dengan sebuah dokumen. Pendekatan ini menggabungkan dua konsep dalam perhitungan bobot, yaitu frekuensi kemunculan suatu kata dalam sebuah dokumen tertentu (TF) dan invers dari frekuensi dokumen yang mengandung kata tersebut (IDF). Berikut adalah contoh perhitungan TF-IDF, di mana rumus dari TF-IDF adalah (Syadid, 2019):

$$W_{dt} = TF_{dt} \times IDF_{ft} + 1 \quad (3)$$

dimana:

- W_{dt} = bobot dokumen ke-d terhadap kata ke-t
- TF_{dt} = banyaknya kata yang dicari pada sebuah dokumen
- IDF_{ft} = Inversed Document Frequency ($\log(\frac{N}{df})$)
- N = total dokumen
- df = banyak dokumen yang mengandung kata yang dicari.

Model Evaluasi dengan Confusion Matrix

Penelitian ini menggunakan model evaluasi atau pengujian berupa model *Confusion Matrix*, yaitu evaluasi yang dilakukan pada *Text Classification* di antaranya dapat menggunakan akurasi, *precision* dan *recall*. *Precision* adalah ukuran ketepatan antara respons sistem dengan informasi yang diinginkan oleh pengguna. *Recall* adalah ukuran keberhasilan sistem dalam menemukan kembali informasi yang relevan. Sedangkan akurasi adalah tingkat kesesuaian antara nilai prediksi dan nilai aktual. Berikut ini rumus dari *precision*, *recall* dan akurasi (Vidya, 2015):

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

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

Keterangan:

- TP = True Positive
- FN = False Negative
- FP = False Positive
- TN = True Negative

Kemudian, untuk mengukur performansi metode yang digunakan, digunakan AUC (*the area under curve*). AUC memiliki tingkat nilai diagnosa yaitu (Rosandy, 2016):

- 1) Akurasi bernilai 0,90 – 1,00 = *excellent classification*
- 2) Akurasi bernilai 0,80 – 0,90 = *good classification*
- 3) Akurasi bernilai 0,70 – 0,80 = *fair classification*

- 4) Akurasi bernilai 0.60 – 0,70 = *poor classification*
- 5) Akurasi bernilai 0.50 – 0.60 = *failure*

Berikut ini rumus *Area Under Curve* (Suwarno & Abdillah, 2016).

$$AUC = \frac{1}{2} \sum_{i=1}^n (x_{i+1} - x_i) (y_{i+1} - y_i) \quad (7)$$

HASIL PENELITIAN DAN PEMBAHASAN

Text Processing

Pada proses ini, data yang diperoleh dari kuesioner monev pembelajaran yang masih berupa data mentah kemudian diubah ke dalam bentuk yang mudah dipahami. Tahap pertama dari *text processing* adalah *Case Folding*, yaitu tahap pengubahan semua huruf besar (kapital) menjadi huruf kecil pada sebuah dokumen teks serta teks selain huruf tentunya akan dihilangkan (Vidya, 2015). Berikut ini adalah sampel hasil dari *Case Folding*:

Tabel 2. Hasil Case Folding

No	Sebelum Case Folding	Sesudah Case Folding
1	TERIMA KASIH!!	terima kasih
2	Mohon Ditingkatkan lagi ketepatan waktu mengajarnya agar proses pembelajaran dapat berjalan sesuai dengan waktu yang telah ditentukan	mohon ditingkatkan lagi ketepatan waktu mengajarnya agar proses pembelajaran dapat berjalan sesuai dengan waktu yang telah ditentukan
3	Sangat disiplin dalam pembelajaran	sangat disiplin dalam pembelajaran

Tahap kedua yaitu *Tokenization*, yaitu pemotongan string input yang dilakukan dengan memperhatikan setiap kata yang membentuknya. Secara sederhana, ini adalah langkah untuk memecah teks dari suatu kalimat atau paragraf menjadi segmen-segmen tertentu (Vidya, 2015). Berikut ini adalah sampel hasil *Tokenization*:

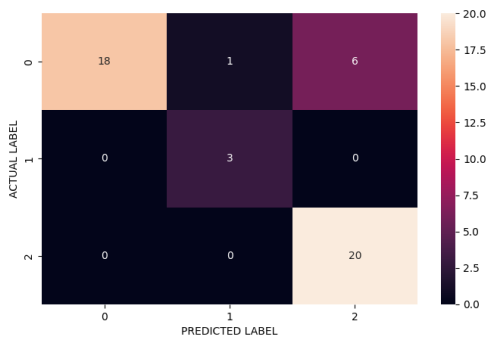
Tabel 3. Hasil Tokenization

No	Sebelum Tokenization	Sesudah Tokenization
1	terima kasih	[terima, kasih]
2	mohon ditingkatkan lagi ketepatan waktu mengajarnya agar proses pembelajaran dapat berjalan sesuai dengan waktu yang telah ditentukan	[mohon, ditingkatkan, lagi, ketepatan, waktu, mengajarnya, agar, proses, pembelajaran, dapat, berjalan, sesuai, dengan, waktu, yang, telah, ditentukan]
3	sangat disiplin dalam pembelajaran	[sangat, disiplin, dalam, pembelajaran]

Evaluasi	
Akurasi : 85.0 %	
Precision : 89.0 %	
Recall : 85.0 %	
Kategori : Counter({'Negatif': 25, 'Positif': 20, 'Netral': 3})	
Total Data Testing : 48	

Gambar 4. Hasil Pengukuran Akurasi, Precision dan Recall

Berdasarkan hasil pengukuran *confussion matrix*, didapatkan bahwa penilaian kinerja algoritma Multinomial Naive Bayes memiliki akurasi sebesar 85%, *precision* sebesar 89%, dan *recall* sebesar 85% dalam melakukan prediksi terhadap saran mahasiswa kepada kinerja dosen Teknik Informatika Unidayan. Berikut ini merupakan grafis dari *confussion matrix* yang ditampilkan oleh sistem.



Gambar 5. Grafik Confussion Matrix Sistem

Gambar 5 merupakan hasil *confussion matrix* yang diproses dalam sistem. Sistem menampilkan *confussion matrix* 3x3 karena terdapat 3 kelas yang digunakan yaitu positif, negatif dan netral. Pada *confussion matrix* di atas, kelas negatif di representasikan dengan angka 0, kelas positif dilambangkan dengan angka 2 dan kelas netral dilambangkan dengan angka 1. Di bawah ini, ditunjukkan hasil perhitungan *confussion matrix* yaitu akurasi, *precision* dan *recall*.

Tabel 6. Perhitungan Manual Confussion Matrix

		PREDICTED			Total Actual
		Negatif	Netral	Positif	
Actual	Negatif	18	1	6	25
	Netral	0	3	0	3
	Positif	0	0	20	20
Total Predicted		18	4	26	48

Dari Tabel 6 di atas, terlihat bahwa *confussion matrix* memiliki ukuran 3x3. Nilai *True Positive* (TP) ditampilkan dengan warna merah muda, dengan 18 TP untuk kelas Negatif, 3 TP untuk kelas Netral, dan 20 TP untuk kelas Positif. TP adalah jumlah dari kelas *True* yang diprediksi benar oleh sistem. Sedangkan *True Negative* (TN) adalah jumlah dari kelas *False* yang diprediksi benar oleh sistem. *False Positive* adalah kelas *True* yang salah diprediksi sebagai kelas *False* oleh sistem, dan *False Negative* adalah kelas *False* yang salah diprediksi sebagai kelas *True* oleh sistem.

Setelahnya dapat dihitung nilai *precision* yang dapat dilihat pada tabel 7 sebagai berikut:

Tabel 7. Hasil Perhitungan Manual Precision

Precision	= $\frac{\text{Correctly Predicted}}{\text{Total Predicted}}$ ATAU $\frac{TP}{(TP+FP)}$	
Kelas	Proses	Hasil
Negatif	18/18	1,0000
Netral	3/4	0,7500
Positif	20/26	0,7692
Weighted Avg Precision	Jumlah dari Total Actual Kelas/Jumlah Data*Precision	
	$(25/48*1,000) + (3/48*0,7500) + (20/48*0,7690)$	

Dari Tabel 7, *precision* dihitung dengan membagi jumlah prediksi benar dengan total keseluruhan yang diprediksi oleh sistem. Hasilnya adalah sebagai berikut: Kelas Negatif memiliki *precision* 1,000, kelas Netral memiliki *precision* 0,7500, dan *precision* untuk kelas Positif adalah 0,7692. Untuk rata-rata *precision* menggunakan *Weighted Average*, didapatkan nilai 0,8882, yang jika dijadikan dalam persentase adalah 89%.

Tabel 8. Hasil Perhitungan Manual Recall

Recall	= $\frac{\text{Corectly Classified}}{\text{Total Actual}}$ ATAU $\frac{TP}{(TP+FN)}$	
Kelas	Proses	Hasil
Negatif	18/25	0,7200
Netral	3/3	1,0000
Positif	20/20	1,0000
Weighted Avg Recall	Jumlah dari Total Actual Kelas/Jumlah Data*Recall	
	$(25/48*0,7200) + (3/48*1,000) + (20/48*1,000)$	

Berdasarkan Tabel 8, *recall* dihitung dengan membagi jumlah klasifikasi yang benar dengan total keseluruhan aktual. Hasilnya adalah sebagai berikut: Kelas Negatif memiliki *recall* sebesar 0,7200, kelas Netral memiliki *recall* sebesar 1,000, dan *recall* untuk kelas Positif adalah 1,000. Untuk rata-rata menggunakan *Weighted Average*, didapatkan nilai 0,8542, yang jika dijadikan dalam persentase adalah 89%.

Kemudian berikut ini merupakan perhitungan manual untuk akurasi:

Tabel 9. Hasil Perhitungan Manual Akurasi

Akurasi	= $\frac{\text{Total Correctly Clasified}}{\text{Actual}}$	
	Proses	Hasil
	41/48	0,8542

Dari Tabel 9, rumus akurasi dinyatakan sebagai jumlah total klasifikasi benar dibagi dengan

jumlah data aktual. Hasilnya adalah 0,8542, yang jika dijadikan dalam persentase adalah 85%.

Setelah itu, dihitung nilai AUC (*the area under curve*) yaitu dengan $(\text{Rata-rata Recall} + \text{Rata-rata Specificity})/2$. Di mana nilai *specificity* didapat dengan rumus $TN/(TN+FP)$. Sehingga hasilnya adalah:

Tabel 10. Hasil Perhitungan AUC

Nilai AUC	= (Rata-rata Recall+ Rata-rata Specificity)/2	
	Proses	Hasil
	$(0,9067+0,9147)/2$	0,9107

Dari hasil perhitungan *precision*, *recall* dan akurasi terhadap tiga label kelas yaitu positif, negatif dan netral, dapat diketahui bahwa akurasi data dipengaruhi oleh data latih dan juga data uji. Semakin banyak referensi kata dari sebuah data latih, yang di mana memungkinkan semakin banyak frekuensi kata dalam data set, maka semakin baik akurasi prediksi terhadap data uji. Melihat hasil *precision* sebesar 89%, *recall* sebesar 85% dan akurasi sebesar 85% serta nilai AUC sebesar 0,91 dapat disimpulkan bahwa algoritma Multinomial Naïve Bayes ini memiliki tingkat *excellent classification*.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan analisis di atas, dapat disimpulkan bahwa algoritma Multinomial Naïve Bayes dengan menggunakan pembobotan TF-IDF untuk faktorisasi dapat melakukan klasifikasi teks dalam kelas positif, negatif dan netral terhadap data saran mahasiswa terhadap kinerja dosen Teknik Informatika Unidayan. Adapun hasil prediksi yang dilakukan oleh algoritma menghasilkan nilai *precision* sebesar 89%, nilai *recall* sebesar 85% dan akurasi sebesar 85%, maka didapatkan nilai AUC sebanyak 0,91 yang menyakan bahwa algoritma memiliki tingkat klasifikasi yang sangat baik.

Saran

Saran untuk penelitian berikutnya adalah agar data set yang digunakan lebih banyak lagi misalnya data set dapat diambil selama 4 semester. Selain itu, dapat melakukan kombinasi algoritma klasifikasi lainnya yang dapat membantu Multinomial Naïve Bayes dalam proses pengklasifikasian dokumen.

DAFTAR PUSTAKA

Afshoh, F. (2017). Analisa Sentimen Menggunakan Naïve Bayes Untuk Melihat Persepsi Masyarakat Terhadap Kenaikan Harga Jual Rokok Pada Media Sosial Twitter. Informatika, Program Studi Komunikasi, Fakultas Informatika, D A N Surakarta, Universitas Muhammadiyah, 12.

Arieska, P. K., & Herdiani, N. (2018). Pemilihan Teknik Sampling Berdasarkan Perhitungan

Efisiensi Relatif. *Jurnal Statistika*, 6(2).

Firdaus, A. F., & Firdaus, W. I. (2021). Text mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi : (Sebuah Ulasan). *JUPITER: Jurnal Penelitian Ilmu Dan Teknologi Komputer*, 13(1).

Informatikologi. (2021). Pembobotan Kata atau Term Weighting TF-IDF. *Informatikologi.Com Portal Belajar, Berbagi Ilmu Informatika Dan Komputer*.

Nurhuda, F., Widya Sihwi, S., & Doewes, A. (2016). Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier. *Jurnal Teknologi & Informasi ITSmart*, 2(2), 35. <https://doi.org/10.20961/its.v2i2.630>

Rosandy, T. (2016). Perbandingan Metode Naive Bayes Classifier Dengan Metode Decision Tree (C4.5) Untuk Menganalisa Kelancaran Pembiayaan (Study Kasus : KSPPS / BMT AL-FADHILA. *Jurnal Teknologi Informasi Magister Darmajaya*, 2(01), 52–62.

Rozi, I. F., Hamdana, E. N., & Iqbal Alfahmi, M. B. (2018). Pengembangan Aplikasi Analisis Sentimen Twitter Menggunakan Metode Naïve Bayes Classifier (Studi Kasus SAMSAT Kota Malang). *Jurnal Informatika Polinema*, 4(2). <https://doi.org/10.33795/jip.v4i2.164>

Sabrani, A., Wedashwara W., I. G. W., & Bimantoro, F. (2020). Multinomial Naïve Bayes untuk Klasifikasi Artikel Online tentang Gempa di Indonesia. *Jurnal Teknologi Informasi, Komputer, Dan Aplikasinya (JTika)*, 2(1), 89–100. <https://doi.org/10.29303/jtika.v2i1.87>

Salam, N. S. N., Supianto, A. A., & Perdanakusuma, A. R. (2019). Analisis Sentimen Opini Mahasiswa Terhadap Saran Kuesioner Penilaian Kinerja Dosen dengan Menggunakan TF-IDF dan K-Nearest Neighbor. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(6).

Singh, G., Kumar, B., Gaur, L., & Tyagi, A. (2019). Comparison between Multinomial and Bernoulli Naïve Bayes for Text Classification. 2019 International Conference on Automation, Computational and Technology Management, ICACTM 2019. <https://doi.org/10.1109/ICACTM.2019.8776800>

Sugiyono. (2015). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (22nd ed.). Alfabeta.

Suwarno, & Abdillah, A. (2016). Penerapan Algoritma Bayesian Regularization Backpropagation Untuk Memprediksi Penyakit Diabetes. 39(2), 98–106.

Syadid, F. (2019). Analisis Sentimen Komentar Netizen Terhadap Calon Presiden Indonesia 2019 Dari Twitter Menggunakan Algoritma Term Frequency-Invers Document Frequency (Tf-Idf) Dan Metode Multi Layer Perceptron (Mlp) Neural Network. In *Skripsi Universitas Islam Negeri Syarif Hidayatullah Jakarta. Univeristas Islam Negeri Syarif Hidayatullah*.

Vidya, N. A. (2015). *Twitter Sentiment Analysis*

Terhadap Brand Reputation: Studi Kasus PT XL AXIATA Tbk. [Univeristas Indonesia]. In Tesis Univeritas Indonesia (Vol. 151). <https://doi.org/10.1145/3132847.3132886>

Warni, E., Alimuddin, A. A. P., & Firman, R. (2018). Sosial Media Mining Untuk Pengawasan Penyakit Demam Berdarah Di Indonesia. 4(November), 193–199.

Widyawati, W., & Sutanto, S. (2020). Perbandingan Kinerja Variasi Naïve Bayes Multivariate Bernoulli Dan Naïve Bayes Multinomial Dalam Pengklasifikasian Dokumen Teks. *Journal of Innovation And Future Technology (IFTECH)*, 2(1). <https://doi.org/10.47080/ifttech.v2i1.859>

Yuyun, Nurul Hidayah, & Supriadi Sahibu. (2021). Algoritma Multinomial Naïve Bayes Untuk Klasifikasi Sentimen Pemerintah Terhadap Penanganan Covid-19 Menggunakan Data Twitter. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(4). <https://doi.org/10.29207/resti.v5i4.3146>