Klasifikasi Sentimen pada *Movie Review* dengan Metode Multinomial Naïve Bayes

Sentiment Classification Movie Review Using Multinomial Naïve Bayes Method

Jenepte Wisudawati[1], Adiwijaya[2], Said Al Faraby[3]

[1,2,3]Prodi S1 Teknik Informatika, Fakultas Informatika, Universitas Telkom

[1]**jeneptesimanullang@student.telkomuniversity.ac.id**, [2]**adiwijaya@telkomuniversity.ac.id,**
[3]**saidalfaraby@telkomuniversity.ac.id**

Abstrak

Opini orang lain terhadap suatu *movie review* di media sangat penting dalam membuat suatu keputusan. Untuk mengetahui opini seseorang terhadap suatu *movie review* di media diperlukan sistem yang dapat mempermudah dalam mengetahui opini seseorang. Klasifikasi sentimen dapat membantu dalam membangun sistem untuk mengetahui opini seseorang terhadap *movie review.* Dataset yang digunakan dalam proses klasifikasi sentimen ini adalah *Internet Movie Database* (IMDb). Namun yang menjadi permasalahan dalam mengetahui polaritas suatu opini dalam proses klasifikasi sentimen pada dataset *movie review* adalah adannya data yang tidak terstuktur, atribut data yang begitu banyak serta adannya negasi yang menyebabkan polaritas suatu kata akan berbeda pada konteks teks yang berbeda. Degan permasalahan tersebut maka proses klasifikasi pada dataset tersebut akan di klasifikasikan ke dalam dua kelas polaritas yaitu positif dan negatif. Metode klasifikasi yang digunakan adalah dengan menggunakan metode multinomial naïve bayes. Untuk meningkatkan nilai akurasi metode multinomial naïve bayes dilakukan dengan memecahkan masalah diatas. Dalam memecahkan permasalahan tersebut yang dilakukan adalah pertama, akan dilakukan proses *preprocessing* untuk menangani data *noisy*. Kedua, dilakukan penanganan negasi, adapun lingkup permasalah negasi yang akan dilakukan adalah negasi dengan kata "*not*", "*n't*", "*no*". Ketiga, dilakukan penghitungan bobot setiap kata dengan menggunakan TF-IDF. Berdasarkan hasil pengujian yang telah dilakukan diperoleh nilai akurasi terbesar 85,16%. Hal tersebut dikarenakan multinomial naïve bayes dengan *negation handling* berdasarkan *punctuation, preprocessing* dan TF-IDF dapat meningkatkan nilai akurasi terhadap metode multinomial naïve bayes.

Kata kunci : Multinomial Naïve Bayes, TF-IDF, *Preprocessing,* Negasi

Abstract

Another person's opinion of a movie review in the media is very important in making a decision. To know opinion or polarity someone to a movie review in the media required a system that can facilitate in knowing the polarity of a person. Sentiment classification is one that can help in building the system to know the polarity of a person against a movie review. The dataset used in this sentiment classification process is IMDb (Internet Movie Database). However, the problem of knowing the polarity of an opinion in sentiment classification process on this dataset are their movie review unstructured data, many attributes of data as well as the negation that causes the polarity of a word will be different in the context of different texts. With these problems then the classification process on the dataset will be classified into two classes of polarity that is positive and negative. Before doing the classification process, there are some things to do to reduce the attributes on the dataset. First, there will be a preprocessing process for handling unstructured data. Second, negation handling, to be handled here is the use of the word "not", "no" "n't". The word "not" will be combined with the next word so it becomes a new word. Negation handling of "not","no", "n't" with the technique can determine the polarity of a word in different text context. Third, the TF-IDF process is used as an option to be able to select any terms or words to be used for interior sentiment process. After the three phases, sentiment classification is carried out by using either a machine learning method that is Multinomial Naïve Bayes. Result of reseach show that using preprocessing, negation handling based on punctuation in multinomial naive bayes is 85.16 %.

Keywords: Multinomial Naïve Bayes, TF-IDF, Preprocessing, Negation

1. Pendahuluan

Pertumbuhan populasi pengguna internet semakin meningkat dengan cepat. Populasi tersebut mengakibatkan banyaknya opini di media. Opini orang lain di media sangat penting untuk diketahui polaritasnya karena opini dapat membantu dalam membuat suatu keputusan. Banyak situs yang menyediakan *review* terhadap suatu produk yang dapat mencerminkan opini pengguna, salah satunya adalah situs *Internet Movie Database* (IMDb). IMDb merupakan salah satu situs yang menyediakan database *review movie*. Opini terhadap *movie* sangat penting terhadap pembuatan keputusan.

Klasifikasi sentimen merupakan salah satu sistem yang mampu mendeteksi polaritas dari suatu opini *movie* dengan otomatis dengan bantuan komputerisasi. Namun yang menjadi tantangan terbesar dalam klasifikasi ini adalah pertama, dalam dataset terdapat banyak atribut. Atribut tersebut terdiri dari atribut penting dan atribut *noisy* pada dataset. Hal tersebut mempengaruhi performansi akurasi yang akan dihasilkan oleh suatu metode klasifikasi. Kedua, negasi dimana suatu kata bisa saja polaritasnya positif namun pada konteks teks yang berbeda polaritasnya bisa negatif, seperti kata "*bad*" pada kalimat dengan dua konteks teks yang berbeda yaitu "*the movie is bad"* dan "*the movie is not bad"*. Negasi tersebut merupakan suatu hal yang harus ditangani, karena negasi merupakan salah satu permasalahan yang paling umum yang dapat mengubah polaritas dari suatu teks. Ketiga, format teks yang tidak terstruktur. Ketiga masalah tersebut sangat mempengaruhi terhadap penentuan polaritas terhadap suatu opini.

Sebelumnya telah dilakukan penelitian oleh peneliti [12] pada *movie review* dengan menggunakan metode *mechine learning* yaitu multinomial naïve bayes, maximum entropy, support vector mechine, stochastic gradient descent. Setiap metode tersebut dikombinasikan dengan menggunakan n-gram. Hasil diperoleh bahwa dengan kombinasi n-gram dapat menigkatkan nilai akurasi terdahap setiap metode yang diimplementasikan. Oleh sebab itu penulis mencoba untuk menyelesaikan masalah pada data *movie review* dengan menggunakan metode kombinasi lain dengan menyelesaikan permasalahan diatas untuk meningkatkan nilai akurasi terhadap metode multinomial naïve bayes. Adapun kombinasi yang dilakukan adalah dengan menerapkan proses *preprocessing, negation handling* dan melakukan pembobotan pada setiap fitur. Proses klasifikasi sentimen dengan metode multinomial naïve bayes akan dibagi menjadi dua kelas yaitu kelas sentimen positif dan kelas sentimen negatif. *Preprocessing* dilakukan sebelum dilakukan tahap klasfikasi. Tujuan dari *preprocessing* ini adalah untuk mengurangi atribut *noisy*. Dalam memecahkan permasalahan negasi maka lingkup negasi yang akan diselesaikan yaitu penggunaan kata *"not", "no", "n't".* Untuk mengatasi negasi tersebut penulis akan melakukan dua teknik dalam memecahkan permasalahan negasi. Teknik yang dilakukan adalah pertama, penanganan negasi berdasarkan *punctuation* dimana ketika ditemukan kata "*not*", "*no*", *"n't"* maka akan digabungkan dengan setiap kata setelah negasi sampai

ditemukan *punction*. Teknik kedua, yaitu dengan menggunakan teknik yang sudah diakukan oleh peneliti [8] yaitu dengan metode *rules* dengan bantuan identifikasi kata dengan *pos tagging*. TF-IDF digunakan untuk melakukan pembobotan pada setiap fitur. Diharapkan dengan menggunakan teknik-teknik tersebut dapat meningkatkan hasil nilai akurasi dengan metode klasifikasi yang digunakan yaitu multinomial naïve bayes.

2. Dasar Teori

2.1 Sentimen Analisis

Sentimen analisis sering juga disebut dengan *opinion minding* yang digunakan untuk mengetahui polaritas suatu opini seseorang terhadap suatu domain tertentu. Sentimen analisis berfokus pada opini yang menyiratkan sentimen positif atau negatif terhadap entitas tertentu. Entitas tersebut dapat berupa produk, organisasi, politik dan topik lainnya. Sentimen analisis dikategorikan kedalam dua bagian kategori yaitu pertama, *binary sentiment classification* dan *multi-class sentiment classification*. *Binary sentiment classification* terdiri dari dua label kelas yaitu positif dan negatif, sedangkan *multi-class sentiment classification* terdapat label kelas yang terdiri *dari strong positive, positive, neural, negative, strong negative* [12]. Klasifikasi sentimen dibagi kedalam tiga berdasarkan sumber katanya yaitu level dokumen, level kalimat dan level aspek [12].

1. Level dokumen
   Pada level ini klasifikasi opini dilakukan pada keseluruhan isi dokumen sebagai sebuah sentimen positif dan sentiment negatif. Proses klasifikasi pada pada isi keseluruhan dokumen tersebut dilakukan untuk mengecek apakah dokumen tersebut positif, netral atau negatif. Pada level ini biasanya level dokumen berfokus pada satu entitas.
2. Level Kalimat
   Pada level ini klasifikasi opini dilakukan pada setiap kalimat dalam suatu dokumen untuk mengecek apakah kalimat tersebut positif, netral atau negatif.
3. Level aspek
   Pada level ini klasifikasi opini dilakukan pada suatu dokumen untuk menentukan sentimen pada suatu entitas pada aspek yang berbeda.

Dengan adanya klasifikasi sentimen dapat membantu secara komputerisasi dalam menentukan polaritas terhadap opini seseorang pada suatu produk *movie review*. Klasifikasi *movie review* akan menggunakan *binary sentiment classification* yaitu dengan label kelas positif dan negatif.

2.2 *Natural Languange Preprocessing (NLP)*

*Natural language processing (*NLP*)* merupakan suatu bagian dari bidang ilmu komputer atau komputasi linguistik yang berfokus untuk memproses bahasa alami manusia secara otomatis [6]. Tantangan dalam NLP tersebut adalah memproses bahasa alami manusia itu sendiri dimana dalam bahasa manusia tersebut memiliki ambiguitas, bahasa yang tidak terstruktur, dan bahasa yang beragam, sehingga dibutuhkan NLP dalam membagun sistem untuk memahami bahasa manusia.

2.3 *Preprocessing*

Preprocessing merupakan salah satu bagian natural language preprocessing (NLP) untuk mengelola data yang akan digunakan dalam proses klasifikasi. Preprocessing tujuannya untuk mereduksi fitur atau atribut yang terdapat di dalam data inputan. Peneliti [7] Mengatakan bahwa manfaat dari teks *preprocessing* ini adalah:

a. Mengurangi atribut terhadap dokumen teks dimana atribut yang akan dikurangi adalah atribut yang tidak memiliki pengaruh besar terhadap proses klasifikasi.
b. Melakukan transformasi atribut dengan mentransformasi atribut ke dalam suatu format data yang lebih mudah untuk meningkatkan efektifitas dan efisiensi yang kemudian di proses dengan teknik klasifikasi, misalnya mereduksi jumlah variasi kata dengan melakukan proses stemming dan lematization.

Adapun beberapa teknik *preprocessing* yang sering dilakukan dalam melakukan proses klasifikasi sentimen, adalah :

a. *Cleaning*
   Cleaning yaitu proses membersihkan dokumen dari kata yang tidak diperlukan untuk mengurangi *noisy*. Kata yang dihilangkan adalah *punctuation* dan karakter spesial yaitu (@, #, $, %, angka) yang ada pada data *movie review* . Dalam melakukan proses *cleaning* ini salah satu tools yang sering digunakan adalah dengan penggunaan *library regex*.

b. *Case Folding*
   Case Folding merupan tahapan *preprocessing* untuk mengubah huruf data inputan berupa teks menjadi teks huruf kecil. Huruf yang diterima adalah huruf a-z. Dalam melakukan proses cleaning ini salah satu tools yang sering digunakan adalah dengan penggunaan *library regex*.

c. *Tokenization*

Tokenization merupakan proses yang memotong kalimat atau dokumen teks menjadi bagian-bagian kata yang lebih kecil. Pemotongan setiap kata dalam dokumen teks dapat dilakukan berdasarkan pemisahan kata seperti titik(.) , koma(,), tanda tanya(?) dan spasi.

d. *Stopword Removal*

Stopword Removal merupakan proses untuk menghapus kata yang tidak memiliki pengaruh terhadap proses klasifikasi sentimen. *Stopword Removal* biasannya kata umum yang sering muncul yang tidak memiliki makna. *Stopword Removal* pada teks bahasa inggris seperti *the, is, a, an, that, on etc.*

e. *Stemming*

Stemming merupakan proses untuk mereduksi jumlah variasi dari sebuah kata [7]. *Stemming* akan mengidentifikasikan dan menghapus prefiks serta suffiks dari tiap-tiap kata. Adapun keuntungan dalam proses stemming adalah bisa meningkatkan kemampuan untuk melakukan recall. Namun yang menjadi resiko dari proses stemming adalah hilangnya informasi dari kata yang di-stem sehingga mengakibatkan menurunya akurasi dari proses klasifikasi tersebut.

f. *Lematization*

Lematization merupakan proses untuk mengubah bentuk dasar dari setiap kata sesuai dengan kamus yang ada.

g. *Part-of-speech Tagging (Pos tagging)*

P os tagging adalah salah satu model bahasa yang tujuannya untuk mengidentifikasi, memahami dan memprediksi makna suatu teks. Untuk dapat mengidentifikasi, memahami dan memprediksi suatu teks dengan *Pos tagging* dapat dilakukan terlebih dahulu dengan pemberian kelas kata atau tag setiap kata dalam suatu teks. Identifikasi *pos tagging* yang paling umum ditemukan dalam bahasa Inggris adalah *nouns, verb, adjectives, adverbs, pronouns, prepositions, conjunction, interjections.* Informasi makna suatu teks dapat diketahui dengan melakukan proses identifikasi kata dengan *pos tagging*. Berikut merupakan daftar kelas kata yang digunakan dalam *pos tagging* oleh Penn Treebank POS Tags pada bahasa Inggris [10].

Tabel 1 Penn Treebank POS Tag

Tag	Description	Example	Tag	Description	Example
CC	Coordinating conjunction	and, but, or	PRP$	possessive pronoun	your, one's
CD	Cardinal number	one, two	RB	Adverb	quickly, never
DT	Determiner	a, the	RBR	adverb, comparative	Faster
EX	Existential there	There	RBS	adverb, superlative	Fastest
FW	Foreign word	Mea culpa	RP	Particle	Up, off
IN	Preposition or subordinating conjunction	Of, in, by	SYM	Symbol	+, % , &
JJ	Adjective	Yellow	TO	"to"	To
JJR	Adjective, comparative	Bigger	UH	Interjection	ah, oops
JJS	Adjective, superlative	Wildest	VB	Verb base form	Eat
LS	List item market	1,2,One	VBD	verb past tense	Ate
MD	Modal	Can, should	VBG	verb gerund	Eating
NN	Noun, singular or mass	Illama	VBN	verb past participle	Eaten
NNS	Noun, plural	Ilamas	VBP	Verb non-3sg pres	Eat
NNP	Proper noun, singular	IBM	VBZ	Verb 3sg pres	Eats

NNPS	Proper noun, plural	Carolinas	WDT	wh-determiner	which, that
PDT	Predeterminer	All, both	WP	wh-pronoun	what, who
POS	Possessive ending	‘s	WP$	possessive wh-	Whose
PRP	Personal pronoun	I, you, he	WRB	Wh-adverb	How, where

2.4 Pembobotan

Pembobotan adalah salah satu proses yang dilakukan untuk memberi bobot pada kata dalam dokumen. Salah satu pembobotan yang sering dilakukan dalam proses klasifikasi sentimen adalah TF-IDF (*Term Frequency–Inverse Document Frequency*). TF-IDF merupakan salah satu metode pembobotan pada setiap kata dalam setiap dokumen teks. Bobot ini merepresentasikan seberapa pentingnya suatu kata dalam dokumen teks. TF merupan jumlah nilai kemunculan kata atau *term* dalam suatu dokumen dibagi dengan jumlah seluruh *term* dalam dokumen sedangkan untuk nilai IDF merupakan jumlah seluruh dokumen dalam corpus dibagi dengan banyaknya dokumen suatu *term* muncul. Semakin sering term muncul maka nilai IDF semakin lebih kecil, hal ini disebabkan karena kecilnya kepentingan fitur tersebut didalam dokumen. Untuk rumus menghitung TF-IDF dapat dilakukan dengan persamaan rumus berikut ini:

$$\mathrm{TF} = \left(\frac{\text{jumlah kemunculan term pada satu dokumen}}{\text{jumlah seluruh term dalam satu dokumen}}\right) \quad (2.1)$$

$$IDF = \log\left(\frac{\text{jumlah seluruh dokumen}}{\text{jumlah dokumen suatu term muncul}}\right) \quad (2.2)$$

Berdasarkan rumus (2.1) dan (2.2) maka diperoleh rumus TF-IDF sebagai berikut:

$$\mathrm{TF} - \mathrm{IDF} = \mathrm{TF} \times \mathrm{IDF} \quad (2.3)$$

2.5 *Negation Handling*

Negation atau sering disebut juga negasi menjadi perhatian penting dalam proses klasifikasi pada sentiment, dimana negasi tersebut dapat mengubah polaritas suatu kata dalam konteks kalimat yang berbeda. Pada penelitian [4] dijelaskan bahwa terdapat beberapa lingkup permasalahan negasi yang sering di temukan pada teks, yaitu:

1. *Syntactic*

   Syntatic merupakan salah satu lingkup negasi dimana negasi *syntactic* ini merupakan negasi yang paling umum terdapat dalam suatu teks. Salah satu ciri paling sering ditemukan pada negation *syntactic* ini merupakan suatu negasi yang pada umumnya melekat pada auxiliary contohnya adalah negasi *“not”* pada auxiliary *“was”.*
2. *Diminisher*

   Diminisher merupakan salah satu lingkup negasi. Salah satu lingkup negasi *diminisher* pada umumnya merupakan negasi dengan fitur teridentifikasi *adjective + “ly”, “less”* dan lain sebagainnya.
3. *Morphology*

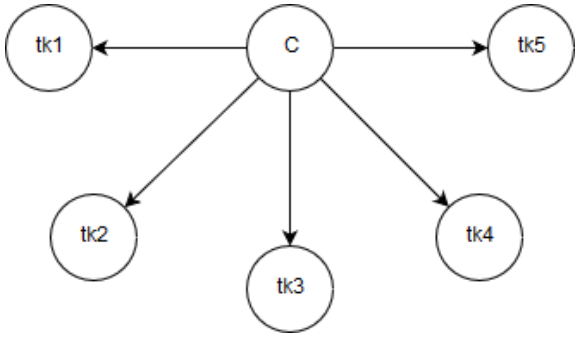
   Morphology merupakan salah satu lingkup negasi dimana suatu fitur ditambah dengan prefixes dan suffix seperti “*un*” pada fitur “*good*”.

Negasi *syntactic* merupakan salah satu lingkup negasi yang paling umum dijumpai dalam suatu teks dimana suatu kata pada konteks teks yang berbeda dapat menghasilkan prediksi kelas yang berbeda. Adapun salah satu contoh kata dalam konteks teks yang berbeda adalah kata “*bad*”. Kata “bad” akan dimasukkan kedalam kalimat dengan dua konteks teks yang berbeda seperti “*the movie is bad”* dan “*the movie is not bad”*.

Lingkup permasalahan negasi yang akan dilakukan dalam pengerjaan tugas akhir ini adalah negasi yang pada umumnya melekat pada auxiliary yaitu negasi “*not*”, *“n’t”,“no”*.

2.6 Naïve Bayes

Naïve Bayes merupakan salah satu metode statistik probabilitas berdasarkan penerapan Teorema Bayes dengan asumsi independen (naif) yang kuat, untuk memprediksi kelas dari suatu dokumen berdasarkan probabilitasnya [13].



Gambar 1 Independen Naïve Bayes

Naïve Bayes *classifier* berasumsi bahwa kehadiran fitur tertentu dari suatu kelas tidak berhubungan dengan kehadiran fitur lainya (independen). Naïve bayes merupakan salah satu pendekatan yang sangat sederhana pada teks klasifikasi. Ada beberapa macam metode Naïve Bayesian yaitu diantaranya adalah naïve bayes, multinomial naïve bayes, bayesian network. Multinomial naïve bayes merupakan salah satu metode spesifik dari metode naïve bayes.

a. Teorema Bayes
Pada teorema bayes ditentukan P(c|d) dimana c merupakan kelas dan d sebagai objek (dokumen) yang akan di klasifikasikan. P(c) merupakan prior probabilitas dari kelas c, P(d|c) merupakan probabilitas dokumen (d) dalam suatu kelas (c). Teorema bayes memiliki rumus sebagai berikut [13].

$$P(c|d) = \frac{P(d|c)P(c)}{P(d)} \quad (2.4)$$

b. Multinomial Naïve Bayes
Multinomial Naïve Bayes merupakan salah satu metode spesifik dari metode Naive Bayes. Multinomial naïve bayes ini juga merupakan salah satu mechine learning dalam *supervised learning* pada proses pengklasifikasian teks dengan menggunakan nilai probabilitas suatu kelas dalam suatu dokumen. Menurut Multinomial Naïve Bayes, secara umum probabilitas suatu dokumen d, sebagai bagian dari anggota kelas c. Probabilitas dari suatu dokumen d terhadap kelas c dapat dihitung dengan rumus sebagai berikut [13].

$$P(c \mid d) \ \alpha \ P(c) \prod_{1}^{n} P(t_k \mid c) \quad (2.5)$$

Dimana:

- $P(t_k \mid c)$ adalah probabilitas kemunculan suatu *term* t_k dalam dokumen pada kelas c dimana t_k adalah *term* dalam dokumen d.
- P(c) adalah prior probabilitas suatu dokumen pada kelas c.

Perhitungan nilai P(c) dan $P(t_k|c)$ dilakukan pada saat melatih data. Probabilitas suatu kelas dapat dilakukan dengan jumlah suatu kelas dokumen dalam kelas latih atau N_c dibagi dengan jumlah total dokumen kelas yang ada atau N dalam dokumen latih [13], sebagai berikut:

$$P(c) = \frac{N_c}{N} \quad (2.6)$$

Untuk perhitungan condisional probabilitas dilakukan untuk menghitung probabilitas kemunculan suatu kata dalam setiap kelas. Conditional probability dapat dilakukan dengan menggunakan frekuensi kemunculan suatu kata pada suatu kelas.

$$P(t|c) = \frac{T_{ct} + 1}{\sum_{t' \epsilon V}(T_{ct'} + 1)} \quad (2.7)$$

Pada proses klasifikasi yang dilakukan pada *movie review* akan menggunakan Multinomial Naïve Bayes, dikarenakan metode ini bagus digunakan dalam proses klasifikasi dan cocok digunakan terhadap dokumen teks. Maximum a posterior (MAP) digunakan untuk menentukan kelas suatu dokumen testing dengan

mengambil nilai maksimum probabilitas setiap dokumen. Adapun rumus untuk MAP adalah sebagai berikut :

$$c_{map} = argmax_{c \in C} \hat{P}(c|d) = argmax_{c \in C} \hat{P}(c) \prod_{1}^{nk} \hat{P}(t_k|c) \quad (2.8)$$

Pada rumus MAP diatas setiap *conditional probability* atau setiap probabilitas suatu kata dikalikan. Perkalian tersebut menghasilkan floating point underflow. Dalam hal ini untuk menghindari floating point underflow maka akan dilakukan proses penjumlahan setiap probabilitas kata dengan menggunakan logaritma dimana log(x,y) =log(x)+ log(y). Untuk mencari MAP pada Multinomial Naïve Bayes adalah sebagai berikut:

$$c_{map} = argmax_{c \in C} [\log \hat{P}(c)] + \sum_{1 \leq k \leq nd} \log \hat{P}(t_k|c)] \quad (2.9)$$

2.7 Evaluasi Performansi

Evaluasi performansi merupakan salah satu parameter yang digunakan untuk mengukur seberapa akurat suatu metode yang di implementasikan. Mengukur parameter evaluasi performansi pada *superviced machine learning* dapat digunakan dengan *confusion matriks* [12], sebagai berikut:

Tabel 2 Confusion Matriks

	Sistem		
Aktual		Positive	Negative
	Positive	TP	FN
	Negative	FP	TN

Dimana :

- TP (*True Positive*) adalah jumlah dokumen review positif di prediksi positif oleh sistem.
- FN (*False Negative*) adalah jumlah dokumen review positif di prediksi negatif oleh sistem.
- FP (*False Positive*) adalah jumlah dokumen review negatif di prediksi positif oleh sistem.
- TN (*True Negative*) adalah jumlah dokumen review negatif di prediksi negatif oleh sistem.

Berdasarkan *confusion matriks* tersebut parameter yang akan digunakan adalah dengan menggunakan nilai *accuracy*, dimana untuk mengukur nilai *accuracy* dilakukan dengan menggunakan rumus dibawah ini:

$$\text{accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (2.10)$$

3. Pembahasan

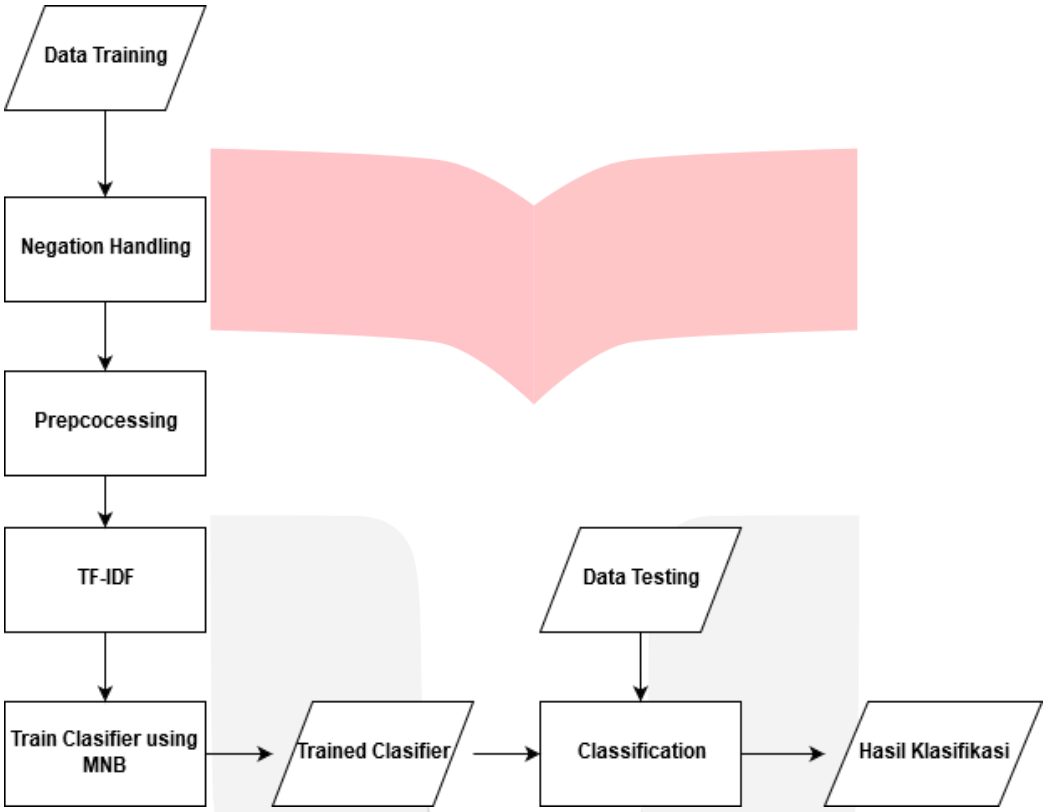
3.1 Dataset

Dalam tugas akhir ini dataset yang digunakan terhadap pengujian dan analisis klasifikasi sentimen adalah dataset *movie review*. Dataset tersebut terdiri dari dua label polaritas dengan jumlah file 2000 yaitu 1000 file dengan polaritas positif dan 1000 file dengan polaritas negatif. Data dibagi kedalam dua bagian yaitu data *training* dan data *testing*. Pembagian data tersebut terdiri atas 30% data digunakan sebagai data *testing* yaitu 292 label positif dan 308 label negatif dan 70% digunakan sebagai data *training* yang terdiri dari 708 label positif dan 692 label negatif.

3.2 Alur Rancangan Sistem

Alur rancangan Sistem akan dibangun dalam tugas akhir ini adalah sistem yang dapat menentukan klasifikasi suatu opini terhadap dokumen *movie review* kedalam kelas negatif dan positif dengan *baseline* menggunakan metode multinomial naïve bayes. Proses penentuan kelas terhadap suatu dokumen *review* dapat dilakukan dalam berbagai tahap. Sebelum melakukan proses penentuan kelas terhadap *review* maka hal yang perlu dilakukan adalah melakukan proses ekstaksi fitur. Dalam ekstraksi fitur tersebut dilakukan untuk memilih fitur-fitur yang akan digunakan di dalam proses klasifikasi. Untuk memilih fitur-fitur yang akan digunakan pada proses klasifikasi akan dilakukan dengan beberapa tahapan yaitu, (1) *Negation handling*, (2) *Preprocessing*, (3) TF-IDF. Fitur yang di ektrak tersebut akan digunakan pada proses klasifikasi dengan menggunakan metode multinomial naïve bayes.

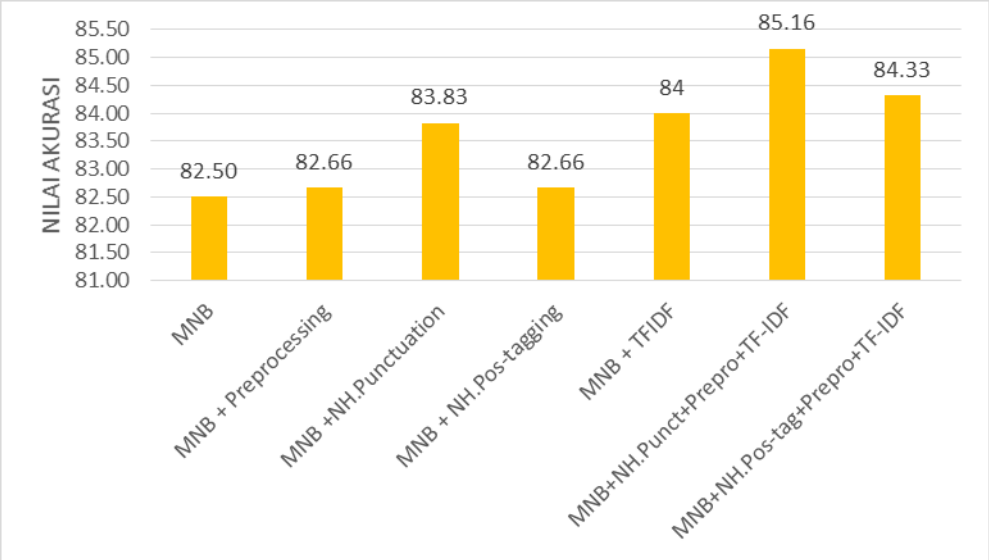
Dalam menguji nilai performansi yang dilakukan dengan menggunakan nilai akurasi yang akan di hitung berdasarkan *confuse matrik* yang dihasilkan oleh sistem. Adapun gambaran umum sistem yasng akan dibuat adalah sebagai berikut:



Gambar 2 Alur rancangan sistem

3.3 Hasil Pengujian dan Analisis

Berdasarkan skenario pengujian yang telah dijabarkan diatas maka akan dihasilkan nilai performansi akurasi pada setiap skenario sesuai dengan tujuan masing-masing skenario dilakukan. Adapun hasil nilai performansi akurasi dari berbagai skenario yang dilakukan adalah sebagai berikut:



Gambar 3 Hasil akurasi setiap skenario pengujian

Pada gambar 3 diatas merupakan hasil nilai akurasi yang dihasilkan oleh setiap skenario pengujian. Dapat dilihat bahwa hasil nilai performansi akurasi yang dihasilkan oleh metode multinomial naïve bayes itu sendiri adalah 82.50%. Nilai performansi akurasi tersebut akan digunakan sebagai *baseline* dalam menganalisis pengaruh setiap

skenario lainnya pada metode multinomial naïve bayes. Berikut merupakan hasil dan analisis dari masing-masing skenario yang dilakukan adalah sebagai berikut:

- Analisi pengaruh *preprocessing* terhadap multinomial naïve bayes

  Pada gambar 3 dapat dilihat bahwa akurasi yang dihasilkan dengan menggunakan *preprocessing* sebesar 82.66%. Jika dibandingkan selisih nilai akurasi dengan *baseline* multinomial naïve bayes sebesar 0.16%. Meningkatnya nilai akurasi disebabkan karena dihilangkannya fitur yang tidak mempengaruhi terhadap proses klasifikasi pada data *movie.* Fitur-firut yang dihilangkan tersebut merupakan fitur yang teridentifikasi sebagai *noisy* yang mempengaruhi terhadap hasil nilai performansi akurasi.
- Analisis pengaruh *negation handling* terhadap multinomial naïve bayes

  Pada gambar 3 dapat dilihat bahwa nilai akurasi yang dihasilkan sistem dengan menggunakan dua metode *negation handling* sangat berbeda. Pada gambar 3 dihasilkan nilai akurasi multinomial naïve bayes degan *negation handling* berdasarkan *punctuation* sebesar 83,83% sedangkan multinomial naïve bayes dengan *negation handling* berdasarkan *pos tangging* dihasilkan nilai akurasi sebesar 82.66%. Perbedaan nilai akurasi yang dihasilkan terhadap metode *baseline* multinomial naive bayes dengan multinomial naïve bayes + *negation handling* berdasarkan *punctuation* sebesar 1,33 % dan 0,16% dengan *negation handling* berdasarkan *pos-tagging*. Selisih akurasi terhadap dua metode *negation handling* yaitu *negation handling* berdasarkan *punctuation* dan *negation handling* berdasarkan *pos tagging* sebesar 1.17%. Berdasarkan hasil tersebut dapat disimpulkan bahwa:
  1. Dengan menggunakan metode *negation handling* terhadap multinomial naïve bayes memiliki pengaruh yang dapat meningkatkan nilai performansi akurasi pada multinomial naïve bayes. Meningkatnya nilai performansi akurasi disebabkan karena metode tersebut cukup mampu dalam memprediksi suatu kelas pada dokumen *review* dengan suatu kata pada konteks teks yang berbeda.
  2. *Negation handling* dengan metode dasar yaitu *punctuation* menghasilkan performansi akurasi yg lebih bagus dibandingkan dengan *negation handling* berdasarkan *pos tagging*. Hal tersebut disebabkan karena suatu kata pada konteks kalimat yang berbeda dengan menggunakan *negation handling* berdasarkan *pos tagging* belum tentu akurat dalam menentukan kelas *review* suatu kata pada suatu kalimat tertentu seperti contoh kalimat " *it isn't nearly as dull as this.*". *Pos tagging* mendeteksi *nearly* sebagai *adverb* (RB), *as* sebagai *adverb* (RB) dan *dull* sebagai *adjective* (JJ). Berdasarkan identifikasi *pos tagging* tersebut dapat dilihat bahwa *rules pos tagging* pada contoh kalimat tersebut dimana suatu kata *dull* tidak dideteksi sebagai *review* kelas negatif hal tersebut disebabkan karena kata setelah negasi tidak terdapat pada ketiga *rules pos tagging* yang digunakan sementara dengan menggunakan *negation handling* berdasarkan *punctuation* mampu menghasilkan kelas suatu *review* tersebut sebagai kelas negatif dimana *dull* tersebut akan digabungkan dengan kata *not* sampai *punctuation* ditemukan yaitu menjadi "*not_dull*". Namun yang menjadi kekurangan dari *negation handling* berdasarkan *punctuation* adalah mengakibatkan banyaknya atribut baru yang menjadikan banyaknya fitur *noisy* atau fitur tidak penting seperti fitur *not_nearly, not_as, not_this* dimana fitur *noisy* tersebut tidak di identifikasi oleh *negation handling* berdasarkan *pos tagging*.
- Analisis pengaruh TF-IDF terhadap multinomial naïve bayes

  Pada gambar 3 dapat dilihat bahwa hasil akurasi yang dihasilkan oleh sistem dengan menggunakan TF-IDF terhadap metode multinomial naïve bayes sebesar 84 %. Jika dibandingkan dengan hasil akurasi baseline yaitu metode multinomial naïve bayes dengan nilai akurasi 82.50 % dengan selisih akurasi 1,50%. Dari hasil tersebut maka dapat disimpulkan bahwa dengan menggunakan metode TF-IDF pada metode multinomial naïve bayes mampu meningkatkan nilai akurasi yang dihasilkan. Hal tersebut disebabkan karena pembobotan terhadap setiap fitur yang dilakukan mampu menurunkan nilai bobot suatu fitur yg sering muncul.
- Analisis pengaruh *negation handling, preprocessing* dan TF-IDF terhadap multinomial naïve bayes

  Pada gambar 3 dapat dilihat bahwa nilai akurasi yang dihasilkan sistem dengan menggunakan dua metode *negation handling* dengan menggunakan *preprocessing* dan TF-IDF sangat berbeda. Pada 3 dihasilkan nilai akurasi multinomial naïve bayes degan *negation handling* berdasarkan *punctuation* sebesar 85.16% sedangkan multinomial naïve bayes dengan *negation handling* berdasarkan *pos tangging* dihasilkan nilai akurasi sebesar 84.33%. Perbedaan nilai akurasi yang dihasilkan terhadap metode *baseline* multinomial naive bayes dengan multinomial naïve bayes + *negation handling* berdasarkan *punctuation* sebesar 2,66% dan 1,83% dengan *negation handling* berdasarkan *pos-tagging*. Selisih akurasi terhadap dua metode *negation handling* yaitu *negation handling* berdasarkan *punctuation* dan *negation handling* berdasarkan pos tagging sebesar 0.83%. Berdasarkan hasil tersebut dapat disimpulkan bahwa penggunaan kedua metode *negation handling* dengan menggunakan kombinasi *preprocessing* dan TF-IDF dapat meningkatkan nilai performansi akurasi. Hal tersebut disebabkan karena *noisy* dalam data tersebut dihilangkan, bobot setiap fitur dengan menggunakan TF-IDF mampu meningkatkan akurasi dengan menurunkan nilai bobot suatu kata yg sering muncul di dokumen *review,*

dan penggunaan *negation handling* mampu meningkatkan nilai performansi akurasi disebabkan karena suatu kata pada konteks teks yang berbeda dapat dideteksi sebagai suatu kelas teks yang berbeda pula.

4. Kesimpulan

Berdasarkan hasil dan analisis yang telah dilakukan dalam penelitian ini, maka dapat diambil kesimpulan bahwa:

1. *Preprocessing* dengan proses cleaning, tokenization, stopword dan lemmatization dapat meningkatkan akurasi dimana *preprocessing* tersebut dapat mengurangi jumlah *noisy* atau fitur yang tidak memiliki pengaruh terhadap proses klasifikasi. Adapun nilai performansi yang dihasilkan adalah 82,66 %.
2. *Negation handling* memiliki pengaruh terhadap nilai performansi akurasi yang dihasilkan dengan metode multinomial naïve bayes. Hal tersebut dikarenakan suatu kata dalam konteks teks yang berbeda dapat diprediksi polaritasnya.
3. Hasil *negation handling* terbaik yang dihasilkan dari kedua metode *negation handling* yang dirancang adalah *negation handling* berdasarkan *punctuation* sebesar 85,16%. Hal tersebut disebabkan karena suatu kata pada konteks kalimat yang berbeda dengan menggunakan *rules negation handling* berdasarkan *pos tagging* belum akurat dalam menentukan kelas *review* suatu kata pada suatu konteks kalimat.
4. TF-IDF dapat mempengaruhi terhadap nilai performansi akurasi yang dihasilkan pada metode multinomial naïve bayes. Hal tersebut disebabkan karena TF-IDF mampu menurunkan nilai bobot suatu fitur yang sering muncul di dokumen.
5. Hasil nilai performansi akurasi terbaik yang diperoleh adalah ketika menggunakan kombinasi *negation handling* berdasarkan pos tagging dimanai nilai performansi akurasi yang dihasilkan adalah 85.16%.

5. Daftar Pustaka

[1] Bilal, M., Israr, H., & Muhammad Shahid, A. K. (2015). Sentiment classification of Roman-Urdu opinions using Navie Baysian, Decision. *Journal of King Saud University - Computer and Information*.

[2] Chandani, V., Wahono, R. S., & Purwanto. (2015, February 1). Komparasi Algoritma Klasifikasi Machine Learning Dan Feature Selection pada Analisis Sentimen Review Film. *Journal of Intelligent Systems*.

[3] developers, s.-l. (2017, Juni 02). scikit-learn user guide.

[4] Ding, Y. (2013, July). A Quantitative Analysis of Words with Implied Negation in Semantics. *Qingdao University of Science and Technology, China*.

[5] Farooq, U., Mansoor, H., Nongaillard, A., Ouzrout, Y., & Abdul, M. (2016, March 20). Negation Handling in Sentiment Analysis at Sentence Level. *Journal of Computers*.

[6] Foundation, W. (2014, December). *Natural language programming*. Retrieved from wikipedia: https://en.wikipedia.org/wiki/Natural_language_programming

[7] Gaigole, P. C., Patil, L. H., & Chaudhari, P. (2013). Preprocessing Techniques in Text Categorization. *National Conference on Innovative Paradigms in Engineering & Technology (NCIPET-2013).*

[8] Garg, S. K., & Meher, R. K. (2015, May). Naive Bayes model with improved negation handling and n-gram method for Sentiment classification. *National Institute of Technology Rourkela*.

[9] Imtiyazi, M. A., Shaufiah, & Bijaksana, M. A. (n.d.). Sentiment Analysis Berbahasa Indonesia Menggunakan Improved Multinomial Naive Bayes Indonesian Sentiment Analysis Using Improved Multinomial Naïve Bayes. *jurnal_eproc*.

[10] Jurafsky, D., & Martin, J. H. (2016., November 7). Speech and Language Processing.

[11] Sorostinean, M., Sana, K., Mohamed, M., & Targhi, A. (2017, March 1). Sentiment Analysis on Movie Reviews. *otto group product classification challenge*.

[12] Tripathy, A., Agrawal, A., & Rath, S. K. (2016). Classification of Sentiment Reviews using N-gram Machine Learning. *Expert Systems With Applications*.

[13] UP, O. e. (2009, April 1). *Naive Bayes text classification* . Retrieved from nlp stanford: https://nlp.stanford.edu/IR-book/html/htmledition/naive-bayes-text-classification-1.html