

PENERAPAN ANALISIS SENTIMEN MELALUI DATA INSTAGRAM UNTUK MENGETAHUI REPUTASI WISATA KULINER DI KOTA BANDUNG MENGUNAKAN METODE KLASIFIKASI NAÏVE BAYES

Sandy Finandra¹, Murahartawaty², Faqih Hamami³

^{1,2,3} Universitas Telkom, Bandung

sandyfinandra@student.telkomuniversity.ac.id¹, murahartawaty@telkomuniversity.ac.id²,
faqihhamami@telkomuniversity.ac.id³

Abstrak

Sosial media adalah media jejaring online yang digunakan satu sama lain untuk melakukan interaksi dan kolaborasi oleh para penggunanya. Instagram adalah suatu platform media social sebagai sarana untuk berbagi momen melalui gambar dan video. Banyaknya interaksi antara pengguna melalui komentar menyebabkan terciptanya berbagai pendapat. Komentar tersebut dijadikan sumber data untuk di olah dengan analisis sentimen untuk mengetahui reputasi wisata kuliner di Kota Bandung. komentar tersebut didapatkan melalui teknik scraping, dimana hasil scraping tersebut berupa data komentar noise yang harus melewati tahap cleaning dan preprocessing agar dataset dapat berfungsi maksimal saat dilakukan pengolahan menggunakan algoritma klasifikasi. Perbandingan algoritma klasifikasi yang dipakai adalah Naïve Bayes, Decision Tree, dan K-NN, dimana hasil akurasi tertinggi didapatkan bahwa Naïve bayes memiliki akurasi yang paling besar dengan akurasi sebesar 68.15%, *precision* 89%, *Recall* 41.60% dan *F1-Measure* 89%. Kemudian dilakukan pembobotan dengan TF-IDF untuk melihat bobot tiap kata yang terkandung dalam komentar tersebut, kemudian dilakukan penjumlahan menggunakan metode *simple additive weighting* (SAW) sebagai metode untuk pendukung keputusan untuk reputasi wisata kuliner kota bandung. Mangkok Manis, Mie Gacoan dan Lotek Alkateri menjadi tiga restoran yang memiliki bobot paling tinggi dibanding restoran lainnya.

Kata Kunci : Analisis Sentimen, Instagram, Naïve Bayes, *Simple additive Weighting*, Kuliner Kota Bandung.

Abstract

Social media is an online network media that is used by each other to interact and collaborate by its users. Instagram is a social media platform as a means to share moments through pictures and videos. The number of interactions between users through comments led to the creation of various opinions. The comments are used as a data source to be processed with sentiment analysis to determine the reputation of culinary tourism in the city of Bandung. The comments are obtained through a scraping technique, where the scraping results are in the form of noise comment data that must pass the cleaning and preprocessing stages so that the dataset can function optimally when processing using a classification algorithm. Comparison of the classification algorithms used are Nave Bayes, Decision Tree, and K-NN, where the highest accuracy results are obtained that Nave Bayes has the greatest accuracy with an accuracy of 68.15%, precision 89%, Recall 41.60% and F1-Measure 89%. Then weighted with TF-IDF to see the weight of each word contained in the comments, then summed using the simple additive weighting (SAW) method as a method for decision support for the reputation of Bandung's culinary tourism. Mangkok Manis, Mie Gacoan and Lotek Alkateri are the three restaurants that have the highest weight compared to other restaurants.

Keywords: *Sentiment Analysis, Instagram, Naïve Bayes, Simple Additive Weighting, Culinary City of Bandung*

1. Pendahuluan

Seiring berjalanya waktu, Instagram menjadi salah satu media sosial yang saat ini menjadi sebuah gaya hidup ditengah-tengah masyarakat khususnya dikalangan anak muda yang memakai sosial media

pada kehidupan sehari-hari. Instagram merupakan sebuah aplikasi foto dan video yang memungkinkan pengguna mengambil foto, menerapkan filter digital, dan membagikannya. Pengguna lain juga dapat mengomentari sebagai bentuk pengakuan atau penilaian terhadap foto atau video yang dibagikan

Analisis sentimen adalah praktik penerapan pemrosesan bahasa alami dan teknik analisis teks untuk mengidentifikasi dan mengekstrak informasi subjektif dari teks (Hussein, 2018). Untuk melakukan analisis sentimen dalam penelitian ini bisa dilihat berdasarkan komentar wisatawan Instagram, baik berupa komentar negatif, positif atau netral. Berdasarkan data kominfo pengguna internet di Indonesia mencapai 150 juta jiwa dengan penetrasi 56% yang tersebar di seluruh wilayah. Jumlah tersebut hanya selisih sedikit dengan pengguna internet *mobile* dengan persentase 53%. Data tersebut dapat dilihat pada Gambar I-1.¹ Sedangkan Berdasarkan data GNFI pengguna Instagram di Indonesia yang tercatat pada bulan Januari sampai Mei 2020 yaitu mencapai 69,2 juta pengguna. Peningkatan tersebut terjadi karena kebijakan di tengah pandemic covid-19, dimana pengguna Instagram di Indonesia didominasi oleh golongan usia produktif pada rentang 18-34 tahun. Penelitian ini dilakukan untuk pengukuran analisis sentimen keterlibatan pengguna Instagram dengan wisata kuliner Kota Bandung. Dimana data komentar Instagram sebagai *input* memungkinkan untuk memperoleh gambaran umum dari opini publik terhadap wisata kuliner Kota Bandung. Hasil dari rekomendasi dapat digunakan oleh Dinas Pariwisata Kota Bandung untuk meningkatkan kualitas wisata kuliner Kota Bandung sekaligus pendapatan daerah Kota Bandung.

2. Dasar Teori

2.1 Instagram

Instagram adalah saluran komunikasi yang relatif baru yang memungkinkan penggunaannya berbagi foto dan video. Sejak diluncurkan pada Oktober 2010, telah terjadi pertumbuhan pesat dalam jumlah penggunaannya di seluruh dunia. Instagram kini memainkan peran dominan tidak hanya sebagai hiburan dan alat interaksi tetapi juga dalam bisnis. Menurut AlGhamdi dan Reilly, 83% pemasar menempatkan nilai tinggi pada aplikasi sosial karena peran penting mereka dalam menjangkau dan mempertahankan pelanggan, dan yang terpenting dalam menciptakan peluang bisnis baru (Alkhowaiter, 2016).

2.2 Analisis Sentimen

Analisis sentimen adalah memberi label pada tubuh teks sebagai pernyataan opini positif atau negatif,

seperti dalam meringkas konten ulasan produk online. Dalam pengertian ini, analisis sentimen dapat dianggap sebagai tantangan dalam membangun pengklasifikasi dari teks (Stine, 2019). Sistem analisis sentimen diterapkan di hampir setiap bisnis dan domain sosial karena opini adalah pusat dari hampir semua aktivitas manusia dan merupakan pengaruh utama dari perilaku kita keyakinan dan persepsi tentang kenyataan, dan pilihan yang dibuat, sebagian besar dikondisikan pada bagaimana orang lain melihat dan mengevaluasi dunia. Untuk alasan ini, ketika dalam membuat keputusan, kita sering mencari pendapat orang lain. Ini berlaku tidak hanya untuk individu tetapi juga untuk organisasi (Liu, 2012).

2.3 Naive Bayes

Naive Bayes (NB) adalah salah satu algoritma data *mining* paling terkenal untuk klasifikasi. Ini menyimpulkan probabilitas contohnya baru milik beberapa kelas berdasarkan asumsi itu semua atribut tidak bergantung satu sama lain. Asumsi ini didorong oleh kebutuhan untuk memperkirakan probabilitas *multivariat* dari data pelatihan *Naive Bayes* adalah pengklasifikasi yang sangat kompeten di banyak aplikasi dunia nyata (Chen et al., 2020). Persamaan naive bayes dapat dilihat sebagai berikut.

$$P(c_j|w_i) = \frac{P(c_j) \times P(w_i|c_j)}{P(w_i)}$$

Keterangan:

$P(c_j|w_i)$: *Posterior* merupakan peluang kategori j ketika terdapat kemunculan kata i.

$P(c_j)$: *Conditional probability* merupakan peluang sebuah kata i masuk. Kedalam kategori j.

$P(w_i|c_j)$: *Prior* merupakan peluang kemunculan sebuah kategori j.

$P(w_i)$: Peluang kemunculan sebuah kata

I : Indeks kata yang dimulai dari 1 hingga kata ke-k

J : Indeks kategori yang dimulai dari 1 hingga kategori ke-n

(Gunawan et al., 2017).

2.4 Simple Additive Weighting (SAW)

Metode Simple Additive Weighting (SAW) sering juga dikenal istilah metode penjumlahan terbobot.

Konsep dasar metode SAW adalah mencari penjumlahan terbobot dari rating kinerja pada setiap alternative pada semua atribut (Darmastuti D., 2013). Proses saw secara umum dapat dilihat pada Persamaan berikut (Bukori et al., 2015).

$$V_i = \sum_{j=1}^n w_j r_{ij}$$

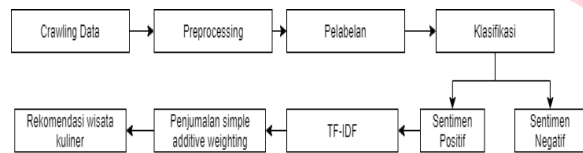
Keterangan:

V_i = ranking untuk setiap alternatif

w_j = nilai bobot dari setiap kriteria

r_{ij} = nilai rating kinerja ternormalisasi

3. Metodologi Penelitian



Gambar 3-1 Alur Kerja Penelitian

Berdasarkan Gambar 1, alur yang dipakai dalam penelitian ini adalah dengan cara, crawling data komentar pengguna instagram menggunakan teknik web scraping. selanjutnya data yang diperoleh masuk ke tahap preprocessing dimana pada tahap tersebut data akan di cleaning, case folding, tokenize, stopword removal, dan *stemming*. Kemudian data akan dilakukan pelabelan secara manual. Tahap selanjutnya data diolah menggunakan klasifikasi naïve bayes. Dan hasil dari klasifikasi tersebut akan dilakukan pembobotan dengan fitur ekstraksi TF-IDF dan dari data tersebut akan dilakukan penjumlahan menggunakan metode simple additive weighting (SAW) agar menghasilkan keputusan untuk rekomendasi.

4. Pengumpulan dan Pengolahan

4.1 Pengumpulan

Teknik yang digunakan untuk crawling data pada penelitian ini adalah menggunakan web scraping. Scraping yang dilakukan adalah menggunakan bahasa pemrograman python dan command prompt sebagai *tools* untuk melakukan *run* script python

tersebut. hasil akhir dari scraping ini ini adalah data yang memiliki dua atribut yaitu username dan komentar.

Table IV-1 Contoh Hasil Scraping

No	Username	Komentar
1	Ikhamei	Blm ada di gofood?
2	Astimaswati	@peni.sabani ada dibanjaran??
3	rikacam	Yang di batununggal kapan buka niih? Pulang pergi ngantor lewat situ belum buka ² aja.. gak sabar nicccc
4	diah.keakanze	kalo sekarang udh harga normal ya min?
5	wfebrianti_	@seni18sw @listrianinurazizah hayu ih ngiler ??
6	hanifah_juna	@hasnaamelliaa @lestarisoffy @dewipurnama_ ku Baru tau yuu ah ??
7	reniyuliani44	Hayu kesini @nenggomesh ??

Pengumpulan data yang dilakukan secara manual yaitu untuk data kiriman (Postingan) yang menjadi sumber komentar pengguna tersebut terkumpul. Dan menjadi salah satu data yang digunakan dalam penelitian ini untuk melakukan penjumlahan simple additive weighting (SAW) untuk mencari rekomendasi restoran atau menu makanan berdasarkan data komentar pengguna. Atribut yang diambil dari pengumpulan manual ini adalah nama restoran atau nama makanan, *url*, deskripsi postingan

Tabel IV- 2 Daftar Kiriman (Postingan)

NO	Nama Restoran / Makanan	URL
1	Ayam Sawce	https://www.instagram.com/p/CLqaJtOI-7S/
2	Ayam Sawce	https://www.instagram.com/p/CLDrwfQILVo/
3	Ayam Sawce	https://www.instagram.com/p/CKDUE6ylmek/
4	Ayam Sawce	https://www.instagram.com/p/CJcra5TFpGz/

5	Ayam Sawce	https://www.instagram.com/p/CIxBNL2liSw/
6	Ayam Sawce	https://www.instagram.com/p/CHE8CN6IEfw/
7	Ayam Sawce	https://www.instagram.com/p/CEi3LoolYHU/
8	Jigana Suki	https://www.instagram.com/p/CMtPJ3dhYR1/
9	Mie Ayam Bandung	https://www.instagram.com/p/CL15gBvpwA3/
10	Lomie Ciliwung	https://www.instagram.com/p/CLoME18pwZw/
11	Bubur Ayam Alkateri	https://www.instagram.com/p/CLYCPd9J8Rl/
12	Tahu Gejrot	https://www.instagram.com/p/CKySSaFJPHd/
13	Ropang Krimbat	https://www.instagram.com/p/CKiugzXlWNn/
14	Roti Bakar 234	https://www.instagram.com/p/CMo-sR5pp8j/
15	Mie Bakso Mas Eko	https://www.instagram.com/p/CMdyUfOpNSc/
16	Mie Bakso Mas Eko	https://www.instagram.com/p/CKQi6bSpl9y/
17	MIE GACOAN	https://www.instagram.com/p/CMJRUNWJyUd/
18	Sate Damri	https://www.instagram.com/p/CL_tvOxL3BR/
19	POCO LOCO	https://www.instagram.com/p/CL3web3BzVB/
20	BATAGOR	https://www.instagram.com/p/CLlb_45pSml/
21	Mangkok Manis	https://www.instagram.com/p/CLGPVRGBx7B/
22	Nasi Bakar Rambo	https://www.instagram.com/p/CK2rvtZhQUn/
23	Michilo Baksobretel	https://www.instagram.com/p/CKxo9XKAg0S/
24	Mie Mahameru	https://www.instagram.com/p/CKtPhJzhyjp/
25	Mie Mahameru	https://www.instagram.com/p/CIF6gP_pKe3/
26	Baso Rudal Anggrek	https://www.instagram.com/p/CJkyb4Hp9aF/
27	Mie Kocok Persib	https://www.instagram.com/p/CJS1O7lp1IU/
28	Baso Joko Sirod	https://www.instagram.com/p/CI1-1cLJBZD/
29	Gangnam Pancake	https://www.instagram.com/p/CIfaUzFptUe/
30	Sate Pak Mino	https://www.instagram.com/p/CIIdKU31pfN7/
31	Cimol	https://www.instagram.com/p/CIUxACvJ0xE/

32	Daleman Juarasambal!	https://www.instagram.com/p/CINmQijjzQe/
33	Lotek Alkateri	https://www.instagram.com/p/CIP3chtJH-b/
34	Awug	https://www.instagram.com/p/CHhy6DQJYAw/

4.1.1 Dataset

Data-data yang digunakan adalah kumpulan komentar pada media sosial instagram. Komentar tersebut diambil dengan menggunakan teknik web scraping dengan pemrograman python. Hasil scraping tersebut disimpan dengan format csv.

Tabel IV- 3 Jumlah Dataset Sebelum dan Setelah

Jenis Dataset	Jumlah data sebelum dipilih	Jumlah data setelah dipilih
Komentar	3729	864
Kiriman (Postingan)	34	34

4.2 Pengolahan Data

Tahap ini merupakan tahap dimana dataset komentar di olah yang terdiri dari beberapa proses diantaranya adalah data *cleansing*, preprocessing, klasifikasi, dan evaluasi performansi.

4.2.1Preprocessing Data

1. Hapus *Username*

Pada tahap ini menghapus mention yang mencantumkan username pengguna, maka mention tersebut akan dihapus, pada data komentar Tabel IV-3 ini, misalkan terdapat mentions “@username”, maka *mention* akan di hapus

Tabel IV- 4 Hapus *Username*

Sebelum	Sesudah
@intanm_15 ntanada yg lebih murah ini 😊	ntan ada yg lebih murah ini 😊
@niewidyawatileni ini si mas eko yang dulu di rivai bukan sih 😊	ini si mas eko yang dulu di rivai bukan sih 😊

2. Hapus Angka

Pada tahap ini menghapus angka pengguna yang menyertakan angka pada komentar, maka angka

tersebut akan dihapus, pada data komentar Tabel IV-4 ini, misalkan terdapat angka 15k, 2x, 5rb dan 100pcs, maka hanya angka yang akan di hapus.

Tabel IV- 5 Hapus Angka

Sebelum	Sesudah
emang mahal ini roti ngeri banget..biasa dipasaran 15k ini 2x lipatnys	emang mahal ini roti ngeri banget..biasa dipasaran k ini x lipatnys
enak tapi kotor ga pake sarung tangan plastik, padahal cuma 5rb dapat 100pcs. apa rasa enaknya dapat dari keringet ditangan	enak tapi kotor ga pake sarung tangan plastik, padahal cuma rb dapat pcs. apa rasa enaknya dapat dari keringet ditangan

3. Hapus Hashtag

Pada tahap ini menghapus *hashtag* pengguna yang menyertakan *hashtag* pada komentar, maka *hashtag* tersebut akan dihapus, pada data komentar Tabel IV-5 ini, misalkan terdapat *hashtag* “#pendapatpribadi, maka *hashtag* tersebut akan di hapus.

Tabel IV- 6 Hapus Hashtag

Sebelum	Sesudah
Rasa sama aja sama sate lain #pendapatpribadi	Rasa sama aja sama sate lain

4. Hapus Emoticons

Pada tahap ini menghapus *emoticons*, *pengguna* yang menyertakan *emoticons* pada komentar, maka *emoticons* tersebut akan dihapus, pada data komentar Tabel IV-6 ini, misalkan terdapat *emoticons* “😊”, maka *emoticons* tersebut akan di hapus.

Tabel IV- 7 Hapus Emoticons

Sebelum	Sesudah
Bakso langganan zaman masih kerja dulu , wajib cobain lagi nih 😊	Bakso langganan zaman masih kerja dulu , wajib cobain lagi nih

5. Case Folding

Tahapan *case folding* adalah tahapan mengubah semua huruf yang ada pada suatu kalimat menjadi huruf kecil semua. Huruf yang dikecilkan hanya huruf ‘a’ sampai dengan ‘z’, selain huruf tersebut maka tidak akan diterima.

Tabel IV- 8 Contoh *Case Folding*

Sebelum	Sesudah
Suka banget jadi kangen pulang ke bandung kalo dijakarta sebutnya kue dongkal	suka banget jadi kangen pulang ke bandung kalo dijakarta sebutnya kue dongkal

6. Tokenization

Tahapan *tokenization* merupakan proses memecah kalimat menjadi bagian-bagian kata seperti token. Pada proses tersebut dibutuhkan “*whitespace*” sebagai pemisah agar kalimat dapat diubah menjadi beberapa bagian kata. contohnya seperti Tabel IV-8 dimana 1 kalimat diubah menjadi 9 buah bagian kata.

Tabel IV- 9 Contoh *Tokenization*

Sebelum	Sesudah
yang aromanya suka sampe ke pintu ruangan kita ya	yang
	aromanya
	suka
	sampe
	ke
	Pintu
	ruangan
	Kita
	Ya

7. Stopword Removal

Tahapan *stopword removal* merupakan proses menghapus kata-kata yang tidak mengandung makna atau tidak perlu digunakan, misalnya : di, untuk, ke, yang, dll. *library* python yang digunakan pada *stopword* ini adalah Sastrawi.

Tabel IV- 10 Contoh *Stopword Removal*

Sebelum	Sesudah
Yang	aromannya
aromanya	suka
Suka	pintu
sampai	ruangan
Ke	
Pintu	
ruangan	
Kita	
Ya	

8. Stemming

Stemming merupakan tahap terakhir dari *preprocessing*. Pada tahap ini bertujuan untuk menghilangkan imbuhan dalam sebuah kata atau merubah kata yang memiliki imbuhan menjadi kata dasar yang bertujuan agar kalimat mudah di pahami.

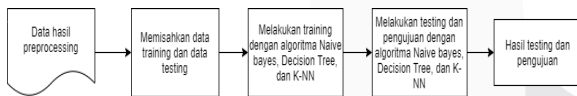
Penghapusan imbuhan ini dilakukan sama seperti pada tahap stopword removal dimana daftar kata yang digunakan menggunakan *library* Sastrawi.

Tabel IV- 10 Contoh Stemming

Sebelum	Sesudah
Aromannya	aroma
Suka	suka
Pintu	pintu
ruangan	Ruangan

4.2.2 Klasifikasi Algoritma

Setelah melalui tahap preprocessing selanjutnya akan dilakukan klasifikasi pada dataset tersebut. Klasifikasi dilakukan dengan cara labeling data komentar secara manual. Pelabelan terdiri dari tiga sentimen yaitu, positif, negatif dan netral. Komentar dikatakan positif apabila komentar mengandung makna baik atau mendukung. Komentar negatif apabila komentar mengandung makna negatif, tidak mendukung, sindiran atau semacamnya. Komentar dikatakan netral apabila tidak mengandung makna positif ataupun negatif. Untuk komentar yang berisi iklan pada kasus ini dianggap sebagai netral. Kemudian dilakukan perbandingan menggunakan tiga algoritma yaitu Naïve Bayes, Decision Tree, dan K-NN. Sampai di dapatkan hasil akurasi.apabila tidak mengandung makna positif ataupun negatif. Untuk komentar yang berisi iklan pada kasus ini dianggap sebagai netral.



Gambar IV- 3 Proses klasifikasi

Pada gambar IV-2 menampilkan tahap yang dilakukan para proses klasifikasi yaitu:

1. Data yang sudah melalui tahap preprocessing digunakan sebagai inputan.
2. Memisahkan dataset menjadi data training dan data testing dengan tiga perbandingan rasio yaitu 70:30, 75:25, dan terakhir 80:20.
3. Melakukan training dengan algoritma Naïve Bayes, Decision Tree dan algoritma K-NN.
4. Melakukan testing dan pengujian algoritma Naïve Bayes, Decision Tree, dan K-NN.

5. Menghitung akurasi berdasarkan hasil dari testing.

4.2.2.1 Data Training

Data training adalah dataset yang digunakan untuk memprediksi atau menjalankan fungsi dari sebuah algoritma dan mesin akan memberikan petunjuk melalui algoritma agar mesin tersebut dapat menemukan korelasi dengan pola yang telah diberikan.

Tabel IV- 11 Pembagian Data Training

Rasio	Jumlah Data Training
70:30	605
75:25	648
80:20	791

4.2.2.2 Data Testing

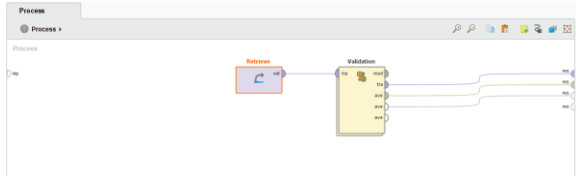
Data testing dalah dataset yang di uji untuk mengukur keberhasilan suatu performa atau keakuratan klasifikasi. Maka untuk mengetahui hal tersebut data training dan data testing harus di bedakan untuk mengetahui apakah algortima tersebut mampu mengetahui pola yang diberikan dengan soal yang berbeda. Tujuan lain dari data testing sendiri adalah untuk mengetahui performansi akurasi dari tiga algoritma yang sudah ditentukan diantaranya adalah algoritma Naïve Bayes, Decision Tree, dan K-Nearest Neighbor (K-NN). Dilakukan pembagian tiga rasio yaitu 70:30, 75:25 dan 80:20 untuk melihat akurasi terbaik yang didapatkan dari pengujian tersebut.

Tabel IV- 12 Pembagian Data Testing

Rasio	Jumlah Data Testing
70:30	259
75:25	216
80:20	173

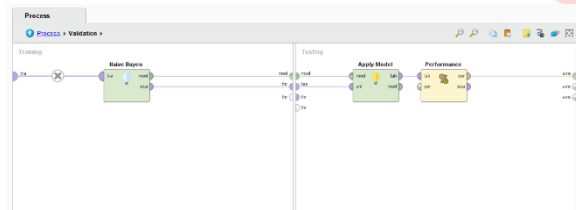
4.2.2.3 Perbandingan Algoritma

Pada tahap perbandingan ini melakukan pemodelan klasifikasi menggunakan tiga algoritma yaitu Naïve Bayes, Decision Tree, dan K-NN. Diawali dengan input berupa data komentar relevan yang berjumlah 864 dan sudah melalui tahap preprocessing. Kemudian dilakukan split validation untuk membagi dua data secara acak, sebagian data training dan sebagian lainnya untuk data testing.



Gambar IV- 4 Model Proses Pada RapidMiner

Dalam proses split validation terdapat proses training, dimana pada proses tersebut melakukan klasifikasi menggunakan tiga algoritma yaitu Naïve Bayes, Decision Tree, dan K-NN secara bergantian. Dan pada proses testing menggunakan apply model untuk menerapkan model yang telah dilatih sebelumnya menggunakan data training. Terakhir performance untuk melihat akurasi dari algoritma tersebut.



Gambar IV- 5 Model Proses Pada RapidMiner

Dalam pengujian dengan rasio yang berbeda dengan tujuan untuk mencari akurasi yang paling baik. Akurasi paling tinggi diperoleh dari algoritma naïve bayes dengan penggunaan rasio 70:30, dimana 70% dijadikan data training dan 30% dijadikan data testing. Akurasi yang didapatkan dari rasio tersebut adalah 86.87%.

Tabel IV- 13 Hasil Pengujian Rasio

		Akurasi	Data Trainin g	Data Testin g	Tota l
Naïve Bayes	70:30	86.87 %	605	259	864
	75:25	86.11 %	648	216	864
	80:20	86.13 %	691	73	864
Decision Tree	80:20	84.56 %	605	259	864
	70:30	84.26 %	648	216	864
	75:25	84.39 %	691	173	864

K-NN					
80:20	80:20	25.48 %	605	259	864
	70:30	25.00 %	648	216	864
75:25	75:25	25.00 %	691	173	864
	80:20	86.87 %	605	259	864

Berdasarkan tabel diatas terbukti bahwa naïve bayes menjadi algoritma yang memiliki tingkat akurasi paling baik dibandingkan algoritma lainnya. Merupakan perbandingan sentimen positif dan negatif dari data, ditampilkan dalam bentuk grafik bar agar mempermudah dalam melihat perbandingan antara sentimen. Hasil tersebut menunjukkan bahwa sentimen positif tersebut berjumlah 730 dan sentimen negatfi berjumlah 134 denga total data yang dipakai pada sentimen ini adalah 864.

5. Analisis Hasil

5.1 Evaluasi Performansi

Hasil dari klasifikasi pelabelan sentimen komentar yang dilakukan secara manual dengan klasifikasi menggunakan algoritma *Naïve Bayes*, *decision tree* dan *k-nn*. Pada tahap evaluasi *performance* akan dihitung nilai akurasi, *precision*, *recall*, dan *f1-measure* yang didapat dari *confusion matrix*.

Tabel 5- 1 Confision Matrix

	Positif	Negatif	Total
Positif	219	253	219
Negatif	0	6	0
Total	219	259	219

Berdasarkan Tabel 5-1, dapat diketahui bahwa:

1. Komentar yang mengandung sentimen positif dan diprediksi secara benar oleh algoritma Naïve Bayes sebanyak 219 (*True positive*) sedangkan jumlah komentar yang salah di prediksi 0 (*False Positive*).
2. Komentar yang mengandung sentimen negatif dan diprediksi secara benar oleh algoritma naïve bayes sebanyak 6 (*True Negative*). Sedangkan jumlah komentar sentimen negatif yang diprediksi secara salah sebanyak 34 (*False Negative*).

Dengan demikian dapat diketahui hasil dari precision, recall dan f1-measure, berikut merupakan hasil perhitungan precision, recall, dan f1-measure dapat dilihat pada tabel 5-2

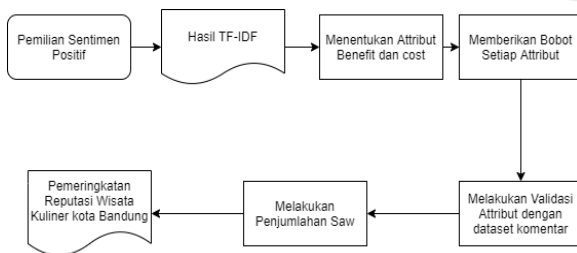
Tabel 5- 2 Confusion Matrix (2)

Label	Precision	Recall	F1-measure
Positif	86.56%	100.00%	92.79%
Negatif	100%	15.00%	26.08%
Rata-Rata	93%	57.50%	75.39%

Berdasarkan Tabel V-3 hasil evaluasi performansi yang dilakukan dengan tools RapidMiner, didapatkan hasil Precision 93%, Recall 57.50%, sehingga menghasilkan F1-Measure sebesar 75.39%. Suatu classifier dikatakan bagus jika memiliki precision dan recall yang tinggi. Namun classifier dapat dikatakan efektif jika nilai precision lebih tinggi dari recall, atau sebaliknya recall lebih tinggi dibandingkan precision dan terakhir precision, recall memiliki nilai yang sama. Hasil recall positif dan negatif menunjukkan hasil lebih kecil jika dibandingkan dengan precision.

5.2 Metode Simple Additive Weighting (SAW)

Pada tahap ini diawali dengan pemilihan sentimen positif kemudian dilakukan TF-IDF. Setelah menjalankan proses tersebut dan mendapatkan hasil TF-IDF maka selanjutnya akan menentukan atribut benefit dan cost, atribut benefit ditentukan berdasarkan kategori *food quality*. Data tersebut dapat dilihat pada Tabel 5-3 dan Tabel 5-4. Pada kategori tersebut dibagikan bobot perkategori hingga bobot tersebut mencapai nilai maksimal satu. Data tersebut dapat dilihat pada Tabel 5-5. Atribut pada kategori yang telah ditetapkan kemudian dicocokkan dengan dataset komentar untuk pembobotan restoran. Data tersebut dapat dilihat pada Tabel V-6. Kemudian dilakukan penjumlahan SAW yang menghasilkan pemeringkatan wisata kuliner di Kota Bandung.



Gambar 2 Gambar Proses Implementasi SAW

1. Atribut Benefit

Tabel 5- 3 Hasil Evaluasi Performansi

No	Kategori <i>Food Quality</i>	Atribut Benefit	Bobot TF-IDF
1	Rasa	Enak	0.994
2	Porsi Makanan	Porsi	0.696
3	Aroma	Rempah	0.396
4	Tingkat Kematangan	Mantap	0.998
5	Bentuk Penyajian	Saji	0.999
6	Temperatur	Hangat	0.651
7	Tekstur Makanan	Kenyang	0.928
8	Warna	Warna	0.549

Tingkat kematangan pada kategori *food quality* diwakili kata mantap karena memungkinkan makanan tersebut disebut mantap karena berdasarkan tingkat kematangan yang pas, makanan tersebut menjadi lebih enak. Kemudian kategori tekstur makanan diwakili kata kenyang karena dari tekstur makanan yang bagus akan membuat rasa kenyang lebih cepat.

2. Atribut Cost

Tabel 5- 4 Hasil Evaluasi Performansi

No	Kategori	Atribut Cost	Bobot TF-IDF
1	Harga	Harga	0.717

3. Kriteria Pembobotan

Tabel 5- 5 Hasil Evaluasi Performansi

No	Kategori <i>Food Quality</i>	Atribut Benefit	Pembobotan
1	Rasa	Enak	0.2
2	Porsi Makanan	Porsi	0.1
3	Aroma	Rempah	0.1
4	Tingkat Kematangan	Mantap	0.1
5	Bentuk Penyajian	Saji	0.05
6	Temperatur	Hangat	0.1

7	Tekstur Makanan	Kenyang	0.1
8	Warna	Warna	0.05
9	Harga	Harga	0,2
Jumlah Total			1

Jumlah total pembobotan pada metode saw ini harus mencapai 1 dan untuk kategori food quality rasa dan harga itu diberi bobo 0.2 karena dua kategori tersebut menjadi salah satu kategori paling penting dalam sebuah makanan. Contoh dengan rasa yang enak dan harga yang murah maka menjadi salah satu keunikan restoran bagi pengunjung

5.2.1 Pembobotan Restoran

Daftar kiriman ini akan digunakan sebagai data utama untuk melakukan pendukung keputusan dalam menggunakan metode simple additive weighting. Dimana kiriman tersebut akan validasi berdasarkan kriteria attribut yang sudah ditentukan sebelumnya. Maka kiriman akan diberikan bobot jika kata pada dataset komentar tersebut mengandung kata attribute benefit dan cost yang bersentimen positif.

5.2.2 Penjumlahan Simple Additive Weighting

Tahap ini merupakan tahap terakhir dari metode SAW. Dimana hasil akhir diperoleh dari penjumlahan perkalian dengan vector bobot sehingga diperoleh nilai terbesar sebagai alternatif untuk menentukan reputasi wisata kuliner di Kota Bandung. Penjumlahan perkalian tersebut dapat menggunakan perususan sebagai berikut:

$$Vi = \sum n wj rij$$

Keterangan:

- Vi= ranking untuk setiap alternatif
- wj= nilai bobot dari setiap kriteria
- rij= nilai rating kinerja ternormalisasi

Maka bisa dipastikan hasil dari penjumlahan dari metode simple additive weighting adalah sebagai berikut:

Tabel 5- 5 Hasil Penjumlahan SAW

NO	Nama Restoran / Menu	Hasil Penjumlahan SAW
1	Mangkok Manis	0.669
2	Mie Gacoan	0.604
3	Lotek Alkateri	0.561
4	Tahu Gejrot	0.539
5	Mie Kocok Persib	0.511
6	Roti Bakar 234	0.481

7	Lomie Ciliwung	0.442
8	Poco Loco	0.442
9	Mie Ayam Bandung	0.439
10	Mie Baso Mas Eko	0.433
11	Cimol	0.413
12	Baso Joko Sirod	0.412
13	Batagor	0.391
14	Awug	0.391
15	Mie Mahameru	0.338
16	Ayam Sawce	0.298
17	Ropang Krimbat	0.298
18	Sate Damri	0.298
19	Michilo Baksobertel	0.298
20	Daleman Juara Sambal	0.298
21	Bubur Ayam Alkateri	0.268
22	Baso Rudal Anggrek	0.226
23	Nasi Bakar Rambo	0.198
24	Gangnam Pancake	0.198
25	Sate Pak Mino	0.198
26	Jigana Suki	0

Pada Tabel 5-5 yang merupakan hasil dari penjumlahan simple additive weighting (SAW). Dari 26 restoran, hanya 5 restoran yang memiliki bobot perhitungan diatas 0.5. Restoran tersebut yaitu Mangkok Manis, Mie Gacoan, Lotek Alkateri, Tahu Gejrot, dan Mie Kocok Persib. Hasil tersebut didapat berdasarkan perhitungan melalui analisis sentimen pengguna Instagram terkait wisata kuliner di Kota Bandung dan mungkin dapat menjadi rekomendasi wisata kuliner untuk masyarakat domestik ataupun luar domestik. Sedangkan untuk restoran lainya yang memiliki bobot kurang dari 0.5 bukan restoran tidak menjadi rekomendasi melainkan disebabkan restoran tersebut mendapat komentar dari pengguna instagram tetapi tidak memiliki penilaian attribut.

6. Kesimpulan dan Saran

6.1 Kesimpulan

1. Menggunakan dengan total 864 data komentar pengguna. Dengan percobaan tiga rasio yaitu 70:30, 75:25 dan 80:20 dengan pengujian tiga metode algoritma sebagai pembanding menghasilkan kesimpulan bahwa algoritma naïve bayes memiliki tingkat akurasi yang paling tinggi dibandingkan dengan algoritma decision tree dan K-NN. Akurasi yang didapat sebesar 86.87% yang di uji menggunakan tools RapidMiner. Kemudian mendapatkan hasil evaluasi performansi yang cukup memuaskan yaitu *Precision* 93%, *Recall* 57.50%, sehingga menghasilkan *F1-Measure* sebesar 75.39%.

Hasil tersebut didapatkan dengan rasio perbandingan 70:30.

6.2 Saran

Diharapkan juga penelitian berikutnya menggali cara merubah kata-kata tidak baku, kata-kata gaul, kata singkat, bahasa daerah menjadi kata baku bahasa indonesia. Selain itu dapat menggali cara untuk melakukan *image processing* apabila ingin melakukan analisis sentimen terhadap sosial media Instagram.

REFERENSI

- Alkhowaiter, W. (2016). The power of instagram in building small businesses. *Social Media. Social Media: The Good, the Bad, and the Ugly Lecture Notes in Computer Science*.
- Bukori, I., Pujiono, P., & Suharnawi, S. (2015). Metode Simple Additive Weighting (Saw) Untuk Penentuan Peringkat Dalam Pembuatan Peta Tematik Daerah Rawan Demam Berdarah Dengue (Studi Kasus Kabupaten Pati). *Techno. Com, 14(4)*, 272–280.
- Chen, S., Webb, G. I., Liu, L., & Ma, X. (2020). A novel selective naïve Bayes algorithm. *Knowledge-Based Systems*.
- Gunawan, F., Fauzi, M. A., & Adikara, P. P. (2017). Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan Naive Bayes dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA Mobile). *Systemic: Information System and Informatics Journal, 3(2)*, 1–6.
- Hussein, D. M. E. D. M. (2018). A survey on sentiment analysis challenges. *Journal of King Saud University - Engineering Sciences*.
- Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies, 5(1)*, 1–167.
- Stine, R. A. (2019). Sentiment analysis. *Annual Review of Statistics and Its Application*.