

OPINION MINING PADA REVIEW BUKU MENGUNAKAN ALGORITMA NAÏVE BAYES

Dinda Ayu Muthia

Abstract — *In the era of widespread use of the internet today, the number of consumers who wrote the opinion and experience of online continues to increase. Read the review as a whole can be time consuming, however, if only a few reviews that read, then the evaluation will be biased. Sentiment analysis aims to address this problem by automatically classifying user review be positive or negative opinion. Naïve Bayes classifier is a popular machine learning techniques for text classification, because it is very simple, efficient and has a good performance in many domains. However, Naïve Bayes has the disadvantage that is very sensitive to feature too much, resulting in a classification accuracy becomes low. Therefore, in this study used the integration method of feature selection, namely Information gain and Genetic algorithm in order to improve the accuracy of Naïve Bayes classifier. This research resulted in the classification of the text in the form of positive or negative review of the book. Measurement is based on the accuracy of Naive Bayes before and after the addition of feature selection methods. The evaluation was done using a 10 fold cross validation. While the measurement accuracy is measured by confusion matrix and ROC curves. The results showed an increase in the accuracy of Naive Bayes from 78.50% to 84.50%.*

Intisari — Di era maraknya penggunaan internet saat ini, jumlah konsumen yang menulis opini dan pengalaman secara online terus meningkat. Membaca review tersebut secara keseluruhan bisa memakan waktu, namun, jika hanya sedikit review yang dibaca, maka evaluasi akan bias. Analisa sentimen bertujuan untuk mengatasi masalah ini dengan secara otomatis mengelompokkan review pengguna menjadi opini positif atau negatif. Pengklasifikasi Naïve Bayes adalah teknik *machine learning* yang populer untuk klasifikasi teks, karena sangat sederhana, efisien dan memiliki performa yang baik pada banyak domain. Namun, Naïve Bayes memiliki kekurangan yaitu sangat sensitif pada fitur yang terlalu banyak, yang mengakibatkan akurasi klasifikasi menjadi rendah. Oleh karena itu, dalam penelitian ini digunakan penggabungan metode pemilihan fitur, yaitu Information gain dan Genetic algorithm agar bisa meningkatkan akurasi pengklasifikasi Naïve Bayes. Penelitian ini menghasilkan klasifikasi teks dalam bentuk positif atau negatif dari review buku. Pengukuran berdasarkan akurasi Naive Bayes sebelum dan sesudah penambahan metode pemilihan fitur. Evaluasi dilakukan menggunakan 10 fold cross validation. Sedangkan pengukuran akurasi diukur dengan confusion matrix dan kurva ROC. Hasil penelitian menunjukkan peningkatan akurasi Naïve Bayes dari 78.50% menjadi 84.50%.

Kata kunci — Buku, Klasifikasi teks, Naive Bayes, *Opinion mining*, Review.

Program Studi Manajemen Informatika AMIK BSI Bekasi, Jln. Cut Mutiah No. 88, Bekasi 17113. Telp (021) 82426027 Fax. (021) 82425638 ; e-mail: dinda.dam@bsi.ac.id.

I. PENDAHULUAN

Di era maraknya penggunaan internet saat ini, jumlah konsumen yang menulis opini dan pengalaman secara online terus meningkat. Membaca review tersebut secara keseluruhan bisa memakan waktu, namun, jika hanya sedikit review yang dibaca, maka evaluasi akan bias. Klasifikasi sentimen bertujuan untuk mengatasi masalah ini dengan secara otomatis mengelompokkan review pengguna menjadi opini positif atau negatif [1].

Ada beberapa penelitian yang sudah dilakukan dalam melakukan *opinion mining* terhadap review yang tersedia secara online diantaranya, *opinion mining* pada opini review film menggunakan pengklasifikasi *Support Vector Machine* dan *Particle Swarm Optimization* yang dilakukan oleh Basari pada tahun 2013. *Opinion mining* pada review restoran di internet yang ditulis dalam bahasa Canton menggunakan pengklasifikasi *Naïve Bayes* dan *Support Vector Machine* yang dilakukan oleh Z. Zhang pada tahun 2011. Analisa sentimen pada review film dari IMDB, review produk seperti GPS, buku, dan kamera dari Amazon.com menggunakan pengklasifikasi *Support Vector Machine* dan *Artificial Neural Network* yang dilakukan oleh Moraes pada tahun 2013.

Penggunaan fitur n-gram dalam menggabungkan kata agar dapat dilihat perbedaan sentimen dari tiap gabungan kata rupanya berpengaruh dengan akurasi yang dihasilkan pengklasifikasi seperti yang telah diteliti oleh Zhang dan kawan-kawan pada tahun 2011. Jika fitur yang digunakan 2-gram (penggabungan 2 kata) atau 3-gram (penggabungan 3 kata), maka otomatis jumlah kata yang diproses akan semakin banyak dan bisa mempengaruhi akurasi klasifikasi.

Pengklasifikasi *Naïve Bayes* sangat sederhana dan efisien, [2]. Di samping kesederhanaannya, pengklasifikasi *Naïve Bayes* adalah teknik *machine learning* yang populer untuk klasifikasi teks, dan memiliki performa yang baik pada banyak domain [3]. Namun, *Naïve Bayes* memiliki kekurangan yaitu sangat sensitif dalam pemilihan fitur [2]. Terlalu banyak jumlah fitur, tidak hanya meningkatkan waktu penghitungan tapi juga menurunkan akurasi klasifikasi [4].

Hal lain yang ditemukan dalam pendekatan *opinion mining* adalah pemilihan fitur. Pemilihan fitur bisa membuat pengklasifikasi baik lebih efisien/efektif dengan mengurangi jumlah data yang dianalisa, maupun mengidentifikasi fitur yang sesuai untuk dipertimbangkan dalam proses pembelajaran. Ada dua jenis utama metode pemilihan fitur dalam *machine learning*: *wrapper* dan *filter*. *Wrapper* menggunakan akurasi klasifikasi dari beberapa algoritma sebagai fungsi evaluasinya. Metode *filter* terdiri dari

Document Frequency, *mutual information*, *Information Gain*, dan *Chi-Square*. Tidak ada dari keempat metode tersebut yang secara luas diterima sebagai metode penyeleksi fitur terbaik untuk klasifikasi sentimen atau kategorisasi teks, namun, *Information Gain* sering lebih unggul dibandingkan yang lain. *Wrapper* mengevaluasi fitur secara berulang dan menghasilkan akurasi klasifikasi yang tinggi. Salah satu metode *wrapper* yang bisa digunakan dalam pemilihan fitur adalah *Genetic algorithm* (GA).

Umumnya metode pemilihan fitur yang lebih disukai adalah filter dikarenakan waktu pemrosesannya yang relatif rendah. *Information Gain* mengukur berapa banyak informasi kehadiran dan ketidakhadiran dari suatu kata yang berperan untuk membuat keputusan klasifikasi yang benar dalam *class* apapun. *Information Gain* adalah salah satu pendekatan filter yang sukses dalam pengklasifikasian teks [4]. Untuk mengurangi kerumitan perhitungan dilakukan pemilihan fitur dengan menghitung *Information Gain* [1].

Pada penelitian ini penggunaan fitur 2-gram dan 3-gram akan dibandingkan untuk mengetahui seberapa besar pengaruhnya terhadap akurasi klasifikasi dan pengklasifikasi *Naïve Bayes* dengan *Information Gain* dan *Genetic algorithm* sebagai metode pemilihan fitur akan diterapkan untuk mengklasifikasikan opini dari review suatu buku untuk meningkatkan akurasi analisa sentimen.

Pengklasifikasi *Naïve Bayes* sangat sederhana, efisien dan merupakan teknik *machine learning* yang populer untuk klasifikasi teks, serta memiliki performa yang baik pada banyak domain. Namun, *Naïve Bayes* memiliki kekurangan yaitu sangat sensitif pada fitur yang terlalu banyak, yang mengakibatkan akurasi klasifikasi menjadi rendah.

Seberapa besar efek metode pemilihan fitur *Information Gain* dan *Genetic Algorithm* pada akurasi analisa sentimen pada review buku menggunakan pengklasifikasi *Naïve Bayes*?

Tujuan dari penelitian ini adalah untuk mengintegrasikan metode pemilihan fitur *Information Gain* dan *Genetic algorithm* dalam menganalisa sentimen pada review buku menggunakan pengklasifikasi *Naïve Bayes*.

Manfaat penelitian ini dapat diuraikan sebagai berikut:

1. Membantu para calon pembaca dalam mengambil keputusan saat ingin membaca suatu buku agar bisa mengurangi waktu dalam membaca review dan komentar dari suatu buku.
2. Membantu para calon penonton dalam mengambil keputusan saat ingin menonton film yang dibuat berdasarkan buku-buku *best-seller* agar bisa mengurangi waktu dalam membaca review yang umumnya membandingkan mana yang lebih bagus ceritanya, versi buku atau versi filmnya.
3. Membantu para pengembang sistem yang berkaitan dengan review film, baik dari sumber IMDB, Goodreads, maupun dari sosial media lainnya seperti Twitter, Blog, dan lain-lain.

Mengklasifikasikan teks analisa sentimen pada review suatu buku dengan menggunakan pengklasifikasi *Naïve Bayes*

yang menerapkan metode pemilihan fitur *Information gain* dan *Genetic algorithm*.

II. KAJIAN LITERATUR

2.1. Opini

Dalam banyak kasus keputusan yang kita buat dipengaruhi oleh opini dari orang lain. Sebelum kesadaran akan internet menjadi tersebar luas, banyak dari kita yang biasanya menanyakan opini teman atau tetangga mengenai alat elektronik ataupun film sebelum benar-benar membelinya. Dengan berkembangnya ketersediaan dan popularitas akan sumber yang kaya opini seperti *website* review online dan blog pribadi, kesempatan baru dan tantangan muncul semenjak orang-orang sekarang bisa dan menggunakan secara aktif informasi dan teknologi untuk mencari dan memahami opini orang lain [5].

2.2. Opinion Mining

Opinion mining didefinisikan sebagai tugas menemukan opini dari penulis tentang entitas tertentu. Menurut Tang dalam Haddi [7], analisa sentimen pada review adalah proses menyelidiki review produk di internet untuk menentukan opini atau perasaan terhadap suatu produk secara keseluruhan. Menurut Thelwall dalam Haddi, analisa sentimen diperlakukan sebagai suatu tugas klasifikasi yang mengklasifikasikan orientasi suatu teks ke dalam positif atau negatif. Menurut Mejova dalam Basari [8], tujuan dari analisa sentimen adalah untuk menentukan perilaku atau opini dari seorang penulis dengan memperhatikan suatu topik tertentu. Perilaku bisa mengindikasikan alasan, opini atau penilaian, kondisi kecenderungan (bagaimana si penulis ingin mempengaruhi pembaca) [6].

2.3. Pemilihan Fitur (*Feature Selection*)

Pemilihan fitur digunakan untuk menghilangkan fitur yang tidak relevan dan berulang, yang mungkin menyebabkan kekacauan, dengan menggunakan metode tertentu. Menurut John, Kohavi, dan Pflieger dalam Chen [2] ada dua jenis metode pemilihan fitur dalam *machine learning*, yaitu *wrapper* dan *filter* [9].

Wrapper menggunakan akurasi klasifikasi dari beberapa algoritma sebagai fungsi evaluasinya [2]. Salah satu metode *wrapper* yang bisa digunakan dalam pemilihan fitur adalah *Genetic algorithm* (GA) [10].

1. *Genetic Algorithm*

Genetic algorithm berusaha untuk menggabungkan ide-ide evolusi alam [11]. Secara umum, pembelajaran genetika dimulai sebagai berikut:

- a. Sebuah populasi awal dibuat terdiri dari aturan acak. Setiap aturan bisa diwakili oleh string bit. Sebagai contoh sederhana, misalkan bahwa sampel dalam satu set pelatihan yang diberikan dijelaskan oleh dua atribut Boolean, A1 dan A2, dan bahwa ada dua kelas, C1 dan C2. Aturan "*If A1 And Not A2 Then C2*" dapat

dikodekan sebagai string bit "100," di mana dua bit paling kiri mewakili atribut A1 dan A2, masing-masing, dan bit paling kanan mewakili kelas. Demikian pula, aturan "If Not A1 And Not A2 Then C1 "dapat dikodekan sebagai" 001. "Jika atribut memiliki nilai-nilai k, di mana $k > 2$, maka k bit dapat digunakan untuk mengkodekan nilai-nilai atribut itu. Kelas dapat dikodekan dengan cara yang sama.

- b. Berdasarkan gagasan ketahanan dari yang paling sesuai, populasi baru terbentuk terdiri dari aturan yang paling sesuai dalam populasi saat ini, serta keturunan aturan ini. Biasanya, fitness aturan dinilai dengan akurasi klasifikasi pada satu set sampel pelatihan.
- c. Keturunan diciptakan dengan menerapkan operator genetika seperti *crossover* dan mutasi. Dalam *crossover*, substring dari sepasang aturan ditukar untuk membentuk pasangan aturan baru. Dalam mutasi, bit yang dipilih secara acak dalam aturan string dibalik.
- d. Proses menghasilkan populasi baru berdasarkan aturan populasi sebelumnya berlanjut sampai populasi, P, berkembang di mana setiap aturan dalam P memenuhi ambang batas *fitness* yang sudah ditentukan.

Genetic algorithm mudah disejajarkan dan telah digunakan untuk klasifikasi seperti masalah optimasi lainnya. Dalam data mining, algoritma genetika dapat digunakan untuk mengevaluasi fitness algoritma lainnya.

Metode filter terdiri dari *Information gain*, *Term frequency*, *Chi-square*, *Expected cross entropy*, *Odds ratio*, *The weight of evidence of text*, *Mutual information*, dan *Gini index* [2]. Menurut Santosa dalam [12] ada pula metode *Gain ratio*. Ada pula metode *Document Frequency*. Tidak ada dari keempat metode tersebut yang secara luas diterima sebagai metode penyeleksi fitur terbaik untuk klasifikasi sentimen atau kategorisasi teks, namun, *information gain* sering lebih unggul dibandingkan yang lain [12].

1. *Information Gain*

Tahapan dalam proses perhitungan *Information gain* sebagai berikut:

- a. Cari nilai entropi sebelum pemisahan dengan rumus berikut:

$$Entropi(y) = -P_i \log_2 P_i$$

P_i : proporsi data y dengan kelas i

- b. Cari nilai entropi setelah pemisahan berdasarkan atribut A dengan rumus berikut:

$$\text{Total entropi} = \sum_{C \in \text{nilai}(A)} \frac{y_C}{y} \text{entropi } y_C$$

- c. Cari nilai *information gain* dengan rumus berikut:

$$\text{gain}(y, A) = \text{entropi}(y) - \sum_{C \in \text{nilai}(A)} \frac{y_C}{y} \text{entropi}(y_C)$$

2.4. Algoritma *Naïve Bayes*

Tahapan dalam algoritma *Naive Bayes* [13]:

1. Hitung probabilitas bersyarat/*likelihood*:

$$P(x | C) = P(x_1, x_2, \dots, x_n | C)$$

C = class

x = vektor dari nilai atribut n

$P(x_i|C)$ = proporsi dokumen dari class C yang mengandung nilai atribut x_i

2. Hitung probabilitas *prior* untuk tiap class:

$$P(C) = \frac{N_j}{N}$$

N_j = jumlah dokumen pada suatu class

N = jumlah total dokumen

3. Hitung probabilitas *posterior* dengan rumus:

$$P(C|x) = \frac{P(x|C) P(C)}{P(x)}$$

Menurut Santoso [14] dengan kata-kata yang lebih umum, rumus Bayes bisa diberikan sebagai berikut:

$$\text{Posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

III. METODE PENELITIAN

Metode penelitian yang penulis lakukan adalah metode penelitian eksperimen, dengan tahapan sebagai berikut:

- a. Pengumpulan Data

Penulis menggunakan data review buku-buku *best-seller* yang di tahun mendatang diangkat menjadi sebuah film, diantaranya buku *Divergent*, *Horns*, *Hunger Games: Mockingjay*, dan *Vampire Academy* yang didapat dari situs khusus review buku yaitu www.goodreads.com yang terdiri dari 100 review positif dan 100 review negatif.

- b. Pengolahan Awal Data

Dataset ini dalam tahap *preprocessing* harus melalui 4 proses, yaitu:

- 1) *Tokenization*

Yaitu mengumpulkan semua kata yang muncul dan menghilangkan tanda baca maupun simbol apapun yang bukan huruf.

- 2) *Stopwords Removal*

Yaitu penghapusan kata-kata yang tidak relevan, seperti *the*, *of*, *for*, *with*, dan sebagainya.

- 3) *Stemming*

Yaitu mengelompokkan kata ke dalam beberapa kelompok yang memiliki kata dasar yang sama, seperti *drug*, *drugged*, dan *drugs* di mana kata dasar dari semuanya adalah kata *drug*.

- 4) *Generate N-grams*

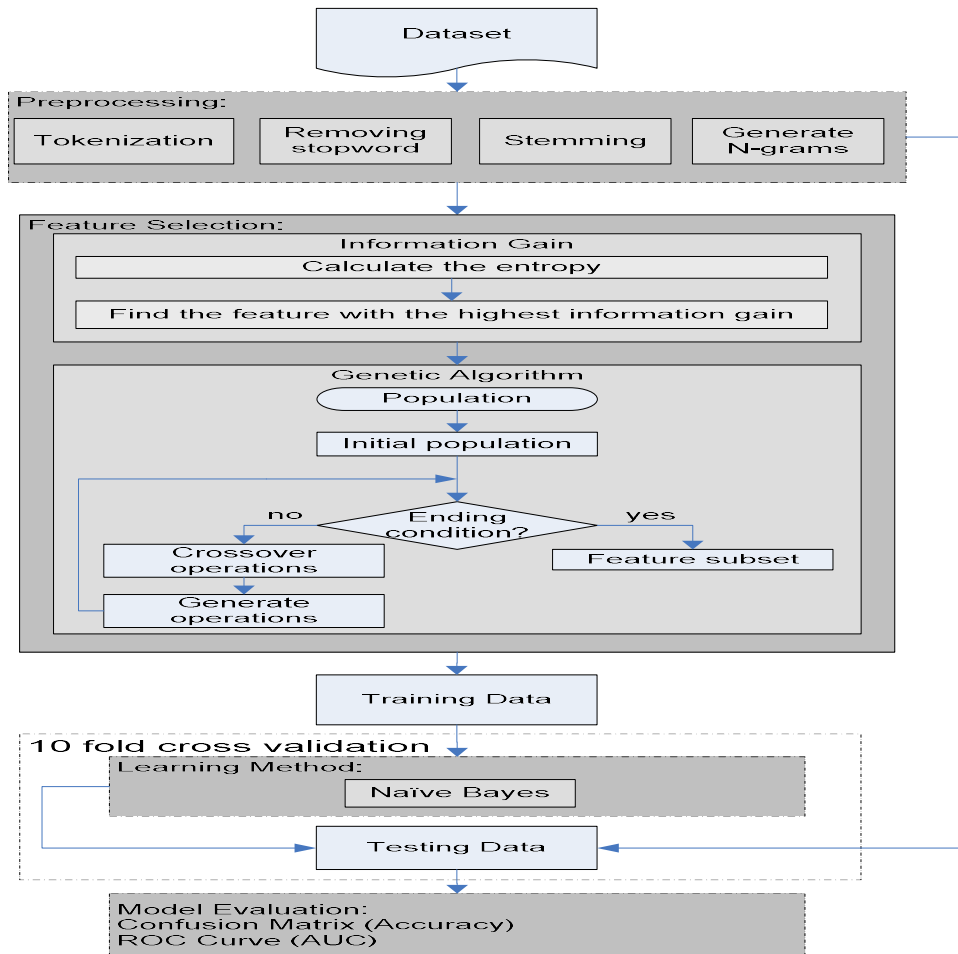
Yaitu menggabungkan kata sifat yang seringkali muncul untuk menunjukkan sentimen, seperti kata *definitely* dan kata *recommended*. Kata *recommended* memang sudah menunjukkan sentimen bentuk opini positif. Kata *definitely* tidak akan berarti jika berdiri sendiri. Namun jika dua kata tersebut digabung menjadi "*definitely recommended*", maka akan sangat menguatkan opini positif tersebut. Penulis hanya menggunakan penggabungan dua kata, yang disebut 2-grams (bigrams).

Sedangkan untuk tahap *transformation* dengan melakukan pembobotan TF-IDF pada masing-masing kata. Di mana prosesnya menghitung kehadiran atau ketidakhadiran sebuah kata di dalam dokumen. Berapa kali sebuah kata muncul di dalam suatu dokumen juga digunakan sebagai skema pembobotan dari data tekstual.

c. Metode Yang Diusulkan

Metode yang penulis usulkan adalah menggabungkan 2 jenis metode pemilihan fitur, yaitu *filter* dan *wrapper*. Dari jenis *filter*, digunakan *information gain* dan dari

jenis *wrapper*, digunakan *genetic algorithm* sebagai metode pemilihan fitur agar akurasi pengklasifikasi *Naïve Bayes* bisa meningkat. Penulis menggunakan pengklasifikasi *Naïve Bayes* karena sangat sederhana, efisien dan merupakan teknik *machine learning* yang populer untuk klasifikasi teks, serta memiliki performa yang baik pada banyak domain. *Genetic algorithm* yang penulis terapkan adalah menggunakan *Naïve Bayes* yang diuji di dalam tahap *wrapper*. Lihat gambar 1 untuk model yang diusulkan secara lebih detail.



Sumber: Hasil Penelitian (2014)

Gambar 1. Model yang diusulkan

Hasil yang dibandingkan adalah akurasi *Naïve Bayes* dengan penggunaan fitur 2-gram dan 3-gram. Selain itu dibandingkan pula akurasi *Naïve Bayes* sebelum menggunakan metode pemilihan fitur dengan akurasi *Naïve Bayes* setelah menggunakan metode pemilihan fitur gabungan, yaitu *Information gain* dan *Genetic algorithm*. Di mana pada *Genetic algorithm*, *Naïve Bayes* diuji di dalam tahap *wrapper*.

d. Eksperimen dan Pengujian Metode

Untuk eksperimen data penelitian, penulis menggunakan RapidMiner 5 untuk mengolah data.

e. Evaluasi dan Validasi Hasil

Validasi dilakukan menggunakan *10 fold cross validation*. Sedangkan pengukuran akurasi diukur dengan *confusion matrix* dan kurva ROC untuk mengukur nilai AUC. Dengan *confusion matrix*, akurasi *Naïve Bayes* sebelum menggunakan metode pemilihan

fitur dan setelah menggunakan metode pemilihan fitur. Tabel 1 berikut adalah tampilan *confusion matrix* dan rumus perhitungannya menurut Gorunescu (Gorunescu, 2011):

Tabel 1. *Confusion Matrix*

Classification <i>n</i>	Predicted Class			
	Class = Yes	Class = No		
Observed Class	Class = Yes	<table border="1"> <tr> <td><i>a</i> (True positive - TP)</td> <td><i>b</i> (False Negative - FN)</td> </tr> </table>	<i>a</i> (True positive - TP)	<i>b</i> (False Negative - FN)
	<i>a</i> (True positive - TP)	<i>b</i> (False Negative - FN)		
Class = No	<table border="1"> <tr> <td><i>c</i> (False positive - FP)</td> <td><i>d</i> (True negative - TN)</td> </tr> </table>	<i>c</i> (False positive - FP)	<i>d</i> (True negative - TN)	
<i>c</i> (False positive - FP)	<i>d</i> (True negative - TN)			

Sumber: Gorunescu (2011)

$$\text{Akurasi} = \frac{TP + TN}{TP + FN + FP + TN}$$

IV. HASIL DAN PEMBAHASAN

4.1. Hasil

4.1.1. Klasifikasi Teks Menggunakan Algoritma *Naïve Bayes*

Data training yang digunakan dalam pengklasifikasian teks ini terdiri dari 100 review buku positif dan 100 review buku negatif. Data tersebut masih berupa sekumpulan teks yang terpisah dalam bentuk dokumen. Sebelum diklasifikasikan, data tersebut harus melalui beberapa tahapan proses agar bisa diklasifikasikan dalam proses selanjutnya, berikut adalah tahapan prosesnya:

1. Pengumpulan Data

Data review positif disatukan dalam folder dengan nama pos. Sedangkan data review negatif disatukan penyimpanannya dalam folder dengan nama neg. Tiap dokumen berekstensi .txt yang dapat dibuka menggunakan aplikasi Notepad.

2. Pengolahan Awal Data

Proses yang dilalui terdiri dari *tokenization*, *stopwords removal*, *stemming*, dan *generate N-grams*. Hasil pengolahan data awal dapat dilihat pada tabel 2.

Tabel 2. Proses *tokenization*

Review	Tokenization
<i>i need to make</i>	<i>i need to make</i>
<i>something perfectly</i>	<i>something perfectly</i>
<i>clear. i am well aware</i>	<i>clear i am well aware</i>
<i>that i gave 4 stars to Daughter of Smoke and Bone.</i>	<i>that i gave stars to Daughter of Smoke and Bone and i am giving</i>
<i>and i am giving 5 stars to this one.</i>	<i>stars to this one</i>

Sumber: Hasil Penelitian (2015)

Tabel 3. Proses *Stopword removal*

<i>Tokenization</i>	<i>Stopwords removal</i>
<i>i need to make</i>	<i>i make perfectly</i>
<i>something perfectly</i>	<i>clear i aware i</i>
<i>clear i am well aware</i>	<i>gave stars</i>
<i>that i gave stars to</i>	<i>Daughter Smoke</i>
<i>Daughter of Smoke and</i>	<i>Bone i giving</i>
<i>Bone and i am giving</i>	<i>stars</i>
<i>stars to this one</i>	

Sumber: Hasil Penelitian (2015)

Tabel 4. Proses *Stemming*

<i>Stopwords removal</i>	<i>Stemming</i>
<i>i make perfectly clear</i>	<i>i make perfect</i>
<i>i aware i gave stars</i>	<i>clear i awar i</i>
<i>Daughter Smoke</i>	<i>gave star</i>
<i>Bone i giving stars</i>	<i>daughter smoke</i>
	<i>bone i give star</i>

Sumber: Hasil Penelitian (2015)

Tabel 5. Hasil pengolahan 2-gram dan 3-gram

<i>Generate 2-grams</i>	<i>Generate 3-grams</i>
<i>i i_make make</i>	<i>i i_make i_make perfect</i>
<i>make perfect</i>	<i>make make perfect</i>
<i>perfect</i>	<i>make perfect clear</i>
<i>perfect_clear</i>	<i>perfect perfect_clear</i>
<i>clear clear_i i</i>	<i>perfect_clear_i clear</i>
<i>i_aware aware</i>	<i>clear_i clear_i_aware i</i>
<i>aware_i i i_gave</i>	<i>i_aware i_aware_i aware</i>
<i>gave gave_star</i>	<i>aware_i aware_i_gave i</i>
<i>star</i>	<i>i_gave i_gave_star gave</i>
<i>star_daughter</i>	<i>gave_star</i>
<i>daughter</i>	<i>gave_star_daughter star</i>
<i>daughter_smoke</i>	<i>star_daughter</i>
<i>smoke</i>	<i>star_daughter_smoke</i>
<i>smoke_bone bone</i>	<i>daughter</i>
<i>bone_i i i_give</i>	<i>daughter_smoke</i>
<i>give give_star</i>	<i>daughter_smoke_bone</i>
<i>star</i>	<i>smoke smoke_bone</i>
	<i>smoke_bone_i bone</i>
	<i>bone_i bone_i_give i</i>
	<i>i_give i_give_star give</i>
	<i>give_star</i>

Sumber: Hasil Penelitian (2015)

3. Klasifikasi

Proses klasifikasi di sini adalah untuk menentukan sebuah kalimat sebagai anggota *class* positif atau *class* negatif berdasarkan nilai perhitungan probabilitas dari rumus *Bayes* yang lebih besar. Jika hasil probabilitas kalimat tersebut untuk *class* positif lebih besar dari pada *class* negatif, maka kalimat tersebut termasuk ke dalam *class* positif. Jika probabilitas untuk *class* positif lebih kecil dari pada *class* negatif, maka kalimat tersebut termasuk ke dalam *class* negatif. Penulis mendapatkan 4 kata dan 1 gabungan kata yang berhubungan dengan sentimen yang paling sering muncul, yaitu *amaz* yang berasal dari kata *amazing, favorit, great, love* dan *i_love*.

4.1.2. Pengujian Model dengan 10 Fold Cross Validation

Pada penelitian ini, penulis melakukan pengujian model dengan menggunakan teknik 10 *cross validation*, di mana proses ini membagi data secara acak ke dalam 10 bagian. Proses pengujian dimulai dengan pembentukan model dengan data pada bagian pertama. Model yang terbentuk akan diujikan pada 9 bagian data sisanya. Setelah itu proses akurasi dihitung dengan melihat seberapa banyak data yang sudah terklasifikasi dengan benar.

4.1.3. Optimasi Model dengan Gabungan Metode Pemilihan Fitur

Dengan menggabungkan metode pemilihan fitur *filter* dan *wrapper*, di mana dalam penelitian ini metode yang digunakan adalah *Information gain* dari *filter* dan *Genetic algorithm* dari *wrapper*. Data yang akan diolah diberikan bobot dari *Information gain* untuk meningkatkan akurasi pengklasifikasi *Naïve Bayes*.

4.1.4. Eksperimen Terhadap Indikator Model

Untuk mendapatkan model yang baik, beberapa indikator disesuaikan nilainya agar didapatkan hasil akurasi yang tinggi. Dalam penyesuaian indikator pada *Genetic algorithm*, akurasi paling tinggi diperoleh dengan kombinasi *population size=50*, *p initialize=0.8*, *p crossover=0.8*, dan *p generate=1.0*. Hasil akurasi mencapai 84.50%. Jika indikator lainnya turut diubah nilainya, dapat menyebabkan proses pengolahan data menjadi semakin lama.

4.2. Pembahasan

Dengan memiliki model klasifikasi teks pada review, pembaca dapat dengan mudah mengidentifikasi mana review yang positif maupun yang negatif. Dari data review yang sudah ada, dipisahkan menjadi kata-kata, lalu diberikan bobot pada masing-masing kata tersebut. Dapat dilihat kata mana saja yang berhubungan dengan sentimen yang sering muncul dan mempunyai bobot paling tinggi. Dengan demikian dapat diketahui review tersebut positif atau negatif. Dalam penelitian ini, hasil pengujian model akan dibahas melalui *confusion matrix* untuk menunjukkan seberapa baik model yang terbentuk.

Pada gambar 2 bisa dilihat confusion matrix yang dihasilkan di mana algoritma *Naïve Bayes* menggunakan fitur 2-gram. Sedangkan pada gambar 3 bisa dilihat confusion matrix yang dihasilkan di mana algoritma *Naïve Bayes* menggunakan fitur 3-gram.

accuracy: 78.50% +/- 14.33% (mikro: 78.50%)

	true Positive	true Negative	class precision
pred. Positive	88	31	73.95%
pred. Negative	12	69	85.19%
class recall	88.00%	69.00%	

Sumber: Hasil Penelitian (2015)

Gambar 2. Confusion matrix algoritma *Naïve Bayes* menggunakan fitur 2-gram

accuracy: 76.00% +/- 12.00% (mikro: 76.00%)

	true Positive	true Negative	class precision
pred. Positive	75	23	76.53%
pred. Negative	25	77	75.49%
class recall	75.00%	77.00%	

Sumber: Hasil Penelitian (2015)

Gambar 3. Confusion matrix algoritma *Naïve Bayes* menggunakan fitur 2-gram

Tabel 6. Model algoritma *Naïve Bayes* menggunakan fitur 2-gram dan 3-gram

	Algoritma <i>Naïve Bayes</i> dan fitur 2-gram	Algoritma <i>Naïve Bayes</i> dan fitur 3-gram
Sukses klasifikasi review positif	88	75
Sukses klasifikasi review negatif	69	77
Akurasi model	78.50%	76%

Sumber: Hasil Penelitian (2015)

Dari hasil yang ditampilkan pada tabel 6 bisa dilihat bahwa penggunaan fitur n-gram berpengaruh terhadap akurasi. Saat menggunakan fitur 2-gram, akurasi pengklasifikasi *Naïve Bayes* mencapai 78.50%, lebih besar dibandingkan pada saat menggunakan fitur 3-gram yang hanya mencapai 76%.

Tanpa menggunakan metode pemilihan fitur, algoritma *Naïve Bayes* dengan fitur 2-gram sudah menghasilkan akurasi sebesar 78.50%. Akurasi tersebut masih berada pada level *fair classification*, masih kurang akurat, sehingga perlu ditingkatkan lagi menggunakan metode pemilihan fitur.

Untuk pengujian data selanjutnya, akan dilanjutkan dari pengujian data menggunakan fitur 2-gram. Setelah menggunakan metode pemilihan fitur dari *filter* dan *wrapper* yang digabungkan, akurasi algoritma *Naïve Bayes* meningkat menjadi 84.50% seperti yang bisa dilihat pada tabel 7.

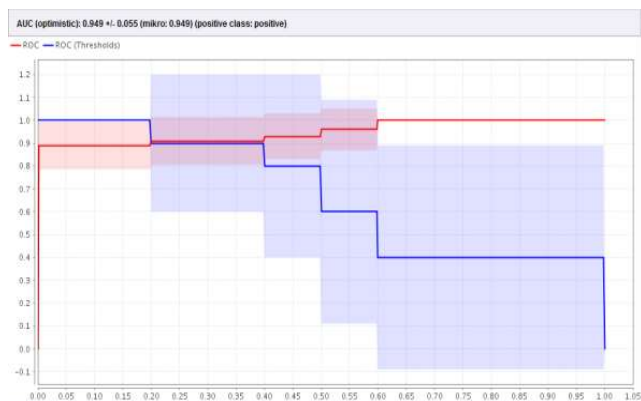
Tabel 7. Model algoritma *Naïve Bayes* sebelum dan sesudah menggunakan metode pemilihan fitur

	Algoritma <i>Naïve Bayes</i> + 2-gram	Algoritma <i>Naïve Bayes</i> + 2-gram + <i>Information Gain</i> & <i>Genetic Algorithm</i>
Sukses klasifikasi review positif	88	77
Sukses klasifikasi review negatif	69	92

Akurasi model	78.50%	84.50%
---------------	--------	--------

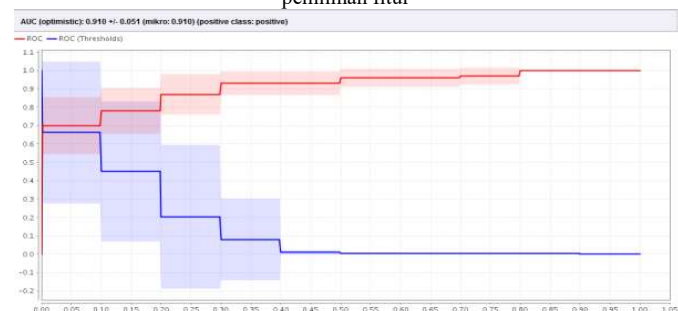
Sumber: Hasil Penelitian (2015)

Berikut adalah tampilan kurva ROC dari hasil uji data. Gambar 4 adalah kurva ROC untuk model *Naïve Bayes* sebelum menggunakan metode pemilihan fitur dan gambar 5 adalah kurva ROC untuk model *Naïve Bayes* setelah menggunakan metode pemilihan fitur.



Sumber: Hasil Penelitian (2015)

Gambar 4. Kurva ROC model *Naïve Bayes* sebelum menggunakan metode pemilihan fitur



Sumber: Hasil Penelitian (2015)

Gambar 5. Kurva ROC model *Naïve Bayes* setelah menggunakan metode pemilihan fitur

4.2.1. Pengukuran dengan *Confusion Matrix*

Pengukuran dengan *confusion matrix* di sini akan menampilkan perbandingan dari hasil akurasi model *Naïve Bayes* sebelum ditambahkan metode pemilihan fitur yang bisa dilihat pada tabel 8 dan setelah ditambahkan metode pemilihan fitur, yaitu penggabungan *Information gain* dan *Genetic algorithm* yang bisa dilihat pada tabel 9.

Tabel 8. *Confusion matrix* model *Naïve Bayes* sebelum penambahan metode pemilihan fitur

Akurasi <i>Naïve Bayes</i>: 78.50% +/- 11.93% (mikro: 78.50%)			
	<i>True negative</i>	<i>True positive</i>	<i>Class precision</i>
<i>Pred. negative</i>	88	31	73.95%
<i>Pred. positive</i>	12	69	85.19%

<i>positif</i>		
<i>Class recall</i>	88.00%	69.00%

Sumber: Hasil Penelitian (2015)

Tabel 9. *Confusion matrix* model *Naïve Bayes* sesudah penambahan metode pemilihan fitur

Akurasi <i>Naïve Bayes</i>: 84.50% +/- 5.22% (mikro: 84.50%)			
	<i>True negative</i>	<i>True positive</i>	<i>Class precision</i>
<i>Pred. negative</i>	92	23	80.00%
<i>Pred. positif</i>	8	77	90.59%
<i>Class recall</i>	92.00%	77.00%	

Sumber: Hasil Penelitian (2015)

4.3. Implikasi Penelitian

Implikasi penelitian ini mencakup beberapa aspek, di antaranya:

1. Implikasi terhadap aspek sistem
Hasil evaluasi menunjukkan penerapan *Information gain* dan *Genetic algorithm* untuk seleksi fitur dapat meningkatkan akurasi *Naïve Bayes* dan merupakan metode yang cukup baik dalam mengklasifikasi teks review buku. Dengan demikian penerapan metode tersebut dapat membantu para calon pembaca dalam mengambil keputusan saat ingin membaca suatu buku, maupun calon penonton film yang kisahnya diangkat dari buku-buku tersebut, agar bisa mengurangi waktu dalam membaca review dan komentar dari buku tersebut.
2. Implikasi terhadap aspek manajerial
Membantu para pengembang sistem yang berkaitan dengan review buku, baik dari sumber www.goodreads.com maupun dari sosial media lainnya seperti Twitter, Blog, dan lain-lain agar menggunakan aplikasi RapidMiner dalam membangun suatu sistem.
3. Implikasi terhadap aspek penelitian lanjutan
Penelitian selanjutnya bisa menggunakan metode pemilihan fitur ataupun dataset dari domain yang berbeda, seperti review produk, review restoran, dan sebagainya.

V. KESIMPULAN

Dari pengolahan data yang sudah dilakukan, terbukti bahwa penggunaan fitur n-gram bisa mempengaruhi akurasi klasifikasi. Penggunaan n-gram yang dapat memperbanyak jumlah kata yang diproses bisa menguntungkan dan bisa juga merugikan di sisi lain. Menguntungkan karena bisa memperlihatkan banyak gabungan kata yang memperkuat sentimen. Sedangkan jumlah kata yang bertambah banyak bisa merugikan dari sisi waktu yang digunakan dalam memproses data. Selain itu jumlah data yang banyak bisa menurunkan akurasi klasifikasi.

Penggabungan metode pemilihan fitur, yaitu *filter* dan *wrapper*, terbukti dapat meningkatkan akurasi pengklasifikasi *Naïve Bayes*. Data review buku dapat diklasifikasi dengan baik ke dalam bentuk positif dan negatif. Akurasi *Naïve Bayes* sebelum menggunakan penggabungan metode pemilihan fitur mencapai 78.50%. Sedangkan setelah menggunakan penggabungan metode pemilihan fitur, yaitu *Information gain* dan *Genetic algorithm*, akurasinya meningkat hingga mencapai 84.50%. Peningkatan akurasi mencapai 6%.

Model yang terbentuk dapat diterapkan pada seluruh data review buku, sehingga dapat dilihat secara langsung hasilnya dalam bentuk positif dan negatif. Hal ini dapat membantu seseorang untuk menghemat waktu saat mencari review suatu buku yang khususnya buku-buku *best-seller* yang akan diangkat ke layar lebar menjadi sebuah film tanpa harus mengkhawatirkan adanya *spoiler* dan pemberian rating yang tidak sesuai dengan reviewnya.



Dinda Ayu Muthia, M.Kom. Tahun 2011 lulus dari Program Strata Satu (S1) Program Studi Sistem Informasi STMIK Nusa Mandiri Jakarta. Tahun 2013 lulus dari Program Strata Dua (S2) Program Studi Magister Ilmu Komputer STMIK Nusa Mandiri Jakarta. Tahun 2015 sudah memiliki Jabatan Fungsional Akademik Asisten Ahli di AMIK BSI Jakarta.

REFERENSI

- [1] 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, Jun. 2011.
- [2] J. Chen, H. Huang, S. Tian, and Y. Qu, "Feature selection for text classification with Naïve Bayes," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 5432–5435, Apr. 2009.
- [3] Q. Ye, Z. Zhang, and R. Law, "Expert Systems with Applications Sentiment classification of online reviews to travel destinations by supervised machine learning approaches," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 6527–6535, 2009.
- [4] A. K. Uysal and S. Gunal, "A novel probabilistic feature selection method for text classification," *Knowledge-Based Syst.*, vol. 36, pp. 226–235, Dec. 2012.
- [5] S. R. R. V, D. V. L. N. Somayajulu, and A. R. Dani, "Classification of Movie Reviews Using Complemented Naive Bayesian Classifier," vol. 1, no. 4, pp. 162–167, 2010.
- [6] R. Feldman, "Techniques and applications for sentiment analysis," *Commun. ACM*, vol. 56, no. 4, p. 82, Apr. 2013.
- [7] E. Haddi, X. Liu, and Y. Shi, "The Role of Text Pre-processing in Sentiment Analysis," *Procedia Comput. Sci.*, vol. 17, pp. 26–32, Jan. 2013.
- [8] A. S. H. Basari, B. Hussin, I. G. P. Ananta, and J. Zeniarja, "Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization," *Procedia Eng.*, vol. 53, pp. 453–462, Jan. 2013.
- [9] F. Gorunescu, *Data Mining Concept Model Technique*. 2011.
- [10] S. Gunal, "Hybrid feature selection for text classification "," vol. 20, 2012.
- [11] J. Han and M. Kamber, *Data Mining Concepts and Techniques*. 2007.
- [12] R. Moraes, J. F. Valiati, and W. P. Gavião Neto, "Document-level sentiment classification: An empirical comparison between SVM and ANN," *Expert Syst. Appl.*, vol. 40, no. 2, pp. 621–633, Feb. 2013.
- [13] Z. Markov and T. Daniel, *Uncovering Patterns in*. 2007.
- [14] Santoso, Budi, *Data Mining Teknik Pemanfaatan Data Untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu. 2007.