

Analisis Sentimen pada Indeks Kinerja Dosen Fakultas SAINTEK UIN Sunan Kalijaga Menggunakan Naive Bayes Classifier

Bama Abpama Sevsa¹, M. Didik R Wahyudi²

^{1,2}Program Studi Teknik Informatika, Fakultas Sains dan Teknologi

UIN Sunan Kalijaga

Jl. Marsda Adisucipto Yogyakarta, DIY, Indonesia

Email: ¹bama4651@gmail.com, ²m.didik@uin-suka.ac.id

Masuk: 11 Juni 2019 ; Direvisi: 22 Oktober 2019; Diterima: 23 Oktober 2019

Abstract. *UIN Sunan Kalijaga Yogyakarta has a lecturer performance appraisal system commonly referred to as IKD (Lecturer Performance Index) in the Academic Information System. At the end of semester, students give the assessment by filling in a questionnaire concerning judgement and suggestion towards the campus learning activities. The questionnaire results may reveal which are the negative, neutral or positive sentiments by conducting sentiment analysis. The sentiment analysis used is Naive Bayes Classifier. In this study, 8249 data were used, with the composition of 3000 training data with labels and 5249 test data without labels. The sentiment analysis of the critical data and suggestions of the IKD resulted in greater accuracy using the TF-IDF weighting, which was 73.9% compared to TF weighting which was 72.8%. This accuracy value is obtained from 3000 training data that has been labeled using the evaluation method of the K-Fold Cross Validation classification model.*

Keywords: *Sentiment Analysis, IKD (Lecturer Performance Index), Naive Bayes Classifier, K-fold cross validation.*

Abstrak. *UIN Sunan Kalijaga Yogyakarta memiliki sistem penilaian kinerja dosen yang biasa disebut dengan IKD (Indeks Kinerja Dosen) pada Sistem Informasi Akademik. Pada akhir semester, para mahasiswa akan memberikan penilaian dengan mengisi kuesioner mengenai kritik dan saran bagi kegiatan perkuliahan. Data-data dari hasil kuesioner yang didapat bisa menunjukkan apakah adanya sentimen negatif, netral, ataupun positif dengan cara melakukan analisis sentimen. Analisis sentimen yang dipergunakan adalah Naive Bayes Classifier. Penelitian ini menggunakan 8249 data dengan komposisi 3000 data latih dengan label dan 5249 data uji tanpa label. Analisis sentimen data kritik dan saran IKD ini menghasilkan akurasi yang lebih besar menggunakan pembobotan TF-IDF yaitu 73,9% dibandingkan dengan pembobotan TF yaitu 72,8%. Nilai akurasi ini didapatkan dari 3000 data latih yang sudah diberi label menggunakan metode evaluasi model klasifikasi K-Fold Cross Validation.*

Kata Kunci: *Analisis Sentimen, Indeks Kinerja Dosen, Naive Bayes Classifier, K-Fold Cross Validation*

1. Pendahuluan

Penilaian kinerja secara sistematis kepada setiap karyawan dalam melaksanakan pekerjaannya, dapat memberikan gambaran secara sistematis tentang kekuatan dan kelemahan kerjanya [1]. UIN Sunan Kalijaga Yogyakarta memiliki sistem penilaian kinerja dosen yang mengajar dalam Sistem Informasi Akademik yang biasa disebut dengan IKD (Indeks Kinerja Dosen).

Penilaian IKD (Indeks Kinerja Dosen) UIN Sunan Kalijaga Yogyakarta diisi oleh mahasiswa pada akhir semester sebelum UAS (Ujian Akhir Semester) dilaksanakan. Pada IKD terdapat kuesioner pengisian yang terdiri dari pilihan-pilihan dari rentang 1-5 dan kritik dan saran. Kritik dan saran ini berupa komentar teks yang terdiri dari komentar negatif maupun positif. Namun, untuk menilai apakah komentar yang ditulis mahasiswa termasuk negatif atau positif

masih menggunakan penilaian manual. Penilaian manual ini tentu saja menyita banyak waktu karena komentar ditulis oleh banyak mahasiswa dan untuk setiap matakuliah yang ada.

Seiring dengan berkembangnya teknologi informasi dan komunikasi, banyak konsep teknologi yang bermunculan untuk menunjang kebutuhan akan informasi dan komunikasi. Salah satu konsep yang ditawarkan untuk mengolah data yang berjumlah besar adalah data mining. Data mining merupakan proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis [2]. Dalam data mining untuk mengolah data mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini disebut dengan Analisis Sentimen.

Pada penelitian ini akan dilakukan Analisis Sentimen Kritik Dan Saran Pada Indeks Kinerja Dosen Fakultas Sains Dan Teknologi UIN Sunan. Data kritik dan saran tersebut akan diklasifikasi secara otomatis menjadi tiga jenis klasifikasi yaitu positif, negatif, dan netral dengan metode *Naive Bayes Classifier*. Kelebihan metode *Naive Bayes Classifier* adalah sederhana tetapi memiliki akurasi yang tinggi [3]. Selain itu metode ini terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar [4].

2. Tinjauan Pustaka

Penelitian tentang analisis sentimen dilakukan oleh Ahmad Fathan Hidayatullah dan Azhari SN (2014). Penelitian ini menghasilkan klasifikasi tweet berdasarkan sentimen dan kategori dengan *Naive Bayes Classifier*. Hasil akurasi pengujian klasifikasi dengan fitur *term frequency* diperoleh sebesar 79,91% sedangkan fitur TF-IDF didapatkan akurasi sebesar 79,68%. Klasifikasi dengan *Support Vector Machine* menghasilkan akurasi 83,14% untuk fitur *term frequency* dan 82,69% untuk fitur TF-IDF. Metode *Support Vector Machine* menghasilkan akurasi performansi yang lebih baik daripada metode *Naive Bayes* [5].

Penelitian yang dilakukan oleh Juen Ling, (2014) menggunakan seleksi fitur berupa Seleksi Fitur Chi Square yang digunakan untuk mengurangi fitur-fitur yang tidak sesuai dalam proses klasifikasi *Naive Bayes Classifier*. Penelitian ini menghasilkan bahwa seleksi fitur *Chi Square* baik digunakan dalam penyeleksian fitur dibandingkan dengan metode *frequency-based*. Dimana Klasifikasi oleh NBC pada data uji negative memperoleh ketepatan sebesar 72% sedangkan untuk data uji positif memperoleh 96%. Secara keseluruhan data uji, klasifikasi oleh NBC memperoleh akurasi sebesar 83%. Terdapat kesalahan klasifikasi karena pada data uji terdapat fitur yang muncul pada bukan kategorinya [6].

Penelitian yang dilakukan oleh Amir Hamzah (2014), menunjukkan bahwa Algoritma NBC dapat melakukan klasifikasi opini yang diberikan oleh user dengan benar. Akurasi klasifikasi opini rata-rata untuk seluruh dokumen yang diujikan adalah 85,95%. Penurunan penggunaan jumlah dokumen latihan yang sedikit tidak terlalu menurunkan kinerja algoritma dalam klasifikasi, karena dengan dokumen latihan hanya 40% masih memberikan akurasi dia atas 80% [7].

Penelitian yang dilakukan oleh Ghulam Asrofi Buntoro (2016) menunjukkan bahwa banyak sentimen yang ditujukan kepada calon Gubernur DKI Jakarta 2017. Nilai akurasi tertinggi didapat saat menggunakan metode klasifikasi *Naive Bayes Classifier* (NBC) untuk klasifikasi data AHY, dengan nilairata-rata akurasi mencapai 95%, nilai presisi 95%, nilai recall95% nilai TP rate96,8% dan nilai TN rate84,6%. Dalam penelitian ini juga dapat diketahui metode klasifikasi *Naive Bayes Classifier* (NBC) lebih tinggi akurasinya untuk klasifikasi sentimen Tweet Bahasa Indonesia dibandingkan dengan metode klasifikasi *Support Vector Machine* (SVM) [8].

3. Metodologi Penelitian

Metode penelitian yang dipakai adalah metode eksperimen atau penelitian terapan, yaitu dengan menerapkan metode *Naive Bayes Classifier* untuk proses klasifikasi analisis sentimen kritik dan saran pada Indeks Kinerja Dosen (IKD) Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta. Alat bantu yang dipergunakan dalam memproses penelitian ini adalah

bahasa pemrograman python. Langkah-langkah yang dilakukan dalam penelitian ini ialah studi pustaka, pengumpulan data, seleksi data, *preprocessing*, analisis dan validasi, dan implementasi.

Data yang dipergunakan dalam penelitian ini adalah data kritik dan saran pada IKD Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta yang diambil dari pihak Pusat Teknologi Informasi dan Pangkalan Data (PTIPD) UIN Sunan Kalijaga Yogyakarta. Data yang dipakai adalah data 3 tahun ajaran, yaitu 2014/2015, 2015/2016, dan 2016/2017.

Pada tahap seleksi data, proses yang dilakukan adalah melakukan seleksi atau pemilihan data yang dibutuhkan dan akan digunakan dari data mentah yang di berikan PTIPD UIN Sunan Kalijaga. Data yang diberikan PTIPD merupakan data dalam bentuk excel. Data diseleksi dengan membuang data kritik dan saran yang mengalami duplikasi, kosong, dan tidak sesuai. Kemudian data dibagi menjadi 2 bagian, yaitu data latih dan data uji untuk klasifikasi otomatis. Data latih kemudian diberi label sebagai bahan pembelajaran model klasifikasi.

Data *preprocessing* merupakan proses mempersiapkan data seperti membersihkan data dari noise ataupun merubah format data [9]. *Pre-processing* merupakan pemrosesan awal dokumen agar diperoleh suatu data yang baik sebagai bahan pembelajaran analisis sentimen. *Pre-processing* sendiri dimulai dari membuang tanda baca dan diubah menjadi huruf kecil. Selanjutnya dilakukan pemecahan menjadi token-token. Kemudian dilakukan *Convert Slangword* dan *Convert Negasi*. setelah itu membuang kata-kata yang tidak penting dan pengubahan kata menjadi kata dasar.

Setelah data selesai di *pre-processing*. Kemudian data tersebut dilakukan pembobotan dengan melihat tingkat kemunculannya atau term frequency. Kemudian data latih diolah dengan menggunakan metode *Naive Bayes Classifier*. Pada pemrosesan data ini akan menghasilkan bahan pembelajaran yang akan digunakan dalam implementasi klasifikasi otomatis. Pada bagian ini juga dilakukan evaluasi model klasifikasi dengan *k-fold cross validation* untuk menghitung akurasi dari klasifikasi sentimen menggunakan metode *Naive Bayes Classifier*.

Data latih yang sudah diolah akan dijadikan bahan pembelajaran, selanjutnya akan dilakukan proses klasifikasi sentimen secara otomatis menggunakan metode *Naive Bayes Classifier*. Klasifikasi ini akan mengklasifikasikan data kritik dan saran ke dalam positif, negatif, maupun netral.

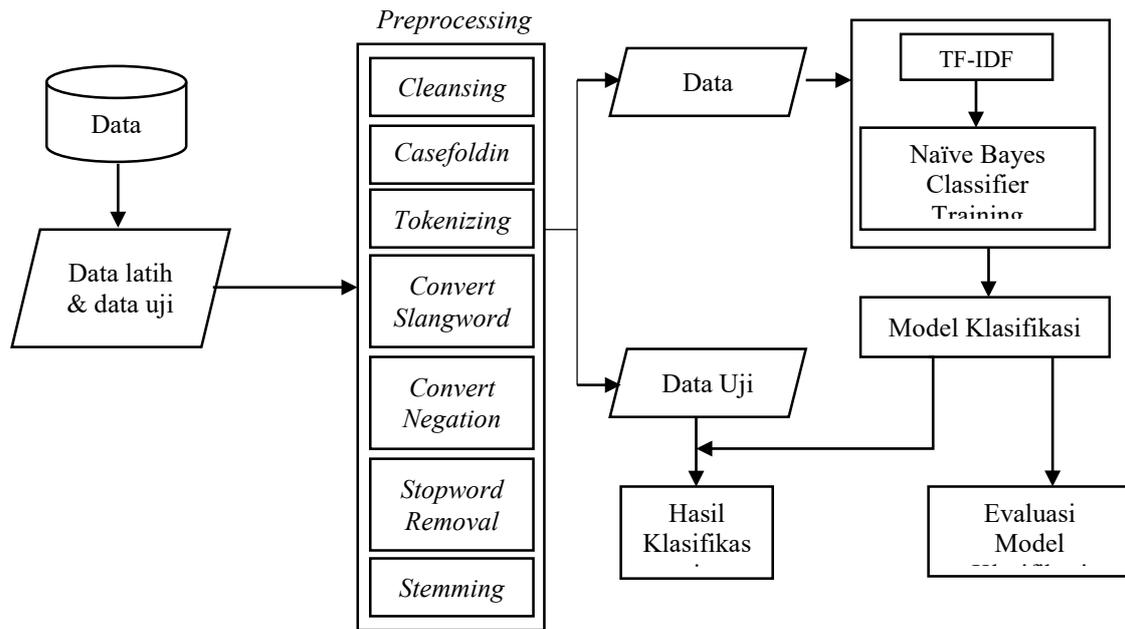
4. Hasil dan Diskusi

Data yang digunakan dalam penelitian ini adalah data kritik dan saran pada IKD selama tiga tahun ajaran, yaitu 2014/2015, 2015/2016, dan 2016/2017 sebanyak 8249 data. Data ini dibagi menjadi data latih sebanyak 3000 data dengan label dan data uji tanpa label sebanyak 5249 data.

Data kritik dan saran dari PTIPD UIN sunan Kalijaga Yogyakarta berbentuk *file excel* dirubah terlebih dahulu menjadi *file csv* untuk dimasukkan ke dalam database. Data *file excel* diseleksi terlebih dahulu untuk membuang data-data yang kosong dan tidak sesuai untuk diproses. Selanjutnya data kritik dan saran yang ada pada database dilakukan *preprocessing* dan disimpan dalam database hasil *pre-processing*.

Data latih kritik dan saran yang berjumlah sebanyak 3000 data, dibuat untuk model klasifikasi. Model klasifikasi ini selanjutnya disimpan dan akan digunakan pada pengujian data. Setelah dilakukan proses menghitung akurasi dari klasifikasi sentimen pada data kritik dan saran menggunakan uji validasi *k-fold cross validation* dengan *10-fold*.

Pada pengujian, hasil *pre-processing* data uji dilakukan klasifikasi atau pelabelan secara otomatis dengan cara menerapkan model klasifikasi yang telah dibuat pada data latih sebelumnya. Hasil klasifikasi ini kemudian ditampilkan dalam grafik dan siap digunakan sebagai informasi. Secara umum proses klasifikasi sentimen disajikan pada gambar 1:



Gambar 1. Langkah-langkah Penelitian

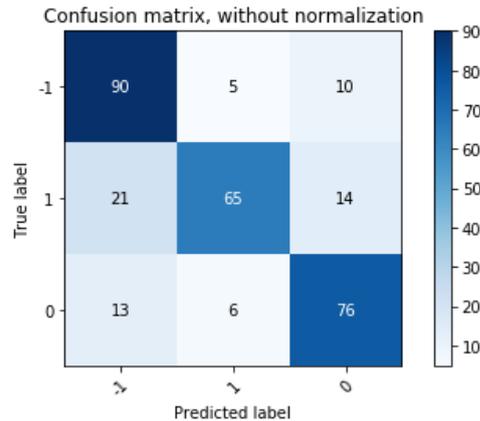
Data tersebut akan diseleksi dengan membuang data kritik dan saran yang mengalami duplikasi, kosong, dan tidak sesuai. Kemudian data dibagi menjadi 2 bagian, yaitu data latih sebanyak 3000 data dan data uji untuk klasifikasi otomatis sebanyak 5249 data.

Data ini kemudian akan dilakukan *pre-processing*. Berikut ini langkah-langkah dalam *pre-processing*: (1). *Cleansing*. Pada bagian *cleansing* ini dilakukan pembersihan terhadap tanda baca dan komponen non-huruf lainnya. (2). *Casefolding*. *Casefolding* digunakan untuk menguah huruf besar menjadi huruf kecil. Hal ini sangat diperlukan karena besar kecilnya huruf dapat mempengaruhi perhitungan TF (*Term Frequency*). (3). *Tokenizing*. *Tokenizing* atau tokenisasi merupakan proses proses untuk memisah-misahkan kata atau token. Tahapan tokenisasi ini dimulai dari membagi kritik dan saran yang dipisahkan dengan karakter spasi. Penelitian ini mempergunakan proses tekenisasi unigram. (4). *Convert Slangword*. Data kritik dan saran yang ditulis oleh mahasiswa terkadang menggunakan kata-kata yang tidak baku. Oleh karena itu diperlukan *Convert Slangword*. *Convert Slangword* adalah tahapan untuk mengubah kata-kata yang tidak baku menjadi kata baku yang sesuai dengan EYD (Ejaan Yang Disempurnakan). (5). *Convert Negation*. Beberapa kata yang mengandung negasi dapat mempengaruhi nilai sentimen. Kata-kata yang bersifat negasi seperti “bukan”, “tidak”, “enggak”, “ga”, “jangan”, “nggak”, “tak”, dan “gak” [10]. Dalam penelitian ini penulis menggunakan dua kata yang bersifat negasi, yaitu tidak dan kurang. Misalkan saja ada kata “tidak telat”, telat termasuk sentimen negatif padahal kata “tidak telat” seharusnya termasuk sentimen positif. Oleh karena ini diperlukan *convert negation* agar diantara kata-kata negasi tidak mempengaruhi hasil sentimen dengan cara menghilangkan spasi kata negasi dengan kata yang mengikutinya. (6). *Stopword Removal*. Proses *Stopword Removal* adalah proses untuk menghilangkan kata-kata tidak memiliki pengaruh apapun dalam proses sentimen analisis dan sering muncul. Kata yang termasuk *stopword* adalah kata hubung, kata ganti orang, penunjuk waktu, dan kata tanya. Contohnya, yaitu saya, kamu, di, dengan, dan apa. (7). *Stemming*. *Stemming* merupakan proses mengubah bentuk kata yang berimbuhan menjadi kata dasar, dengan cara menghilangkan imbuhan pada kata tersebut. Contohnya, yaitu kata “bermanfaat” menjadi kata “manfaat”.

Data yang sudah melalui tahap *cleansing*, kemudian diberi label sebagai *supervised learning* untuk model klasifikasi. Sebanyak 3000 data diberikan label secara manual. Setiap data kritik dan saran akan diklasifikasikan ke dalam 3 (tiga) kelas, yaitu positif (1) berjumlah 1000 data, negatif (-1) berjumlah 1000 data, dan netral (0) berjumlah 1000 data. Training set yang

diberikan label manual ini kemudian diuji dengan mempergunakan *Confusion Matrix*. *Confusion Matrix* adalah suatu matriks yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi. Setiap baris *confusion matrix* mewakili hasil klasifikasi yang seharusnya dan setiap kolom mewakili hasil klasifikasi yang dilakukan oleh sistem [11]. Hasil perhitungan *confusion matrix* menunjukkan bahwa training set tersebut memiliki akurasi 77%. Distribusi akurasi prediksi untuk setiap kelas sebagaimana ditunjukkan pada gambar 2.

```
Confusion matrix, without normalization
[[90  5 10]
 [21 65 14]
 [13  6 76]]
```



Gambar 2: Confusion matrix training set

Jumlah data untuk setiap contoh data kritik dan saran yang telah diberi label disajikan dalam tabel 1 berikut ini:

Tabel 1. Data Kritik dan Saran yang telah diberi label

Dokumen	Kritik dan Saran	Label
KS1	dosen sangat tepat waktu kuliah	1
KS2	bapak ajar sangat baik	1
KS3	moga manfaat dosen mahasiswa	0
KS4	mohon maaf lahir batin	0
KS5	jangan sering telat	-1
KS6	sampai materi terlalu cepat	-1

Langkah berikutnya adalah mencari *Term Frequency*. *Term frequency* (tf) merupakan suatu metode yang digunakan untuk mencari bobot suatu kata dalam dokumen kunci di setiap kategori dan mencari kata kunci yang hampir mirip dengan kategori yang tersedia [12]. Berikut ini tabel *Term Frequency* dari data kritik dan saran ditunjukkan pada tabel 2 berikut ini :

Tabel 2. Term Frequency

No.	term	Dokumen					
		ks1	ks2	ks3	ks4	ks5	ks6
1	ajar	0	1	0	0	0	0
2	baik	0	1	0	0	0	0
3	bapak	0	1	0	0	0	0
4	batin	0	0	0	1	0	0
5	cepat	0	0	0	0	0	1

6	dosen	1	0	1	0	0	0
7	jangan	0	0	0	0	1	0
8	kuliah	1	0	0	0	0	0
9	lahir	0	0	0	1	0	0
10	maaf	0	0	0	1	0	0

Selanjutnya dilakukan proses klasifikasi *Naive Bayes Classifier* berdasarkan data latih sebanyak 3000 dan pembobotan TF dengan cara menghitung nilai peluang pada masing-masing kelas berdasarkan contoh dokumen data kritik dan saran diatas dengan rumus sebagai berikut:

$$P(i) = \frac{N(i)}{N} \tag{1}$$

Keterangan:

P(i) = peluang sentimen i

N(i) = jumlah dokumen yang masuk dalam sentimen i

N = jumlah keseluruhan dokumen

Berdasarkan contoh pada tabel 1, maka didapat:

$$P(\text{positif}) = \frac{N(\text{positif})}{N} = 2/6 = 0,333$$

$$P(\text{negatif}) = \frac{N(\text{negatif})}{N} = 2/6 = 0,333$$

$$P(\text{netral}) = \frac{N(\text{netral})}{N} = 2/6 = 0,333$$

Langkah berikutnya adalah dengan mencari nilai peluang pada setiap *term* pada dokumen label sentimen positif, negatif dan netral. Nilai peluang term sentimen positif, netral dan negatif sebagaimana ditunjukkan dalam tabel 3, 4 dan 5. Rumus yang dipergunakan menghitung nilai peluang setiap term pada setiap sentimen sebagaimana ditunjukkan pada rumus nomor 2.

$$P(\text{ajar} | \text{positif}) = \frac{\text{count}(\text{training_set, sentimen})+1}{\text{count}(\text{sentimen})+|V|} \tag{2}$$

Tabel 3. Peluang Term Sentimen Positif

No.	term	Dokumen		df	peluang Term Positif
		ks1	ks2		
1	ajar	0	1	1	0,064516129
2	baik	0	1	1	0,064516129
3	bapak	0	1	1	0,064516129
4	batin	0	0	0	0,032258065
5	cepat	0	0	0	0,032258065
6	dosen	1	0	1	0,064516129
7	jangan	0	0	0	0,032258065
8	kuliah	1	0	1	0,064516129
9	lahir	0	0	0	0,032258065
10	maaf	0	0	0	0,032258065

Tabel 4. Peluang Term Sentimen Netral

No.	term	Term Frequency		Df	peluang Term netral
		ks3	ks4		
1	ajar	0	0	0	0,033333

2	baik	0	0	0	0,033333
3	bapak	0	0	0	0,033333
4	batin	0	1	1	0,066667
5	cepat	0	0	0	0,033333
6	dosen	1	0	1	0,066667
7	jangan	0	0	0	0,033333
8	kuliah	0	0	0	0,033333
9	lahir	0	1	1	0,066667
10	maaf	0	1	1	0,066667

Tabel 5. Peluang Term Sentimen Negatif

No.	term	Term Frequency		Df	peluang Term Negatif
		ks5	ks6		
1	ajar	0	0	0	0,034483
2	baik	0	0	0	0,034483
3	bapak	0	0	0	0,034483
4	batin	0	0	0	0,034483
5	cepat	0	1	1	0,068966
6	dosen	0	0	0	0,034483
7	jangan	1	0	1	0,068966
8	kuliah	0	0	0	0,034483
9	lahir	0	0	0	0,034483
10	maaf	0	0	0	0,034483

Setelah dilakukan perhitungan *term frequency* pada masing-masing label sentimen, kemudian dilakukan proses klasifikasi dengan contoh dokumen baru. Proses klasifikasi ini dengan cara mengalikan semua variabel kelas pada masing-masing label sentimen. Contohnya data kritik dan saran berikut:

A= kuliah sering telat

Label Sentimen Positif

$$\begin{aligned} P(\text{positif} | A) &= P(\text{positif}) * P(\text{kuliah} | \text{positif}) * P(\text{sering} | \text{positif}) * P(\text{telat} | \text{positif}) \\ &= 0,333 * 0,0645 * 0,0323 * 0,0323 \\ &= 0,0000238 \end{aligned}$$

Label Sentimen Netral

$$\begin{aligned} P(\text{netral} | A) &= P(\text{netral}) * P(\text{kuliah} | \text{netral}) * P(\text{sering} | \text{netral}) * P(\text{telat} | \text{netral}) \\ &= 0,333 * 0,033 * 0,033 * 0,033 \\ &= 0,0000123 \end{aligned}$$

Label Sentimen Negatif

$$\begin{aligned} P(\text{negatif} | A) &= P(\text{negatif}) * P(\text{kuliah} | \text{negatif}) * P(\text{sering} | \text{negatif}) \\ &\quad * P(\text{telat} | \text{negatif}) \\ &= 0,333 * 0,0345 * 0,0690 * 0,0690 \\ &= 0,0000547 \end{aligned}$$

Dari hasil perhitungan dari ketiga label kelas diatas kemudian bandingkan hasil label sentimen dari masing-masing label sentimen. Nilai yang paling besar menentukan label sentimennya, seperti data A yang memiliki nilai paling tinggi pada Label sentimen negatif. Oleh karena itu A termasuk dalam sentimen negatif.

Selanjutnya dilakukan proses training menggunakan klasifikasi *Naive Bayes Classifier* dengan pembobotan *term frequency* terhadap data latih yang berjumlah 3000 data, yang memiliki komposisi masing-masing label positif, negatif, dan netral sebanyak 1000 data. proses pelatihan/training ini akan menghasilkan model klasifikasi untuk melabeli data uji yang belum memiliki label.

Dalam proses skoring, terkadang semakin banyak suatu kata dalam suatu kelas terkandung dalam suatu dokumen, maka skor tf akan menjadi lebih tinggi pada kelas tersebut, namun hal ini tidak bisa dipastikan bahwa suatu kata tersebut memiliki makna khusus yang relevan. Solusi dari masalah ini adalah dengan menggunakan TF-IDF (*Term Frequency-Invers Document Frequency*) dengan menggabungkan metode tf dan idf (*inverse document frequency*) [12]. Rumus TF-IDF bisa dilihat pada rumus 3 dan 4:

$$idf(t) = \log \frac{n_d}{df(a,t)} + 1 \tag{3}$$

$$Tf-idf(t,d) = tf(t,d) \times idf(t) \tag{4}$$

Keterangan: Nd = Jumlah dokumen

Proses training menggunakan klasifikasi *Naive Bayes Classifier* dengan pembobotan *term frequency-inverse document frequency* terhadap 3000 data latih, yang memiliki komposisi masing-masing label positif, negatif, dan netral sebanyak 1000 data. Proses pelatihan/training ini akan menghasilkan model klasifikasi untuk melabeli data uji yang belum memiliki label. Penggunaan pembobotan TF-IDF dilakukan sebagai pembandingan dengan menggunakan pembobotan TF dengan melihat akurasi masing-masing pembobotan. Hasil perhitungan tersebut disajikan pada tabel 6.

Tabel 6. Perhitungan TF-IDF

N	Term	Dokumen						df	idf	tf*idf					
		k	ks	ks3	ks	ks	k			k	ks2	ks3	k	ks5	k
0	.	1	2		4	5	6			1			4		6
1	ajar	0	1	0	0	0	0	1	1,7 7	0	1,7 7	0	0	0	0
2	baik	0	1	0	0	0	0	1	1,7 7	0	1,7 7	0	0	0	0
3	bapak	0	1	0	0	0	0	1	1,7 7	0	1,7 7	0	0	0	0
4	batin	0	0	0	1	0	0	1	1,7 7	0	0	0	1, 7 7	0	0
5	cepat	0	0	0	0	0	1	1	1,7 7	0	0	0	0	0	1, 7 7
6	dosen	1	0	1	0	0	0	2	1,4 7	1, 4 7	0	1,4 7	0	0	0
7	jangan	0	0	0	0	1	0	1	1,7 7	0	0	0	0	1,7 7	0
8	kuliah	1	0	0	0	0	0	1	1,7 7	1, 7 7	0	0	0	0	0

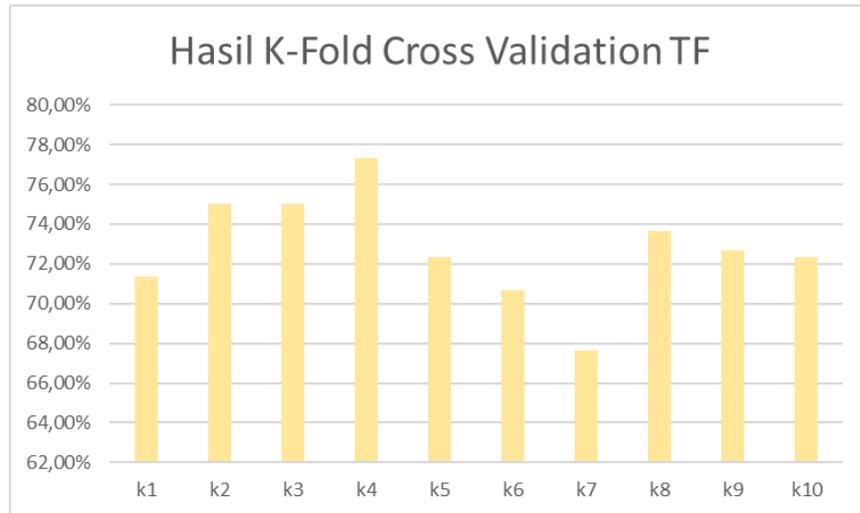
9	lahir	0	0	0	1	0	0	1	1,7 7	0	0	0	1, 7 7	0	0
10	maaf	0	0	0	1	0	0	1	1,7 7	0	0	0	1, 7 7	0	0
11	mahasiswa	0	0	1	0	0	0	1	1,7 7	0	0	1,7 7	0	0	0
12	manfaat	0	0	1	0	0	0	1	1,7 7	0	0	1,7 7	0	0	0
13	materi	0	0	0	0	0	1	1	1,7 7	0	0	0	0	0	1, 7 7
14	moga	0	0	1	0	0	0	1	1,7 7	0	0	1,7 7	0	0	0
15	mohon	0	0	0	1	0	0	1	1,7 7	0	0	1,7 7	0	0	0
16	sampai	0	0	0	0	0	1	1	1,7 7	0	0	0	0	0	1,7 7
17	sangat	1	1	0	0	0	0	2	1,4 7	1,47	1,4 7	0	0	0	0
18	sering	0	0	0	0	1	0	1	1,7 7	0	0	0	0	1,77	0
19	telat	0	0	0	0	1	0	1	1,7 7	0	0	0	0	1,77	0
20	tepat	1	0	0	0	0	0	1	1,7 7	1,77	0	0	0	0	0
21	terlalu	0	0	0	0	0	1	1	1,7 7	0	0	0	0	0	1,7 7
22	waktu	1	0	0	0	0	0	1	1,7 7	1,77	0	0	0	0	0
Nilai Bobot tiap dokumen									8,28	6,8	6,81	7,1	5,33	7,1	
										1		1		1	

Evaluasi model klasifikasi diperlukan untuk menghitung akurasi dari klasifikasi sentimen menggunakan metode *Naive Bayes Classifier*. Evaluasi model klasifikasi ini dilakukan dengan metode *k-fold cross validation* yang berfungsi menentukan data training dan testing yang baik. *Fold* yang digunakan adalah *10-fold*, yaitu membagi data latih menjadi sepuluh bagian.

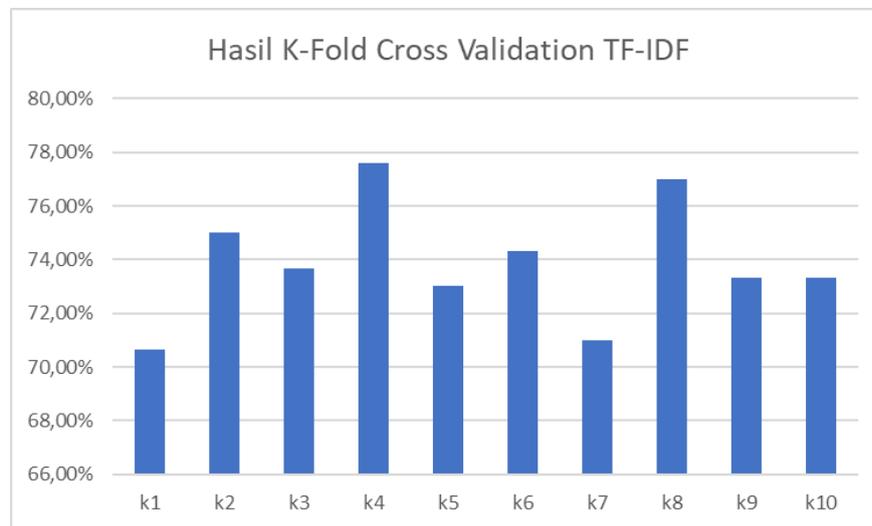
Rata-rata akurasi analisis sentimen menggunakan metode *Naive Bayes Classifier* dengan pembobotan *term frequency* (TF) pada data latih IKD sebesar 72,8%. Sedangkan Rata-rata akurasi sentimen menggunakan sentimen menggunakan metode *Naive Bayes Classifier* dengan pembobotan *term frequency-inverse document frequency* (TF-IDF) pada data latih IKD sebesar 73,9%. Hal ini menunjukkan bahwa akurasi dengan pembobotan TF-IDF lebih besar daripada dengan pembobotan TF saja.

Hasil akurasi menggunakan TF-IDF lebih besar daripada menggunakan TF dikarenakan TF-IDF memberikan hasil pembobotan dengan mempertimbangkan seberapa penting sebuah *token* atau *term* dalam kumpulan dokumen daripada TF. TF hanya menghitung kemunculan *term* pada dokumen tanpa mempertimbangkan seberapa penting sebuah *token* atau *term* dalam kumpulan dokumen.

Hasil *K-fold cross validation* dengan pembobotan TF dan TF-IDF bisa dilihat pada gambar 3 dan 4.



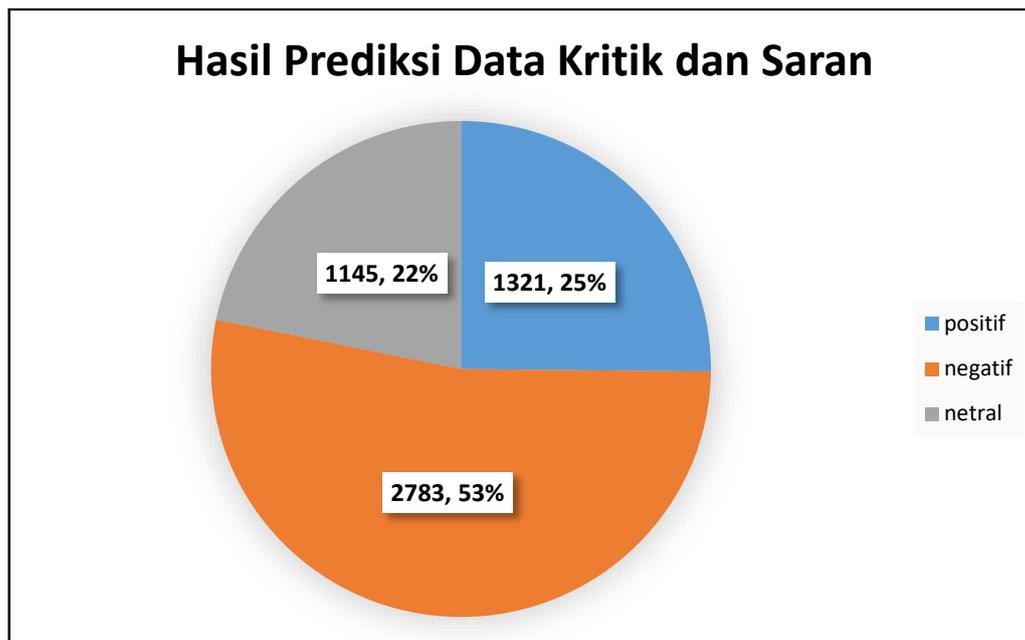
Gambar 3. Grafik Akurasi *K-fold cross validation* TF



Gambar 4. Grafik Akurasi *K-fold cross validation* TF-IDF

Tingkat akurasi yang dihasilkan sebesar 72,8% dan 73,9% ini masih bisa menyebabkan kesalahan klasifikasi. Kesalahan klasifikasi ini disebabkan oleh kalimat yang rancu atau ambigu.

Klasifikasi dilakukan secara otomatis pada data uji yang belum memiliki label dengan menggunakan *Naïve Bayes Classifier* dan pembobotan TF-IDF. Data uji berjumlah 5249. Setelah diberikan label secara otomatis, maka didapatkan hasil klasifikasi sentimen negatif sebanyak 53% dengan 2783 data, sentimen positif sebanyak 25% dengan 1321 data, dan sentimen netral sebanyak 22% sebanyak 1145 data. Dari grafik pada gambar 5 menunjukkan bahwa data kritik dan saran Fakultas Sains dan Teknologi pada 3 tahun ajaran paling banyak mendapatkan sentimen negatif.



Gambar 5. Hasil Implementasi Model Klasifikasi pada Data Uji

5. Kesimpulan dan Saran

Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa analisis sentimen dengan metode *Naïve Bayes Classifier* dapat digunakan untuk mengklasifikasi data kritik dan saran IKD dengan akurasi yang lebih besar menggunakan pembobotan TF-IDF yaitu 73,9% dibandingkan dengan pembobotan TF yaitu 72,8%. Akurasi sebesar 73,9% dan 72,8% ini merupakan rata-rata akurasi dari evaluasi model klasifikasi menggunakan *k-fold cross validation* pada 3000 data latih.

Hasil implementasi pada data uji sebanyak 5249 data kritik dan saran pada IKD Fakultas Sains dan Teknologi UIN Sunan kalijaga Yogyakarta dengan memanfaatkan model klasifikasi dari data latih menggunakan *Naïve Bayes Classifier* dan Pembobotan TF-IDF menghasilkan 53% sentimen negatif, 25% sentimen positif, dan 22% sentimen netral. Penelitian ini mempergunakan tokenisasi unigram, perlu dipertimbangkan untuk mempergunakan jenis tokenisasi Bigram dan Trigram untuk mengetahui hasil maksimal.

Referensi

- [1] A. P. M. Adji Sigit Sutedjo, "Pengaruh Kompetensi dan Motivasi Kerja terhadap Kinerja Karyawan di PT. Inti Kebun Sejahtera," *BISMA, J. Bisnis dan Manaj.*, vol. 5, no. 2, 2013.
- [2] F. A. Hermawati, *Data Mining*. Yogyakarta: ANDI, 2013.
- [3] L. Oktasari, Y. H. Chrisnanto, and R. Yuniarti, "Text Mining Dalam Analisis Sentimen Asuransi Menggunakan Metode Naive Bayes Classifier," *Pros. SNST*, vol. 7, pp. 37–42, 2016.
- [4] E. T. L. Kusriani, *Algoritma Data Mining*. Yogyakarta: ANDI, 2009.
- [5] A. Fathan Hidayatullah, A. Sn, J. Teknik, I. Fakultas, and T. Industri, "Analisis Sentimen Dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter," *Semin. Nas. Inform.*, vol. 2014, no. semnasIF, pp. 115–122, 2014.
- [6] J. LING, I. P. E. N. KENCANA, and T. B. OKA, "Analisis Sentimen Menggunakan Metode Naïve Bayes Classifier Dengan Seleksi Fitur Chi Square," *E-Jurnal Mat.*, vol. 3, no. 3, p. 92, 2014.

- [7] A. Hamzah, "Sentiment Analysis Untuk Memanfaatkan Saran Kuesioner Dalam Evaluasi Pembelajaran Dengan Menggunakan Naive Bayes Classifier (NBC)," *Snast*, vol. 3, no. November, pp. 211–216, 2014.
- [8] G. A. Buntoro, "Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter," *Integer J. Maret*, vol. 1, no. 1, pp. 32–41, 2017.
- [9] D. Gunawan, "Evaluasi Performa Pemecahan Database dengan Metode Klasifikasi Pada Data Preprocessing Data mining," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 2, no. 1, p. 10, 2017.
- [10] R. Deden and R. Nina, "Analisis Sentimen Pasar Otomotif Mobil: Tweet Twitter Menggunakan Naïve Bayes," *Simetris*, vol. 8, no. 1, pp. 113–120, 2017.
- [11] E. Prasetyo, *Data Mining Konsep dan Aplikasi Menggunakan Matlab*. ANDI Publisher, 2012.
- [12] P. H. Saputro, M. Aristian, and D. ListianingTyas, "Klasifikasi lagu Daerah Indonesia Berdasarkan lirik menggunakan metode Tf-Idf Dan Naïve Bayes," *J. Teknol. Inf. dan Terap.*, vol. 4, no. 1, pp. 47–52, 2019.