

Klasifikasi Sentimen Ulasan Tempat Makan Berbahasa Indonesia Menggunakan *Lexicon* dan *Improved Naive Bayes*

Agi Maulana ^{#1}, Yuliant Sibaroni ^{*2}

School of Computing, Telkom University

Jl. Telekomunikasi Terusan Buah Batu, Bandung, West Java 40257, Indonesia

¹ agimaulanaxyz@student.telkomuniversity.ac.id

² yuliant@telkomuniversity.ac.id

Abstract

Restaurant reviews on online sites often provide reviews with scores that do not match the sentiment of the review. A review could have a low score but the review contain positive meanings and a high-score review could have negative meanings. Various attempts to classify sentiment using sentiment analysis have been carried out in many studies. However, sentiment analysis by relying solely on the supervised learning approach gives the results of one class tends to appear more frequently which mean this decreases the performance of the classifier. In this paper, the Improved Naive Bayes approach, Naive Bayes with unigram and bigram features combined with the approach using lexicon is proposed to improve the performance of classifiers. The feature is obtained by extracting the POS TAG pattern that contains words or phrases that express emotions that are relevant to the restaurant review. The Lexicon is built manually by collecting unigram and bigram words and phrases that show relevant emotions expressed in the restaurant review. Classifiers using Improved Naive Bayes show better performance than classifiers using Naive Bayes. Improved Naive Bayes evaluation obtains precision 80%, recall 77%, and F1 76%. While Naive Bayes evaluation obtains precision 68%, recall of 60%, and F1 56%.

Keywords: bigram, improved naive bayes, lexicon, naive bayes, sentiment analysis, unigram

Abstrak

Ulasan tempat makan pada situs daring seringkali memberikan skor yang tidak sesuai dengan makna pada ulasan. Ulasan dapat berskor rendah namun ulasan mengandung makna positif dan ulasan berskor tinggi dapat mengandung makna yang negatif. Berbagai upaya klasifikasi sentimen ulasan dengan menggunakan analisis sentimen telah dilakukan pada banyak penelitian. Namun analisis sentimen dengan hanya mengandalkan pendekatan *supervised learning* memberikan hasil salah satu kelas cenderung lebih sering muncul sehingga berakibat pada menurunnya kinerja pengklasifikasi. Dalam makalah ini, pendekatan *Improved Naive Bayes* yaitu Naive Bayes dengan fitur *unigram* dan *bigram* dipadukan dengan pendekatan menggunakan *lexicon* diusulkan untuk meningkatkan kinerja pengklasifikasi. Fitur diperoleh dengan mengekstrak pola *POS TAG* yang mengandung kata atau frasa yang mengekspresikan emosi yang relevan dengan ulasan tempat makan. *Lexicon* dibangun secara manual dengan mengumpulkan kata dan frasa *unigram* dan *bigram* yang menunjukkan emosi yang relevan diungkapkan pada ulasan tempat makan. Pengklasifikasi dengan menggunakan *Improved Naive Bayes* menunjukkan kinerja yang lebih baik dibandingkan pengklasifikasi menggunakan *Naive Bayes*. *Improved Naive Bayes* memperoleh rata – rata skor *precision* 74%, *recall* 68%, dan *F1* 65%. Sedangkan *Naive Bayes* memperoleh skor *precision* 69%, *recall* 60%, dan *F1* 52%.

Kata Kunci: analisis sentimen, *bigram*, *improved naive bayes*, *lexicon*, *naive bayes*, *unigram*

I. PENDAHULUAN

INTERNET dan teknologi informasi memberikan kesempatan bagi setiap orang untuk berbagi pendapat mengenai sesuatu pada platform daring. Beberapa situs muncul untuk memberikan kemudahan bagi konsumen untuk memberikan ulasan atau penilaian terhadap suatu entitas. Ulasan daring menjadi hal baru dalam *word-of-mouth* dan memberikan pengaruh terhadap keputusan pembelian seseorang [1]. Ulasan pada situs daring seringkali antara makna dan penilaian tidak relevan, pada ulasan dengan rentang penilaian 1 sampai 5 ulasan bernilai 1 dapat mengandung makna yang positif dan ulasan bernilai 5 dapat mengandung makna yang negatif.

Untuk memudahkan mengetahui kandungan emosi pada ulasan diperlukan klasifikasi sentimen menggunakan analisis sentimen untuk memperoleh informasi sentimen dari suatu kalimat opini [2]. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek, apakah cenderung beropini positif atau negatif. Ada beberapa pendekatan yang kerap digunakan dalam analisis sentimen, namun dibedakan menjadi dua kelompok, *supervised learning* dan *unsupervised learning*. *Naive Bayes* merupakan pendekatan supervised learning yang dapat mengklasifikasikan data dengan cepat tanpa perlu dataset yang besar dan efektif untuk mengklasifikasikan data teks. *Improved Naive Bayes* memadukan rasio *sentiment lexicon* dan pendekatan *Naive Bayes* dengan fitur unigram dan bigram. *Lexicon* atau *sentiment lexicon* merupakan kata atau frasa yang memiliki polaritas yang menunjukkan orientasi sentimen negatif atau sentimen positif [3]. *Sentiment lexicon* memiliki sensitivitas yang tinggi untuk digunakan dalam analisis sentimen [4].

Pada makalah ini akan membandingkan kinerja analisis sentimen dengan pendekatan *Naive Bayes* dan *Improved Naive Bayes* dalam mengklasifikasikan sentimen ulasan tempat makan. Evaluasi didasarkan pada perolehan skor *precision*, *recall*, dan *F1*.

II. KAJIAN PUSTAKA

Pada penelitian sebelumnya Samal, Behera, dan Panda [5] melakukan analisis kinerja terhadap beberapa algoritma supervised learning pada penelitian analisis sentimen. Secara umum, terdapat 2 tahap proses pada penelitian ini yaitu data *cleansing* dan klasifikasi menjadi positif dan negatif. Pada penelitian ini dapat disimpulkan semakin besar data yang digunakan untuk melatih model maka model akan semakin baik dalam melakukan klasifikasi data.

Penelitian analisis sentimen terhadap tempat makan berbahasa Kantonis dilakukan oleh Ziqiong Zhang, Ye, Zii Zhang dan Li [6] dengan algoritma *Naive Bayes* dan *SVM* serta menggunakan fitur *n-gram* karakter Kantonis. Penelitian memperlihatkan algoritma *Naive Bayes* memperoleh kinerja yang lebih baik dibandingkan *SVM* dan fitur *bigram* memberikan hasil yang lebih baik dibandingkan *unigram* dan *trigram*.

Penelitian analisis sentimen untuk ulasan tempat makan berbahasa Inggris dengan *sentiment lexicon* dilakukan oleh Kang, Yoo, dan Han [7]. Pada penelitian tersebut *sentiment lexicon* umum tidak dapat digunakan secara efektif untuk ulasan tempat makan karena terdapat beberapa kata yang mempengaruhi penilaian terhadap tempat makan akan tetapi dianggap memiliki polaritas yang netral atau kata yang memiliki polaritas negatif termasuk ke dalam polaritas positif atau sebaliknya. *Sentiment lexicon* yang efektif dan efisien untuk penilaian tempat makan dibangun secara manual dengan memperhatikan *unigram* dan *bigram* yang memenuhi kriteria. Kriteria yang harus dipenuhi *sentiment lexicon* antara lain pola *POS Tag* dan mengekspresikan emosi yang relevan pada ulasan tempat makan.

Penelitian ekstraksi *sentiment lexicon* dilakukan oleh Aye dan Aung [8], penelitian difokuskan pada ulasan tempat makan berbahasa Myanmar. Pada penelitian ini *sentiment lexicon* dibangun secara manual dan diekstrak dari ulasan tempat makan yang ditambang dari media sosial. *Sentiment lexicon* dibangun dengan mencari kata antonim dan kata sinonim, diklasifikasikan dalam bentuk numerik netral = 0, negative = -1, dan positif = 1. Suatu kata pada bahasa Myanmar dapat terdiri dari satu atau lebih suku kata. Pendekatan heuristik berbasis

aturan dilakukan untuk mengekstrak suku kata pada bahasa atau teks Myanmar kemudian suku kata yang telah diekstrak dibangun menjadi suatu kata dengan mengacu pada kamus. Kelas kata yang digunakan dalam POS tagging antara lain adjective, verb, adverb, noun dan emoticons.

III. METODE PENELITIAN

A. Dataset

Data yang dikumpulkan berasal dari situs Zomato dan *Google Place Review*. Zomato merupakan situs yang memfasilitasi ulasan tempat makan dan *Google Place Review* menyediakan informasi berbagai tempat beserta ulasan pengguna, dalam mengumpulkan ulasan dari *Google Place Review* menggunakan kata kunci berupa nama restoran atau kafe. Ulasan – ulasan tersebut juga memiliki penilaian bintang, ulasan dengan penilaian bintang 1 dan 2 merupakan ulasan dengan penilaian negatif sedangkan ulasan dengan penilaian bintang 4 dan 5 merupakan ulasan dengan penilaian positif. Untuk memudahkan pelabelan, ulasan yang digunakan akan menggunakan ulasan bernilai 1, 2, 4, dan 5 bintang, ulasan dengan bintang 1 dan 2 akan diberi label negatif, sedangkan ulasan bintang 4 dan 5 akan diberi label positif. Namun, seringkali ulasan dan penilaian bintang tidak konsisten, ulasan berbintang 1 dan 2 yang seharusnya mengandung emosi negatif tetapi ulasan yang diberikan mengandung emosi positif serta ulasan berbintang 4 dan 5 yang seharusnya mengandung emosi positif tetapi ulasan yang diberikan mengandung emosi negatif, maka dari itu akan dilakukan peninjauan kembali terhadap ulasan – ulasan tersebut agar label yang diberikan sesuai dengan sentimen yang terkandung.

Ulasan yang dikumpulkan berjumlah 1.006 ulasan dengan 553 ulasan positif dan 453 ulasan negatif. Dilakukan pengacakan *dataset* sebanyak 10 kali dengan ketentuan setiap pengacakan menggunakan 80% untuk data latih dan 20% untuk data uji seperti ditunjukkan oleh Tabel I.

TABEL I
PENGACAKAN *DATASET* SEBANYAK 10 KALI

<i>Dataset</i>	Data Latih		Data Uji	
	Positif	Negatif	Positif	Negatif
<i>Dataset 1</i>	444	360	109	93
<i>Dataset 2</i>	442	362	111	91
<i>Dataset 3</i>	442	362	111	91
<i>Dataset 4</i>	439	365	114	88
<i>Dataset 5</i>	441	363	112	90
<i>Dataset 6</i>	453	351	100	102
<i>Dataset 7</i>	445	359	108	94
<i>Dataset 8</i>	442	362	111	91
<i>Dataset 9</i>	450	354	103	99
<i>Dataset 10</i>	445	359	108	94

B. Sentiment Lexicon

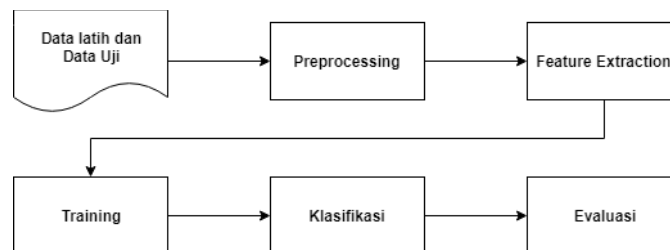
Sentiment lexicon umum kurang efektif jika digunakan sebagai fitur pada ulasan tempat makan karena terdapat *lexicon* yang tidak relevan pada ulasan tempat makan. *Sentiment lexicon* ditentukan secara manual dengan mengidentifikasi kata dan frasa yang mengekspresikan emosi terhadap rasa makanan, fasilitas, harga, layanan dan tempat. *Sentiment lexicon* harus memenuhi kriteria *type*, *target*, *pattern*, dan *polarity*. *Type* merujuk pada jenis kata atau frasa yang dibuat, *unigram* atau *bigram*. *Unigram* terdiri dari satu kata seperti “enak”,

“nyaman” dan “murah”. *Bigram* terdiri dari dua kata seperti “sangat enak” dan “suasana nyaman”. *Target* adalah subjek yang mengekspresikan emosi terhadap rasa makanan, fasilitas, harga, layanan atau tempat. *Pattern* berisi kelas kata dari kata atau frasa *sentiment lexicon*, kelas kata yang digunakan antara lain *adjectives*, *adverbs*, *verbs*, *nouns* dan *negations*. Hanya kelas kata *adjectives* dan *verbs* yang menunjukkan emosi yang relevan terhadap ulasan tempat makan akan diekstrak seperti kata “enak” atau “memuaskan”. Kata dengan kelas *adverbs*, *verbs* dan *nouns* yang diikuti kelas *adjectives* seringkali mengekspresikan emosi, seperti “sangat enak”, “terawat baik”, “tempat nyaman” akan diekstrak sebagai *sentiment lexicon*. Karena terdapat beberapa kata dengan kelas *verbs* dapat mengekspresikan emosi maka kelas kata *adverbs* dan *nouns* yang diikuti oleh kelas kata *verbs* yang mengekspresikan emosi yang relevan terhadap ulasan tempat makan juga diekstrak sebagai *sentiment lexicon*, seperti “sangat memuaskan” atau “layanan memuaskan”. Kelas kata *adjectives* dan *verbs* yang mengekspresikan emosi yang relevan dan diikuti oleh *adverbs* seperti kata “memuaskan sekali” diekstrak sebagai *sentiment lexicon*. Kelas kata *negations* yang diikuti oleh kelas kata *adjectives* dan *verbs* seperti kata “tidak nyaman” yang mengekspresikan emosi diekstrak sebagai *sentiment lexicon*. Tabel II merupakan contoh *sentiment lexicon* yang diekstrak, *sentiment lexicon* terdiri dari pola kelas kata dimana *lexicon unigram* terdiri dari satu kelas kata *adjectives* (JJ), atau *verbs* (VB) dan *lexicon bigram* terdiri dari dua kelas kata, *adverbs* (RB), *verbs* (VB), dan *nouns* (NN) diikuti oleh *adjectives* (JJ), *adverbs* (RB) dan *nouns* (NN) diikuti oleh *verbs* (VB) yang mengekspresikan emosi yang relevan.

TABEL II
 CONTOH SENTIMENT LEXICON UNTUK ULASAN TEMPAT MAKAN

Type	Target	Pattern	Polarity
unigram	Rasa makanan	enak_JJ	Positif
unigram	Pelayanan	kecewa_JJ	Negatif
bigram	Harga	sangat_RB mahal_JJ	Negatif
bigram	Tempat	tidak_NEG nyaman_JJ	Positif

C. System Overview



Gambar 1. System overview

Pada gambaran perancangan sistem terdapat dua proses utama, yaitu *preprocessing* dan *feature extraction*. Pada *preprocessing* data diolah agar menghasilkan data yang berkualitas untuk memudahkan ekstraksi fitur. *Feature extraction* menambang fitur berupa *unigram* dan *bigram* yang mengekspresikan emosi yang relevan untuk ulasan tempat makan. Pada tahap *training* dilakukan perhitungan probabilitas setiap fitur berdasarkan frekuensi fitur terhadap suatu kelas dan probabilitas masing - masing kelas sentimen. Klasifikasi merupakan proses untuk menentukan kelas sentimen ulasan tempat makan menggunakan model yang sudah dibuat pada tahap *training*, setelah melakukan klasifikasi dilakukan evaluasi kinerja dengan menghitung *precision*, *recall*, dan *F1*.

D. Preprocessing

Preprocessing dilakukan sebelum tahap *feature extraction*. *Preprocessing* bertujuan untuk memudahkan *feature extraction* dan meningkatkan kinerja klasifikasi sentimen. Pada tahapan ini teks ulasan diubah menjadi *lower-case*, angka dan simbol yang tidak relevan dihilangkan dari teks, kemudian teks ditokenisasi untuk memudahkan proses pada tahapan selanjutnya. Ulasan yang telah ditokenisasi akan dicari padanan baku pada kamus untuk setiap kata pada ulasan yang tidak dituliskan dalam bentuk baku, kamus menggunakan data *slangword* dan singkatan yang telah dibangun pada penelitian sebelumnya [9], ulasan seperti “Pelayanan bgs dan murh. Tp suasana berisik” yang telah melalui proses sebelumnya akan dirubah menjadi “pelayanan bagus dan murah tapi suasana berisik”. Pada ulasan berbahasa Indonesia seringkali konsumen memasukan penilaian atau berekspresi dengan bahasa asing, oleh sebab itu kata asing seperti “*cozy*”, “*pricey*” dan “*noisy*” serta bentuk kata tidak baku lainnya yang relevan pada ulasan tempat makan dimasukkan ke dalam kamus untuk kemudian diubah menjadi bahasa Indonesia yang sepadan. Setelah kata – kata pada ulasan menjadi bentuk baku, kemudian dilakukan pelabelan kelas kata atau *POS Tagging*, ulasan “pelayanan bagus dan murah tapi suasana berisik” setelah melalui proses *POS Tagging* menjadi “[('pelayanan', 'NN'), ('bagus', 'JJ'), ('dan', 'CC'), ('murah', 'JJ'), ('tapi', 'CC'), ('suasana', 'NN'), ('berisik', 'VB')]”. Data label kelas kata menggunakan data penelitian yang sudah dilakukan sebelumnya [10]. Kata – kata umum yang dianggap tidak relevan untuk ulasan tempat makan kemudian dihilangkan dari ulasan, setiap kata dicocokkan pada kamus, jika terdapat pada kamus kata tersebut dihapus dari ulasan, kamus menggunakan data kamus yang telah dibangun [11], ulasan yang telah melalui pelabelan kelas kata “[('pelayanan', 'NN'), ('bagus', 'JJ'), ('dan', 'CC'), ('murah', 'JJ'), ('tapi', 'CC'), ('suasana', 'NN'), ('berisik', 'VB')]” kemudian menjadi “[('pelayanan', 'NN'), ('bagus', 'JJ'), ('murah', 'JJ'), ('suasana', 'NN'), ('berisik', 'VB')]”.

E. Feature Extraction

Pada proses *feature extraction*, sistem mendeteksi pola kelas kata pada ulasan yang telah melalui tahap *preprocessing* untuk menambang fitur yang akan digunakan pada fase *modeling* dan klasifikasi. Pola – pola tersebut antara lain untuk fitur *unigram* kelas kata *adjectives* (JJ), atau *verbs* (VB) dan *lexicon bigram* terdiri dari dua kelas kata, *adverbs* (RB), *verbs* (VB), dan *nouns* (NN) diikuti oleh *adjectives* (JJ), *adverbs* (RB) dan *nouns* (NN) diikuti oleh *verbs* (VB). Pada proses ini fitur dicocokkan pada kamus *sentiment lexicon*, jika terdapat pada kamus maka fitur akan diekstrak. Jika kandidat fitur bigram tidak terdapat pada kamus maka kata *adjectives* (JJ) atau *verbs* (VB) pada bigram diperiksa pada kamus *sentiment lexicon*, jika merupakan *sentiment lexicon* maka kata *adjectives* (JJ) atau *verbs* (VB) tersebut diekstrak sebagai fitur *unigram*.

F. Training

Pada tahap training dilakukan pembuatan model atau *modeling* dengan menghitung probabilitas untuk setiap kelas sentimen dan fitur yang telah diekstrak berdasarkan frekuensi kemunculan fitur tersebut terhadap masing – masing kelas sentimen. Pada Tabel III terlihat “pelayanan bagus” yang seharusnya memiliki probabilitas pada kelas positif lebih tinggi dibandingkan pada kelas negatif namun terjadi sebaliknya, hal tersebut dapat terjadi karena fitur “pelayanan bagus” lebih sering muncul pada ulasan negatif.

TABEL III
PROBABILITAS FITUR TERHADAP MASING-MASING KELAS SENTIMEN

Fitur	P(fi positif)	P(fi negatif)
pelayanan bagus	0.4	0.6
murah	0.6	0.4
berisik	0.4	0.6

G. Klasifikasi

Penentuan kelas sentimen berdasarkan nilai probabilitas yang lebih besar dibandingkan probabilitas kelas lainnya. Ulasan diklasifikasikan ke dalam kelas positif jika nilai probabilitas kelas positif lebih besar dibandingkan nilai probabilitas kelas negatif dan ulasan diklasifikasikan ke dalam kelas negatif jika nilai probabilitas kelas negatif lebih besar dibandingkan nilai probabilitas kelas positif.

Klasifikasi dilakukan menggunakan fitur yang sudah diekstraksi dan model yang sudah dibuat pada fase *training*. Pada algoritma *Naive Bayes* klasifikasi tidak melibatkan rasio *lexicon* seperti pada formula (1)

$$Class(d_i) = arg\ max\ P(c_j) \prod_{i=1}^d P(p_i|c_j) \quad (1)$$

Class(d_i) merupakan fungsi untuk menentukan kelas sentimen (positif atau negatif) dokumen (d_i). $P(c_j)$ melakukan perhitungan probabilitas kelas (c_j). $P(p_i|c_j)$ menghitung probabilitas fitur *unigram* dan *bigram* terhadap kelas (c_j). Pada *Improved Naive Bayes* klasifikasi melibatkan rasio *lexicon* (R_1) guna meningkatkan kinerja klasifikasi seperti pada formula (2)

$$Class(d_i) = arg\ max\ R_1(p_{ij})P(c_j) \prod_{i=1}^d P(p_i|c_j) \quad (2)$$

Keterangan formula (2) sama dengan keterangan pada formula (1). Hanya saja pada *Improved Naive Bayes* melibatkan rasio *lexicon* (R_1). R_1 ditunjukkan seperti pada formula (3)

$$R_1(p_{ij}) = \frac{\sum_{p_{ij} \in L_j} C(p_{ij})}{\sum_{p_{ij} \in L} C(p_{ij})} \quad (3)$$

Keterangan :

$R_1(p_{ij})$: Rasio pola p_{ij} terhadap kelas (j)

L : *Sentiment lexicon*

$|L|$: Jumlah *sentiment lexicon*

$\sum_{p_{ij} \in L_j} C(p_{ij})$: Jumlah *sentiment lexicon* yang termasuk kelas (j) yang terdapat pada dokumen

$\sum_{p_{ij} \in L} C(p_{ij})$: Jumlah *sentiment lexicon* seluruh kelas yang terdapat pada dokumen

$C(p_{ij})$: Jumlah *sentiment lexicon* dihitung

H. Evaluasi

Skor evaluasi dihitung menggunakan *Confusion Matrix* untuk memperoleh skor *precision*, *recall*, dan *F1*. *Confusion Matrix* adalah tabel yang dihasilkan oleh model pengklasifikasi serta dapat menggambarkan kinerja model pengklasifikasi [12]. Dalam *Confusion Matrix* terdapat beberapa istilah diantaranya *True Positives* yaitu kelas actual dan prediksi benar, *True Negatives* yaitu kelas aktual benar dan prediksi salah, *False Positives* yaitu kelas aktual salah dan prediksi benar, dan *False Negatives* yaitu kelas aktual dan prediksi salah.

- 1) *Preprocessing*: Rasio prediksi benar terhadap kelas aktual benar. Formula *precision* ditunjukkan pada formula (4).

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (4)$$

- 2) *Recall*: Rasio prediksi benar terhadap seluruh kelas aktual. Formula *recall* ditunjukkan pada formula (5).

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (5)$$

3) *F1*: Rata – rata *precision* dan *recall*. Formula *F1* ditunjukkan pada formula (6).

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

IV. HASIL DAN PEMBAHASAN

Perhitungan probabilitas kelas sentimen dengan *Naive Bayes* memungkinkan masing – masing kelas sentimen memperoleh nilai probabilitas yang sama yang dapat menurunkan kinerja model pengklasifikasi. Pada (7) dan (8) menunjukkan ulasan dengan sentimen positif “Pelayanan bgs dan murh. Tp suasana berisik” yang telah melewati *preprocessing* dan *feature extraction* maka akan diperoleh fitur “pelayanan bagus”, “murah”, “suasana berisik” memperoleh nilai probabilitas positif dan probabilitas negatif yang sama.

$$\begin{aligned} \text{likelihood}_{\text{positif}} &= P(C_{\text{positif}}) \times P('pelayanan\ bagus'|C_{\text{positif}}) \times P('murah'|C_{\text{positif}}) \\ &\times P('berisik'|C_{\text{positif}}) \times P(C_{\text{positif}}) \\ &= 0.6 \times 0.4 \times 0.6 \times 0.4 = 0.0576 \end{aligned} \quad (7)$$

$$\begin{aligned} \text{likelihood}_{\text{negatif}} &= P(C_{\text{negatif}}) \times P('pelayanan\ bagus'|C_{\text{negatif}}) \times P('murah'|negatif) \\ &\times P('berisik'|C_{\text{negatif}}) \times P(C_{\text{negatif}}) \\ &= 0.4 \times 0.6 \times 0.4 \times 0.6 = 0.0576 \end{aligned} \quad (8)$$

Hal tersebut dapat terjadi karena fitur “pelayanan bagus” meskipun fitur pada *sentiment lexicon* berada pada *lexicon* positif namun fitur “pelayanan bagus” lebih sering muncul pada ulasan negatif sehingga penentuan kelas ditentukan secara acak yang mengakibatkan penurunan kinerja model pengklasifikasi. Perhitungan pada *Improved Naive Bayes* dengan melibatkan rasio *sentiment lexicon* dapat meningkatkan prediksi kelas sentimen seperti terlihat pada (9) dan (10). Pada ulasan dengan fitur “pelayanan bagus”, “murah”, “suasana berisik” terdapat total 3 *sentiment lexicon*, 2 *sentiment lexicon* positif dan 1 *sentiment lexicon* negative

$$\begin{aligned} \text{likelihood}_{\text{positif}} &= R_1 \times P(C_{\text{positif}}) \times P('pelayanan\ bagus'|C_{\text{positif}}) \times P('murah'|C_{\text{positif}}) \\ &\times P('berisik'|C_{\text{positif}}) \times P(C_{\text{positif}}) \\ &= \left(\frac{2}{3}\right) \times 0.6 \times 0.4 \times 0.6 \times 0.4 = 0.0384 \end{aligned} \quad (9)$$

$$\begin{aligned} \text{likelihood}_{\text{negatif}} &= R_1 \times P(C_{\text{negatif}}) \times P('pelayanan\ bagus'|C_{\text{negatif}}) \times P('murah'|negatif) \\ &\times P('berisik'|C_{\text{negatif}}) \times P(C_{\text{negatif}}) \\ &= \left(\frac{1}{3}\right) \times 0.4 \times 0.6 \times 0.4 \times 0.6 = 0.0192 \end{aligned} \quad (10)$$

Pada perhitungan (9) dan (10) menunjukkan *Improved Naive Bayes* dapat memprediksi kelas sentimen dengan lebih baik. *Improved Naive Bayes* dapat memprediksi ulasan “Pelayanan bgs dan murh. Tp suasana berisik” merupakan ulasan dengan sentimen positif.

Setelah dilakukan pengujian menggunakan 10 *dataset* hasil pengacakan yang sebelumnya sudah dilakukan diperoleh skor evaluasi klasifikasi menggunakan *Naive Bayes* yang ditunjukkan oleh Tabel IV dan *Improved Naive Bayes* ditunjukkan oleh Tabel V.

TABEL IV
 SKOR EVALUASI KLASIFIKASI MENGGUNAKAN *NAÏVE BAYES*

<i>Dataset</i>	<i>Precision</i>	<i>Recall</i>	<i>F1</i>
<i>Dataset 1</i>	76%	63%	56%
<i>Dataset 2</i>	77%	65%	59%
<i>Dataset 3</i>	63%	59%	53%
<i>Dataset 4</i>	70%	62%	54%
<i>Dataset 5</i>	70%	61%	53%
<i>Dataset 6</i>	72%	59%	52%
<i>Dataset 7</i>	64%	56%	46%
<i>Dataset 8</i>	62%	58%	50%
<i>Dataset 9</i>	71%	57%	48%
<i>Dataset 10</i>	68%	58%	49%
Rata - rata	69%	60%	52%

TABEL V
 SKOR EVALUASI KLASIFIKASI MENGGUNAKAN *IMPROVED NAÏVE BAYES*

Dataset	Precision	Recall	F1
<i>Dataset 1</i>	78%	72%	70%
<i>Dataset 2</i>	75%	71%	69%
<i>Dataset 3</i>	69%	66%	63%
<i>Dataset 4</i>	72%	69%	66%
<i>Dataset 5</i>	75%	71%	68%
<i>Dataset 6</i>	75%	68%	66%
<i>Dataset 7</i>	71%	64%	60%
<i>Dataset 8</i>	72%	65%	61%
<i>Dataset 9</i>	76%	67%	63%
<i>Dataset 10</i>	73%	67%	64%
Rata - rata	74%	68%	65%

Pada Tabel IV menunjukkan klasifikasi dengan *Naive Bayes* menggunakan *dataset 2* memperoleh skor evaluasi terbaik dibandingkan menggunakan *dataset* lainnya dengan skor *precision* 77%, *recall* 65%, dan *F1*

59%. Pada Tabel V menunjukkan klasifikasi dengan *Improved Naive Bayes* menggunakan *dataset 1* memperoleh skor evaluasi terbaik dibandingkan menggunakan *dataset* lainnya dengan skor *precision* 78%, *recall* 72% dan *F1* 70%. Rata – rata skor evaluasi dengan *Naive Bayes* memperoleh skor *precision* 69%, *recall* 60% dan *F1* 52% dan rata – rata skor evaluasi dengan *Improved Naive Bayes* memperoleh skor *precision* 74%, *recall* 68% dan *F1* 65%. Berdasarkan perolehan tersebut *Improved Naive Bayes* dapat mengklasifikasikan ulasan tempat makan berbahasa Indonesia dengan lebih baik dibandingkan *Naive Bayes*.

Kesulitan yang dihadapi pada penelitian ini ialah belum tersedianya *POS Tagger* bahasa Indonesia yang memadai sehingga beberapa fitur tidak dapat terekstrak. Seperti pada ulasan “Enak dan nyaman untuk harga relatif menurut saya pribadi.”, kata “enak” pada ulasan tersebut yang seharusnya merupakan kelas kata *adjectives* dapat teridentifikasi sebagai kelas kata *nouns* sehingga *unigram* “enak” tidak dapat diekstrak sebagai fitur. Tabel VI menunjukkan kesalahan *POS Tagger* dalam mengidentifikasi label kelas kata.

TABEL VI
KESALAHAN *POS TAGGING* DALAM MENGIDENTIFIKASI LABEL KELAS KATA

Ulasan	Enak dan nyaman untuk harga relatif menurut saya pribadi.
Setelah preprocessing	[('enak', 'NN'), ('nyaman', 'JJ'), ('harga', 'NN'), ('relatif', 'JJ'), ('pribadi', 'VB')]
Fitur yang terekstrak	['nyaman']

V. KESIMPULAN

Kinerja *Naive Bayes* dapat menurun karena terdapat kemungkinan masing – masing kelas sentimen memperoleh probabilitas yang sama. *Improved Naive Bayes* dengan melibatkan rasio *sentiment lexicon* yang ditentukan secara khusus untuk ulasan tempat makan yang mengekspresikan rasa makanan, fasilitas, harga, layanan dan tempat dapat meningkatkan kinerja klasifikasi. Evaluasi berdasarkan *precision*, *recall*, dan *F1* menunjukkan *Improved Naive Bayes* memiliki kinerja yang lebih baik dibandingkan *Naive Bayes*. Kesulitan yang dihadapi pada analisis sentimen ulasan tempat makan berbahasa Indonesia dengan *sentiment lexicon* ialah belum tersedianya *POS Tagger* bahasa Indonesia yang dapat memberikan label kelas kata dengan keakuratan yang tinggi. Analisis sentimen untuk kasus ulasan tempat makan berbahasa Indonesia masih dapat terus dikembangkan karena pada saat ini belum terdapat *POS Tagger* bahasa Indonesia yang memadai.

REFERENSI

- [1] comScore, “Online Consumer-Generated Reviews Have Significant Impact on Offline Purchase Behavior,” *Study Conduct. by comScore Kelsey Gr. Reveal. that 24 Percent Online Local Serv. Rev. Users Purch. Serv.*, 2007.
- [2] L. Zhang and B. Liu, “Sentiment Analysis and Opinion Mining,” *Encycl. Mach. Learn. Data Min.*, no. May, pp. 1–10, 2016.
- [3] K. Khalifa and N. Omar, “A hybrid method using lexicon-based approach and Naive Bayes classifier for Arabic opinion question answering,” *J. Comput. Sci.*, vol. 10, no. 10, pp. 1961–1968, 2014.
- [4] R. Feldman, “Techniques and applications for sentiment analysis,” *Commun. ACM*, vol. 56, no. 4, p. 82, 2013.
- [5] B. R. Samal, A. K. Behera, and M. Panda, “Performance analysis of supervised machine learning techniques for sentiment analysis,” *Proc. 2017 3rd IEEE Int. Conf. Sensing, Signal Process. Secur. ICSSS 2017*, pp. 128–133, 2017.
- [6] Z. Zhang, Q. Ye, Z. Zhang, and Y. Li, “Sentiment classification of Internet restaurant reviews written in Cantonese,” *Expert Syst. Appl.*, vol. 38, no. 6, pp. 7674–7682, 2011.
- [7] H. Kang, S. J. Yoo, and D. Han, “Senti-lexicon and improved Naive Bayes algorithms for sentiment analysis of restaurant reviews,” *Expert Syst. Appl.*, vol. 39, no. 5, pp. 6000–6010, 2012.
- [8] Y. M. Aye and S. S. Aung, “Senti-Lexicon and Analysis for Restaurant Reviews of Myanmar Text,” no. 5, pp. 380–385, 2018.
- [9] linggom, “GitHub - linggom/BigDataGroupProjects: Group 2 projects on WebBig Data Summer School UI.” 2015.
- [10] R. Manurung, A. Dinakaramani, F. Rashel, and A. Luthfi, “idn-tagged-corpus.” 2016.
- [11] masdevid, “GitHub - masdevid/ID-Stopwords: Stopwords collection of Bahasa Indonesia collected from many sources.” 2016.
- [12] M. N. R, J and P. R, “Performance Analysis of Text Classification Algorithms using Confusion Matrix,” *Int. J. Eng. Tech. Res.*, vol. 0869, no. 4, pp. 75–78, 2016.

