

**PENERAPAN TEKNIK PENANGANAN NEGASI
MENGUNAKAN *SYNTACTIC RULE* PADA ANALISIS SENTIMEN
ULASAN HOTEL DI TRAVELOKA DENGAN METODE *NAIVE
BAYES CLASSIFIER***



SKRIPSI

**Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Komputer
pada Departemen Ilmu Komputer / Informatika**

**Disusun Oleh :
Tiara Galuh Prahasiwi
24010314130083**

**DEPARTEMEN ILMU KOMPUTER/ INFORMATIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO**

2018

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini :

Nama : Tiara Galuh Prahasiwi

NIM : 24010314130083

Judul : Penerapan Teknik Penanganan Negasi menggunakan *Syntatic Rule* pada Analisis Sentimen Ulasan Hotel di Traveloka dengan Metode *Naive Bayes Classifier*

Dengan ini saya menyatakan bahwa dalam tugas akhir/ skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang sepengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.

Semarang, 23 Juli 2018



Tiara Galuh Prahasiwi

NIM. 24010314130083

HALAMAN PENGESAHAN

Judul : Penerapan Teknik Penanganan Negasi menggunakan *Syntatic Rule* pada Analisis Sentimen Ulasan Hotel di Traveloka dengan Metode *Naive Bayes Classifier*
Nama : Tiara Galuh Prahasiwi
NIM : 24010314130083

Telah diujikan pada sidang tugas akhir tanggal 23 Juli 2018 dan dinyatakan lulus pada tanggal 23 Juli 2018.


Semarang, 23 Juli 2018

Mengetahui,
Kepala Departemen Ilmu Komputer/ Informatika



Dr. Retno Kusumaningrum, S.Si, M.Kom.
NIP. 198104202005012001

Panitia Penguji Tugas Akhir
Ketua,



Sukmawati Nur E., S.Si, M.Kom
NIP. 197805022005012002


HALAMAN PENGESAHAN

Judul : Penerapan Teknik Penanganan Negasi menggunakan *Syntatic Rule* pada Analisis Sentimen Ulasan Hotel di Traveloka dengan Metode *Naive Bayes Classifier*
Nama : Tiara Galuh Prahasiwi
NIM : 24010314130083

Telah diujikan pada sidang tugas akhir tanggal 23 Juli 2018.

Semarang, 23 Juli 2018

Dosen Pembimbing



Dr. Ketno Kusumaningrum, S.Si, M.Kom.

NIP. 198104202005012001

ABSTRAK

Peningkatan pariwisata di Indonesia menjadikan peluang untuk perusahaan mendirikan usaha layanan pemesanan tiket hotel, salah satunya yaitu Traveloka. Pengguna Traveloka memberikan penilaian dan ulasan mengenai hotel yang dikemas dalam sebuah komentar. Komentar dinilai penting dikarenakan dapat digunakan untuk pengambilan keputusan maupun pengembangan hotel itu sendiri. Namun, analisis komentar menjadi hal yang sulit karena jumlah komentar yang terlalu banyak. Hal ini menyebabkan dibutuhkan suatu teknologi untuk mengelola data komentar yang menunjukkan kepuasan pengguna terhadap suatu topik. Teknologi ini disebut *sentiment analysis*. Dalam penerapannya, klasifikasi yang kurang tepat masih menjadi masalah dikarenakan terdapat beberapa kata yang memiliki polaritas berbeda pada konteks kalimat yang berbeda. Adanya kata negasi dapat membalikkan nilai polaritas dari suatu teks. Hal ini mempengaruhi kinerja dari *sentiment analysis* jika tidak ditangani. Penelitian ini mengajukan penerapan teknik penanganan negasi menggunakan *syntactic rule* pada analisis sentimen ulasan hotel dengan *Naive Bayes Classifier*. Penelitian ini dilakukan untuk mengetahui pengaruh kinerja dari penerapan negasi dibandingkan dengan *sentiment analysis* tanpa penanganan negasi. Teknik penanganan negasi terbukti dapat meningkatkan kinerja *sentiment analysis* dengan mengadaptasi *syntactic rule* dan memberi tag “NEG_” menggunakan *Naive Bayes Classifier*. Hasil dari eksperimen pada penelitian ini menunjukkan peningkatan rata-rata akurasi sebesar 2,2% dari perbandingan *sentiment analysis* yang menerapkan teknik penanganan negasi dengan *sentiment analysis* yang tidak menerapkan teknik penanganan negasi.

Kata kunci : *Sentiment Analysis*, Penanganan Negasi, *Syntactic Rule*, *Naive Bayes*, Hotel, Traveloka

ABSTRACT

Tourism increased in Indonesia makes an opportunity for companies to build service bussiness in hotel ticket reservations, such as Traveloka. Traveloka users present reviews and assesment about the hotel that contains in a comment. Comment is very important because that can be used for decision-making basic and development of the hotel. However, comment analysis becomes difficult because of the number of comments is too much. This issue causes technology needs that used to classify the comment in a topics. This technology is called sentiment analysis. In practice, inappropriate classification are still a problem because there are several words that have different polarities in different sentence contexts. The existence of the negation word can invert the polarity of a text. If it is not handled, it can affects the performance of sentiment analysis. This research proposes the implementation of negation handling techniques using syntatic rule in the sentiment analysis of hotel reviews with Naive Bayes Classifier. This study was conducted to determine the effect of performance of the implementation of negation compared with sentiment analysis without negation handling. Negation handling techniques can improve the performance of sentiment analysis by adapting syntatic rules and tagging "NEG_" using Naive Bayes Classifier. The results of the experiments in this study show that proposed method improves 2,2% average accuracy of the comparison of sentiment analysis with and without the application of negation handling techniques.

Keyword : Sentiment Analysis, Negation Handling, Syntatic Rule, Naive Bayes, Hotel, Traveloka

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT yang telah melimpahkan segala rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Penerapan Teknik Penanganan Negasi menggunakan Syntactic Rule pada Analisis Sentimen Ulasan Hotel di Traveloka dengan Metode *Naive Bayes Classifier*”. Skripsi ini dibuat dengan tujuan sebagai salah satu syarat untuk memperoleh gelar sarjana komputer pada Departemen Ilmu Komputer/Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang.

Dalam pelaksanaan skripsi serta penyusunan dokumen skripsi ini, banyak pihak yang membantu hingga dokumen ini dapat diselesaikan. Untuk itu, melalui kesempatan ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Ibu Dr. Retno Kusumaningrum, S.Si, M.Kom. selaku Kepala Departemen Ilmu Komputer/ Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang dan selaku dosen pembimbing yang telah membantu dalam proses bimbingan hingga selesainya skripsi ini.
2. Bapak Helmie Arif Wibawa, S.Si, M.Cs, selaku Koordinator Skripsi Departemen Ilmu Komputer/Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang.
3. Bapak Bambang Budiarto, Ibu Dwi Mustianingrum, Melyniar Galih Rizkiani, selaku orang tua dan adik yang telah sabar dan selalu memberikan dukungan dan doa dalam menyelesaikan skripsi ini.
4. Keluarga besar, teman-teman, dan semua pihak yang telah membantu kelancaran dalam penyusunan tugas akhir, yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa dokumen skripsi ini masih banyak kekurangan dari segi materi ataupun dalam penyajiannya. Oleh karena itu, saran dan kritik yang membangun sangat penulis harapkan. Akhir kata, semoga skripsi ini dapat bermanfaat bagi semua pihak.

Semarang, 23 Juli 2018

Tiara Galuh Prahasiwi

DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	Error! Bookmark not defined.
HALAMAN PENGESAHAN	Error! Bookmark not defined.
HALAMAN PENGESAHAN	Error! Bookmark not defined.
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI	viii
DAFTAR GAMBAR.....	x
DAFTAR TABEL	xii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan dan Manfaat	4
1.4 Ruang Lingkup	4
1.5 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
2.1 Perkembangan Penelitian Mengenai <i>Sentiment Analysis</i>	7
2.2 Penanganan Negasi	8
2.3 <i>Preprocessing</i>	10
2.3.1 <i>Case Folding</i>	11
2.3.2 Tokenisasi.....	11
2.3.3 <i>Stopword Removal</i>	12
2.3.4 <i>Stemming</i>	12
2.4 <i>K-Fold Cross Validation</i>	20
2.5 <i>Naive Bayes</i>	22
2.6 Evaluasi.....	23
2.7 Pengembangan Perangkat Lunak.....	24

BAB III METODOLOGI PENELITIAN	27
3.1 Pengumpulan Data Komentar pada <i>Website</i> Traveloka.....	28
3.2 <i>Preprocessing</i> tanpa Menerapkan Penanganan negasi	28
3.2.1 <i>Case Folding</i>	29
3.2.2 Tokenisasi.....	30
3.2.3 <i>Stopword Removal</i>	31
3.2.4 <i>Stemming</i>	32
3.3 <i>Preprocessing</i> dengan Menerapkan Penanganan negasi	38
3.4 Pembentukan Model Klasifikasi menggunakan <i>Naive Bayes</i>	45
3.5 Penentuan Model Klasifikasi Terbaik.....	47
3.6 Pengumpulan Data Komentar pada Traveloka secara <i>Real Time</i>	49
3.7 Menganalisis Klasifikasi Sentimen.....	49
3.8 Visualisasi Hasil	51
3.9 Analisis dan Desain Aplikasi.....	52
BAB IV HASIL DAN ANALISIS	61
4.1 Hasil Pengembangan Aplikasi.....	61
4.1.1 Lingkungan Implementasi	61
4.1.2 Implementasi Antarmuka	61
4.2 Skenario Pengujian Aplikasi.....	67
4.2.1 Skenario Pengujian Fungsional Aplikasi.....	67
4.2.2 Skenario Pengujian Kinerja Aplikasi	68
4.3 Hasil dan Analisis Aplikasi.....	70
4.3.1 Hasil dan Analisis Pengujian Fungsional Aplikasi.....	71
4.3.2 Hasil dan Analisis Pengujian Kinerja Aplikasi	71
BAB V PENUTUP	79
5.1 Kesimpulan	79
5.2 Saran	80
DAFTAR PUSTAKA.....	81

DAFTAR GAMBAR

Gambar 2.1 Diagram <i>10-fold cross validation</i>	21
Gambar 2. 2 Model <i>Waterfall</i> (Sommerville, 2011)	25
Gambar 3.1 Gambaran Umum Penelitian	27
Gambar 3.2 <i>Flowchart Preprocessing</i> Tanpa Menerapkan Penanganan negasi	29
Gambar 3.3 <i>Flowchart</i> Tokenisasi	30
Gambar 3.4 <i>Flowchart Stopword Removal</i>	31
Gambar 3.5 Proses <i>Stemming</i>	32
Gambar 3.6 Sub Proses <i>Stemming</i> Sastrawi (Bashri, 2017)	33
Gambar 3.7 Alur Sub Proses <i>Stemming Plural</i> (Bashri, 2017)	34
Gambar 3.8 Alur Sub Proses <i>Stemming Singular</i> (Bashri, 2017).....	35
Gambar 3.9 Alur Sub Proses <i>loopPengembalianAkhir</i> (Bashri, 2017)	37
Gambar 3.10 <i>Flowchart Preprocessing</i> dengan Menerapkan Penanganan Negasi.....	38
Gambar 3.11 <i>Flowchart</i> Penanganan Negasi	39
Gambar 3.12 <i>Flowchart</i> Pembentukan Model Klasifikasi menggunakan <i>Naive Bayes</i> . 46	
Gambar 3.13 <i>Flowchart</i> Pelatihan menggunakan <i>Naive Bayes</i>	47
Gambar 3.14 <i>Flowchart</i> Proses Menganalisis Klasifikasi Sentimen	50
Gambar 3.15 <i>Flowchart</i> Prediksi Kelas menggunakan <i>Naive Bayes</i>	50
Gambar 3.16 <i>Data Context Diagram</i> (DCD)	54
Gambar 3.17 <i>Data Flow Diagram</i> (DFD) level 1	56
Gambar 3. 18 Desain Antarmuka Halaman Utama	57
Gambar 3.19 Desain Antarmuka Halaman Masukan URL Hotel	57
Gambar 3.20 Desain Antarmuka Halaman Hasil Klasifikasi Sentimen Data Baru.....	58
Gambar 3.21 Desain Antarmuka Halaman Data Komentar	58
Gambar 3.22 Desain Antarmuka Halaman Hasil Penanganan Negasi.....	59
Gambar 3.23 Desain Antarmuka Halaman Hasil <i>Preprocessing</i>	59
Gambar 3.24 Desain Antarmuka Halaman Hasil Pembentukan Model	60
Gambar 4.1 Implementasi Antarmuka Halaman Utama	62
Gambar 4.2 Implementasi Antarmuka Halaman Masukan URL Hotel	63
Gambar 4.3 Implementasi Antarmuka Halaman Hasil Klasifikasi Sentimen Data Baru	63
Gambar 4.4 Implementasi Antarmuka Halaman Data Komentar.....	64

Gambar 4.5 Implementasi Antarmuka Halaman Hasil Penanganan Negasi	65
Gambar 4.6 Implementasi Antarmuka Halaman Hasil <i>Preprocessing</i>	66
Gambar 4.7 Implementasi Antarmuka Halaman Hasil Pembentukan Model	66
Gambar 4.8 Pembagian Data Untuk Setiap Kelas Sentimen.....	69
Gambar 4.9 Grafik Nilai Akurasi Skenario 1	72
Gambar 4.10 Grafik Nilai Akurasi Skenario 2	75
Gambar 4.11 Perbandingan Akurasi Skenario 1 dan Skenario 2	76

DAFTAR TABEL

Tabel 2.1 Penelitian Lain Mengenai <i>Sentiment Analysis</i>	8
Tabel 2.2 Contoh Hasil Proses Penanganan Negasi Menggunakan <i>Punctuation Mark</i> ...	9
Tabel 2.3 Contoh Hasil Proses Penanganan Negasi Menggunakan <i>Part Of Speech</i>	10
Tabel 2.4 Contoh Hasil Proses Tokenisasi	11
Tabel 2.5 Contoh Hasil Proses <i>Stopword Removal</i>	12
Tabel 2.6 Aturan Pemenggalan Awalan <i>Stemmer</i> Nazief-Adriani.....	15
Tabel 2.7 Modifikasi Tambahan Aturan pada Tabel 2.6 oleh Algoritma CS <i>Stemmer</i> ..	17
Tabel 2.8 Daftar Aturan <i>rulePrecedence</i>	17
Tabel 2.9 Modifikasi Aturan pada Tabel 2.6 oleh Algoritma ECS <i>Stemmer</i>	19
Tabel 2.10 Modifikasi Aturan Pemenggalan Awalan dan Modifikasi Penambahan Aturan Pemenggalan Sisipan oleh Algoritma <i>Stemming Modified</i> ECS....	20
Tabel 2.11 Penambahan dan Modifikasi Aturan Pemenggalan Awalan <i>Stemmer</i> Sastrawi	20
Tabel 2.12 Contoh Confusion Matrix.....	24
Tabel 3.1 Contoh Hasil Proses <i>Case Folding</i> pada <i>Dataset</i>	29
Tabel 3.2 Contoh Hasil Proses Tokenisasi	30
Tabel 3.3 Ilustrasi Proses <i>Stopword Removal</i>	32
Tabel 3.4 Contoh Hasil <i>Stemming Singular</i>	34
Tabel 3.5 Contoh Hasil Proses <i>Stemming</i>	37
Tabel 3.6 Daftar Kelas Label pada Korpus (Dinakaramani et al., 2014)	40
Tabel 3.7 Kosakata yang Ditambahkan pada Korpus beserta <i>Tag</i> -nya.....	42
Tabel 3.8 Rule Penanganan Negasi (Amalia et al., 2018).....	44
Tabel 3.9 Rule Penanganan Negasi yang Diusulkan.....	44
Tabel 3.10 Contoh hasil pengindeksan <i>10-Fold Cross-Validation</i>	45
Tabel 3.11 Perbandingan Kelas Prediksi dengan Kelas Sebenarnya.....	48
Tabel 3.12 <i>Confussion Matrix</i> dari data Tabel 3.11	48
Tabel 3.13 Perhitungan Persentase Data Tiap Kategori.....	51
Tabel 3.14 Kebutuhan Fungsional Penerapan Teknik Penanganan Negasi	53
Tabel 3.15 Kebutuhan Non Fungsional Penerapan Teknik Penanganan Negasi	53
Tabel 4.1 Skenario Pengujian Fungsional Penerapan Teknik Penanganan Negasi.....	67

Tabel 4.2 Nilai Akurasi dari Hasil Skenario 1.....	71
Tabel 4.3 Nilai Akurasi dari Hasil Skenario 2.....	73
Tabel 4.4 Perbandingan Hasil Skenario 1 dan Skenario 2	75
Tabel 4.5 Perbandingan Rata-Rata Akurasi dan Akurasi Tertinggi	77

DAFTAR LAMPIRAN

Lampiran 1. Perhitungan Pelatihan dan Prediksi Model <i>Multinomial Naive Bayes</i>	85
Lampiran 2. Deskripsi dan Hasil Pengujian Fungsional Aplikasi.....	92

BAB I

PENDAHULUAN

Bab ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, dan sistematika penulisan skripsi mengenai Penerapan Teknik Penanganan Negasi menggunakan Syntactic Rule pada Analisis Sentimen Ulasan Hotel di Traveloka dengan Metode *Naive Bayes Classifier*.

1.1 Latar Belakang

Pariwisata Indonesia merupakan salah satu sektor yang menjadi fokus pemerintah Indonesia untuk dikembangkan. Meningkatnya pariwisata di Indonesia akan berdampak pada semakin banyak jumlah hotel di Indonesia. Hotel merupakan sarana akomodasi umum untuk tempat menginap tamu dari dalam negeri maupun luar negeri. Akhir-akhir ini hotel tidak hanya digunakan untuk tempat menginap, akan tetapi hotel juga digunakan sebagai tempat pertemuan bisnis, seminar, musyawarah nasional dan kegiatan lainnya. Hal tersebut membuat pengguna hotel akan semakin meningkat. Kondisi ini dibuktikan dengan adanya data statistik yang menyebutkan tingkat penghunian kamar hotel pada tahun 2015 mencapai 53,92% untuk kamar hotel berbintang dan 33,21% untuk hotel nonbintang dan akomodasi lainnya, sedangkan pada tahun 2016 mencapai 54,29% untuk kamar hotel berbintang dan 34,85% untuk hotel nonbintang dan akomodasi lainnya (Badan Pusat Statistik, 2018).

Jumlah pengguna hotel yang meningkat dijadikan beberapa perusahaan untuk mendirikan usaha layanan pemesanan tiket hotel salah satunya yaitu Traveloka. Perusahaan traveloka ini memiliki aplikasi dan *website* www.traveloka.com yang berhasil menjangkau 15 juta pengunduh aplikasinya dengan mayoritasnya pengguna aktif (Herman, 2017). Pengguna dapat memesan dan memberi penilaian atau ulasan mengenai hotel pada aplikasi dan *website* tersebut. Ulasan yang diberikan dikemas menjadi bentuk komentar. Komentar yang ada dapat membantu pengguna memilih hotel yang sesuai dan juga membantu pemilik hotel untuk mengembangkan usahanya. Komentar yang positif akan memberikan peluang meningkatnya pengguna hotel, sedangkan komentar negatif dapat dijadikan referensi atau acuan bagi pemilik

hotel untuk pengembangan dan peningkatan hotel. Namun, terkadang pengguna Traveloka mengalami kesulitan untuk menganalisis komentar dikarenakan terlalu banyak jumlah komentar yang ada. Maka dari itu, diperlukan suatu teknologi pengolahan data dan analisis terhadap komentar.

Sentiment analysis merupakan salah satu bentuk teknologi yang mengelola data komentar untuk menunjukkan kepuasan dan sikap pengguna terhadap suatu topik yang dalam hal ini adalah hotel. *Sentiment analysis* atau disebut juga dengan *opinion mining* merupakan bidang studi yang menganalisa pendapat setiap orang, sentimen, evaluasi, penilaian, sikap, dan perasaan terhadap entitas seperti produk, jasa organisasi, individu, masalah, peristiwa, topik, dan atribut-atributnya (Liu, 2012). Penelitian mengenai *sentiment analysis* sudah banyak dilakukan dengan berbagai metode yang digunakan. Kinerja yang dicapai oleh metode-metode ini sulit untuk dijadikan bahan penilaian metode klasifikasi terbaik. Hal ini dikarenakan setiap metode menggunakan berbagai sumber daya pelatihan, koleksi dokumen untuk pengujian yang berbeda, berbagai metode pemilihan fitur, dan pengolahan teks yang berbeda pula. Klasifikasi data dapat dilakukan dengan berbagai metode *machine learning* yang telah mencapai sukses besar dalam kategorisasi teks, antara lain *Naive Bayes*, *Maximum Entropy* dan *Support Vector Machines* (Vinodhini & Chandrasekaran, 2012).

Pada penerapan *sentiment analysis*, tugas utama yang harus dilakukan yaitu mengekstrak opini, penilaian dan ekspresi sentimen untuk diklasifikasikan ke kelas positif maupun negatif. Dalam pengekstrakan sentimen, suatu kata memiliki polaritas yang berbeda pada konteks kalimat yang berbeda pada keadaan tertentu. Hal ini dapat dicontohkan seperti pada kalimat “Hotel itu bersih”, kata “bersih” akan berbeda polaritas dengan kalimat “Hotel itu tidak bersih”. Hal itu dikarenakan adanya kata negasi dalam kalimat tersebut. Maka dari itu perlu dilakukan penanganan negasi untuk mengurangi kesalahan dalam pengklasifikasian komentar.

Negasi ditangani dengan menentukan terlebih dahulu kata-kata yang terpengaruh dari kata negasi tersebut yang biasanya disebut ruang lingkup negasi. Kebanyakan dari sistem *sentiment analysis* yang telah dilakukan, negasi ditangani menggunakan metode tradisional menggunakan *static window* dan tanda baca (*punctuation mark*) untuk menentukan ruang lingkup dari negasi (Farooq et al., 2016). Semua kata setelah kata negasi (seperti no, not, n't) akan diberi tag “not_” +

kata tersebut hingga ditemukan tanda baca atau kata negasi lain (Narayanan et al., 2013). Akan tetapi, metode ini dirasa kurang efektif dikarenakan jumlah fitur kata yang tidak dibutuhkan akan semakin bertambah. Hal tersebut dapat diatasi dengan menerapkan metode penentuan ruang lingkup negasi menggunakan *Part Of Speech* (POS). POS mengklasifikasikan setiap kata menjadi kategorinya masing-masing seperti kata benda, kata sifat, kata kerja, dan lain-lain. Setelah kata negasi ditemukan, maka akan berhenti jika ditemukan kata dengan label kategori tertentu. Metode penanganan negasi dengan POS berhasil menaikkan akurasi klasifikasi pada *sentiment analysis* (Garg & Meher, 2015). Beberapa penelitian mengenai penanganan negasi dengan POS telah membuktikan adanya peningkatan akurasi *sentiment analysis* dalam dokumen bahasa Inggris (Farooq et al., 2016; Garg & Meher, 2015; Wisudawati et al., 2017).

Penentuan ruang lingkup negasi menggunakan POS juga dapat diterapkan pada dokumen bahasa Indonesia. Hal ini dilakukan dengan menyesuaikan struktur kata pengungkap negasi pada bahasa Indonesia seperti tidak, jangan, bukan, belum, dan lain-lain yang termasuk bentuk negasi *syntactic*. Penanganan negasi *syntactic* dilakukan dengan pendekatan *rule-based* dimana memperhatikan struktur kata negasi yang ada dimana penanganan tersebut biasa disebut dengan *syntactic rule*. Penanganan dengan cara ini dalam dokumen bahasa Indonesia juga telah dilakukan (Amalia et al., 2018). Data yang digunakan dalam penelitian ini yaitu data Twitter. Akan tetapi, penelitian tersebut masih memiliki kekurangan yaitu *rule* yang digunakan kurang lengkap dan belum ada penanganan untuk kata negasi yang tidak baku.

Oleh karena itu, pada penelitian kali ini diusulkan *sentiment analysis* yang menerapkan penanganan negasi menggunakan *syntactic rule* dalam komentar bahasa Indonesia pada Traveloka dengan beberapa modifikasi *rule*. Ruang lingkup negasi ditentukan dengan POS. Sedangkan metode klasifikasi yang digunakan adalah *Naive Bayes*. *Naive Bayes* memiliki beberapa keunggulan yaitu sederhana dan memiliki asumsi yang kuat (Devika et al., 2016). Selain itu *Naive Bayes* hanya membutuhkan sejumlah kecil data pelatihan untuk memperkirakan parameter yang diperlukan untuk klasifikasi. *Naive Bayes* telah terbukti bekerja dengan memuaskan di banyak domain (Dey et al., 2016).

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, dapat dirumuskan permasalahan yaitu bagaimana pengaruh penanganan negasi menggunakan *syntactic rule* pada *sentiment analysis* ulasan hotel dengan menggunakan metode *Naive Bayes Classifier*.

1.3 Tujuan dan Manfaat

Tujuan yang ingin dicapai dari penelitian ini adalah mengetahui pengaruh kinerja dari penerapan penanganan negasi menggunakan *syntactic rule* pada *sentiment analysis* ulasan hotel di Traveloka dengan metode *Naive Bayes Classifier* dibandingkan dengan *sentiment analysis* tanpa penanganan negasi. Oleh karena itu, tujuan tersebut dapat didetailkan menjadi berbagai tujuan khusus sebagai berikut :

1. Mengetahui kinerja *sentiment analysis* dengan menerapkan penanganan negasi pada ulasan hotel
2. Mengetahui kinerja *sentiment analysis* tanpa menerapkan penanganan negasi pada ulasan hotel
3. Membandingkan hasil kinerja *sentiment analysis* dengan dan tanpa menerapkan penanganan negasi pada ulasan hotel

Manfaat yang diharapkan dari penelitian ini yaitu membantu dalam mengetahui opini publik terhadap suatu hotel bagi pengguna hotel maupun pemilik hotel melalui komentar pada *website* Traveloka. Secara khusus, manfaat untuk pengguna hotel yaitu dapat menjadi referensi atau acuan untuk memilih hotel, sedangkan manfaat untuk pemilik hotel yaitu sebagai bahan untuk pengembangan dan peningkatan hotel.

1.4 Ruang Lingkup

Ruang lingkup dalam menerapkan penanganan negasi menggunakan *syntactic rule* pada *sentiment analysis* dengan metode *Naive Bayes Classifier* adalah :

1. Objek penelitian yang digunakan pada penelitian ini adalah beberapa hotel di Indonesia yang diambil secara acak.
2. Data komentar yang digunakan untuk pelatihan pada penelitian ini berjumlah 1000 komentar dengan jumlah yang seimbang yaitu 500 data positif dan 500 data negatif yang diambil dengan proses *scraping* pada *website* Traveloka.
3. Pelabelan *dataset* penelitian dilakukan secara manual.
4. Data komentar yang digunakan adalah data komentar hotel berbahasa Indonesia.

5. Data komentar yang digunakan pada proses *sentiment analysis* secara *real time* diambil pada *range* waktu 1 bulan terakhir dari tanggal dilakukan klasifikasi secara *real time*.

1.5 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam penelitian ini terbagi dalam beberapa pokok bahasan, yaitu :

BAB I PENDAHULUAN

Bab ini memberikan gambaran mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup serta sistematika penulisan skripsi mengenai Penerapan Teknik Penanganan Negasi menggunakan Syntatic Rule pada Analisis Sentimen Ulasan Hotel di Traveloka Dengan Metode *Naive Bayes Classifier*.

BAB II TINJAUAN PUSTAKA

Bab ini memberikan kajian pustaka yang berhubungan dengan tema skripsi sebagai landasan untuk perumusan dan analisis permasalahan pada skripsi. Kajian pustaka yang digunakan meliputi perkembangan penelitian mengenai *sentiment analysis*, penanganan negasi, *preprocessing*, *K-fold Cross Validation*, *Naive Bayes*, evaluasi dan pengembangan perangkat lunak.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan mengenai tahapan dalam penyelesaian masalah skripsi. Tahapan tersebut meliputi pengumpulan data komentar pada *website* Traveloka, *preprocessing* tanpa menerapkan penanganan negasi, *preprocessing* dengan menerapkan penanganan negasi, pembentukan model klasifikasi menggunakan *Naive Bayes*, penentuan model klasifikasi terbaik, pengumpulan data komentar pada Traveloka secara *real time*, menganalisis klasifikasi sentimen, dan visualisasi hasil.

BAB IV HASIL DAN ANALISIS

Bab ini menguraikan hasil penelitian dan analisis pada penelitian yang telah dilakukan dimulai dari teknik pengumpulan data sampai hasil dan analisa.

BAB V PENUTUP

Bab ini menjelaskan kesimpulan dari uraian yang telah dijabarkan pada bab-bab sebelumnya dan saran untuk pengembangan penelitian lebih lanjut.