

**ANALISIS SENTIMEN KEPUASAN PELANGGAN ONLINE SHOP
MENGUNAKAN ALGORITMA NAÏVE BAYES CLASSIFIER DAN K-NEAREST
NEIGHBOR (Studi Kasus : Twitter Resmi ShopeeCare)**

SKRIPSI

Untuk memenuhi sebagian persyaratan

mencapai derajat Sarjana S-1

Program Studi Teknik Informatika



Diajukan Oleh

INARATUL MAHBUBAH

STATE ISLAMIC UNIVERSITY
16650069
SUNAN KALIJAGA
YOGYAKARTA
PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2020

Analisis Sentimen Kepuasan Pelanggan Online Shop Menggunakan *Algoritman Naïve Bayes Classifier* dan *K-Nearest Neighbor* (Studi Kasus : Twitter Resmi Shopee)

Inaratul Mahbubah

16650069

INTISARI

Shopee adalah salah satu onlineshop yang sering digunakan masyarakat di Indonesia. Saat ini masyarakat di Indonesia bahkan dunia lebih senang berbelanja barang-barang yang mereka inginkan menggunakan onlineshop. Hal tersebut terjadi dikarenakan masyarakat menilai menggunakan online shop akan menjadi lebih mudah dan praktis. Selain itu, ada juga karena barang yang dibeli lebih murah dan banyaknya promo yang diberikan oleh toko online.

Dalam penelitian ini dilakukan analisis sentimen dengan menggunakan data latih sebanyak 1015 data tweet dan data testing sebanyak 480 data tweet. Data tersebut diuji menggunakan metode *Naïve Bayes Classifier* dan *K-NN* menggunakan *Confussion Matrix*. Data yang diuji menggunakan *Naïve Bayes Classifier* menggunakan *confusion matrix* menghasilkan nilai akurasi sebanyak 86,27% dengan rasio *split* 90:10, nilai akurasi sebanyak 88,17% dengan rasio *split* 80:20, dan nilai akurasi 85,90% dengan rasio *split* 70:30. Sedangkan pada data yang diproses menggunakan *K-NN* menggunakan *confussion matrix* menghasilkan nilai akurasi sebanyak 84,31% dengan rasio *split* 90:10, nilai akurasi sebanyak 87,68% dengan rasio *split* 80:20, dan nilai akurasi 87,21% dengan rasio *split* 70:30. Hasil implementasi pada data tweet baru sejumlah 480 yang didapatkan pada *Naïve Bayes Classifier* adalah 25,6% data positif dan 74,4% data negative. Sedangkan pada implementasi *K-NN* adalah 40,2% data positif dan 59,8% data negative.

Kata Kunci : Analisis sentiment, *Naïve Bayes Classifier*, *K-NN* (*K-Nearest Neighbor*)

Analisis Sentimen Kepuasan Pelanggan Online Shop Menggunakan *Algoritman Naïve Bayes Classifier* dan *K-Nearest Neighbor* (Studi Kasus : Twitter Resmi ShopeeCare

Inaratul Mahbubah

16650069

ABSTRACT

Shopee is one of the mostly used online marketplace in Indonesia. People around the world are most likely to buy products through online marketplace nowadays, including Indonesians. It happens because the citizens find online shopping is easier and more practical. Other than that, some people think the products are cheaper and the availability of many discounted prices given by the online stores.

In this research, sentiment analysis was carried out using 1015 tweets as training data and 480 tweets as testing data. The data was tested using the Naïve Byes Classifier method and K-NN using the confusion matrix. Data tested using the Naïve Bayes Classifier using confusion matrix obtains an accuracy rate of 86.27% with a split ratio of 90:10, an accuracy rate of 88.17% with a split ratio of 80:20, and an accuracy rate of 85.90% with a split ratio of 70:30. Whilst the data processed using K-NN using a configuration matrix obtains an accuracy rate of 84.31% with a split ratio of 90:10, an accuracy rate of 87.68% with a split ratio of 80:20, and an accuracy rate of 87.21% with a ratio split 70:30. The results of the implementation of the 480 new data tweets obtained on the Naïve Bayes Classifier are 25.6% positive data and 74.4% negative data. Meanwhile, the implementation of K-NN are 40.2% positive data and 59.8% negative data.

Keywords : *Sentiment analysis, Naïve Bayes Classifier, K-NN (K-Nearest Neighbor)*

SURAT PERNYATAAN KEASLIAN SKRIPSI

PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan dibawah ini :

Nama : INARATUL MAHBUBAH

NIM : 16650069

Jurusan : Teknik Informatika

Fakultas : Sains dan Teknologi

Menyatakan bahwa skripsi saya yang berjudul "**Analisis Sentimen Kepuasan Pelanggan Online Shop Menggunakan Algoritman Naïve Bayes Classifier dan K-Nearest Neighbor (Studi Kasus : Twitter Resmi ShopeeCare)**" merupakan hasil penelitian saya sendiri, tidak terdapat karya yang pernah di ajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi dan bukan plagiasi karya orang lain kecuali yang secara tertulis diacu dalam naskah ini disebutkan dalam daftar pustaka.

Yogyakarta, 9 Desember 2020



Inaratul Mahbubah

NIM. 1665006

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

SURAT PERSETUJUAN SKRIPSI



Universitas Islam Negeri Sunan Kalijaga



FM-UINSK-BM-05-03/R0

SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi

Lamp :

Kepada
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga Yogyakarta
di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Inaratul Mahbubah
NIM : 16650069
Judul Skripsi : "Analisis Sentimen Kepuasan Pelanggan Online Shop Menggunkan *Algoritman Naïve Bayes Classifier* dan *K-Nearest Neighbor* (Studi Kasus : Twitter Resmi ShopeeCare)"

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Teknik Informatika

Dengan ini kami mengharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 15 Desember 2020

Pembimbing

M. Didik Rohmad Wahyudi, S.T., M.T.

NIP. 19901113 201903 1 012

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PENGESAHAN TUGAS AKHIR



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-2902/Un.02/DST/PP.00.9/12/2020

Tugas Akhir dengan judul : ANALISIS SENTIMEN KEPUASAN PELANGGAN ONLINE SHOP MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER DAN K - NEAREST NEIGHBOR (STUDI KASUS : TWITTER RESMI SHOPEECARE)

yang dipersiapkan dan disusun oleh:

Nama : INARATUL MAHBUBAH
Nomor Induk Mahasiswa : I6650069
Telah diujikan pada : Selasa, 22 Desember 2020
Nilai ujian Tugas Akhir : A-

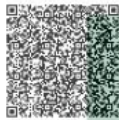
dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang
Muhammad Didik Rohmad Wahyudi, S.T., MT.
SIGNED

Valid ID: 5fedf03e75cc



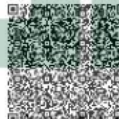
Penguji I
Dr. Shofwatul 'Uyun, S.T., M.Kom.
SIGNED

Valid ID: 5feb7305068e



Penguji II
Nurochman, S.Kom., M.Kom
SIGNED

Valid ID: 5feb915c1aa5d



Yogyakarta, 22 Desember 2020
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi
Dr. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 5fed9070718b

SUNAN KALIJAGA UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan penelitian yang berjudul “Analisis Sentimen Kepuasan Pelanggan Online Shop Menggunakan *Algoritman Naïve Bayes Classifier* dan *K-Nearest Neighbor* (Studi Kasus : Twitter Resmi Shopeecare)” sebagai salah satu syarat untuk mendapatkan gelar sarjana program studi Teknik Informatika Universitas Islam Negeri Sunan Kalijaga Yogyakarta. Sholawat serta salam selalu tercurahkan kepada Rasulullah SAW, keluarga, sahabat, dan para pengikutnya.

Penulis menyadari bahwa dalam proses penulisan penelitian ini masih jauh dari kara sempurna maka dari itu diharapkan ada kritik dan saran untuk memperbaiki penelitian yang akan datang.

Tidak lupa penulis juga mengucapkan terimakasih kepada seluruh pihak yang telah membantu dalam penyelesaian skripsi ini. Ucapan terimakasih penulis sampaikan kepada :

1. Allah SWT yang telah memberikan kekuatan dan kelancaran.
2. Bapak Prof. Dr. Phil. Al Makin, S.Ag., M.A. selaku Rektor UIN Sunan Kalijaga Yogyakarta.
3. Ibu Maria Ulfah Siregar, S.Kom. MIT., Ph.D selaku Kepala Program Teknik Informatika.
4. Bapak Muhammad Didik Rohmad Wahyudi, S.T., M.T. Selaku dosen pembimbing skripsi dan dosen pembimbing akademik yang telah sabar membimbing , memberikan arahan, dan motivasi dalam penyusunan skripsi.

5. Bapak Agus Mulyanto, S.Si., M.Kom., Ph.D., Bapak Aulia Faqih Rifa'i, M.Kom., Bapak Dr. Agung Fatwanto, S.Si., M.Kom, Bapak Nurochman, S.Kom., Bapak M. Taufiq Nuruzzaman, S.T., Bapak Rahmat Hidayat, S.Kom., M.Cs., Ibu Dr. Shofwatul 'Uyun, S.T., M.Kom., selaku dosen pengampu mata kuliah program studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta.
6. Seluruh staff dan karyawan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
7. Kedua orang tua, Bapak Widodo Budiharto dan Alm. Ibu Siti Mukhibah yang telah memberikan dukungan, semangat, dan do'a sehingga penulis dapat menyelesaikan penelitian ini.
8. Adikku Basyar Mahfuzh Suwailim
9. Keluarga besar kedua orang tua.
10. Teman-teman yang telah membantu dan bertukar pikir Alfi, Rifqa, Lia, Alfia, Titik, Yasmin, Aat, Nur, The Kntnk.
11. Teman-teman yang telah berpartisipasi dalam prosesn penelitian Alfi, Rifqa, Yasmin.
12. Teman-teman angkatan 2016 semuanya yang telah berproses bersama
13. Teman-teman KKN Selo Barat, Kokap Angkatan 99 yangtelah menjadi keluarga lebih dari 2 bulan lamanya.
14. Serta seluruh pihak yang tidak dapat disebutkan satu persatu dan telah memberikan banyak do'a dan dukungan sehingga penelitian ini dapat terselesaikan.

Yogyakarta, 9 Desember 2020

Penulis

HALAMAN PERSEMBAHAN

Skripsi ini dipersembahkan kepada orangtua, adik, keluarga besar yang selalu memberikan do'a dan dukungan.

Serta teman-teman yang selalu memberikan motivasi dan arahan sehingga penulis dapat belajar dari banyak hal



HALAMAN MOTTO

“Sesungguhnya sesudah kesulitan itu ada kemudahan, sesungguhnya sesudah kesulitan itu ada kemudahan.”

(Q.S. Al Insyirah ayat 5-6)



DAFTAR ISI

INTISARI	i
ABSTRACT	ii
SURAT PERNYATAAN KEASLIAN SKRIPSI	iii
SURAT PERSETUJUAN SKRIPSI	iv
PENGESAHAN TUGAS AKHIR	v
KATA PENGANTAR	vi
HALAMAN PERSEMBAHAN	viii
HALAMAN MOTTO	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xv
DAFTAR RUMUS	xvii
BAB I PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	3
1.6. Keaslian Penelitian.....	4

1.7.	Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....		7
2.1.	Tinjauan Pustaka.....	7
2.2.	Landasan Teori.....	14
2.2.1.	Text Mining.....	14
2.2.2.	Text Preprocessing.....	14
2.2.3.	Analisis Sentimen.....	16
2.2.4.	Python.....	17
2.2.5.	Twitter.....	17
2.2.6.	Klasifikasi.....	17
2.2.7.	Naïve Bayes Classifier.....	17
2.2.8.	K-Nearest Neighbors.....	20
2.2.9.	TF-IDF.....	22
2.2.10.	Split Validation.....	23
2.2.11.	Confusion Matrix.....	23
2.2.12.	Shopee.....	24
BAB III METODE PENELITIAN		27
3.1.	Metode Penelitian.....	27

3.2.	Alur Penelitian.....	27
3.2.1.	Studi Pustaka.....	28
3.2.2.	Pengumpulan Data.....	28
3.2.3.	Seleksi Data dan Pelabelan Data.....	28
3.2.4.	Preprocessing Data.....	29
3.2.5.	Analisa.....	29
3.2.6.	Implementasi.....	30
3.2.7.	Perangkat Penelitian.....	30
BAB IV HASIL DAN PEMBAHASAN.....		31
4.1.	Pengumpulan Data.....	32
4.2.	Seleksi Data dan Pelabelan Data.....	32
4.3.	Preprocessing Data.....	33
4.3.1.	Cleansing.....	33
4.3.2.	Casefolding.....	34
4.3.3.	Tokenizing dan Convert Slangword.....	35
4.3.4.	Stopword Removal.....	36
4.3.5.	Stemming.....	36
4.4.	Pembobotan Kata (Term Weighthing).....	37

4.4.1. Term Frequency – Invers Document Frequency (TF-IDF).....	37
4.5. Algoritma Naïve Bayes.....	39
4.6. Algoritma K-Nearest Neighbor.....	44
4.7. Evaluasi Model Klasifikasi Naïve Bayes Classifier.....	45
4.8. Evaluasi Model Klasifikasi K-Nearest Neighbor.....	49
4.9. Implementasi Naïve Bayes Classifier.....	52
4.10. Implementasi K-Nearest Neighbor.....	54
BAB V PENUTUP.....	57
5.1. Kesimpulan.....	57
5.2. Saran.....	58
DAFTAR PUSTAKA.....	59
LAMPIRAN.....	61
CURICULUM VITAE.....	75

DAFTAR GAMBAR

Gambar 3.2. 1 Skema Alur Penelitian.....	27
Gambar 4. 1 Flowchart Penelitian.....	31
Gambar 4.9. 1 Grafik Presentase Kelas Hasil Implementasi Naïve Bayes Classiffier.....	53
Gambar 4.9. 2 Visualisasi Data Uji Berlabel Positif dengan Wordcloud NBC.....	53
Gambar 4.9. 3 Visualisasi Data Uji Berlabel Negativ dengan Wordcloud NBC.....	54
Gambar 4.10. 1 Grafik Presentase Kelas Hasil Implementasi K-Nearest Neighbor	55
Gambar 4.10. 2 Visualisasi Data Uji Berlabel Positif dengan Wordcloud K-NN.....	55
Gambar 4.10. 3 Visualisasi Data Uji Berlabel Negativ dengan Wordcloud K-NN.....	56



DAFTAR TABEL

Tabel 2. 1 Tinjauan Putaka	10
Tabel 2. 2 Tinjauan Pustaka (lanjutan).....	11
Tabel 2. 3 Tinjauan Pustaka (lanjutan).....	12
Tabel 2. 4 Tinjauan Pustaka (lanjutan).....	13
Tabel 2.2.11. 1 Confusion Matrix.....	23
Tabel 4. 1 Tabel Contoh Pengumpulan Data Tweet Shopeecare	32
Tabel 4. 2 Tabel Contoh Pengumpulan Data Tweet Yang Sudah Dilabeli Dan Diseleksi	33
Tabel 4.3. 1 Tabel Contoh Hasil Cleansing.....	34
Tabel 4.3. 2 Tabel Contoh Hasil Casefolding	34
Tabel 4.3. 3 Table Contoh Hasil Tokenizing dan Convert Slangword	35
Tabel 4.3. 4 Tabel Contoh Hasil Stopword Remover.....	36
Tabel 4.3. 5 Tabel Contoh Hasil Stemming	37
Tabel 4.4. 1 Contoh Perhitungan TF-IDF	38
Tabel 4.5. 1 Contoh Perhitungan TF-IDF (Lanjutan).....	41
Tabel 4.5. 2 Contoh Untuk Perhitungan Laplace Smoothing.....	41
Tabel 4.5. 3 Tabel contoh perhitungan klasifikasi kelas negatif	43
Tabel 4.5. 4 Tabel Contoh Perhitungan Klasifikasi Kelas Positif	44
Tabel 4.7. 1 Pembagian split data pengujian Naïve Bayes Classifier.....	46
Tabel 4.7. 2 Tabel Confussion Matrix Percobaan pertama Naive Bayes Classifier	46
Tabel 4.7. 3 Tabel Confussion Matrix Percobaan Kedua Naive Bayes Classifier	47
Tabel 4.7. 4 Tabel Confussion Matrix Percobaan ketiga Naive Bayes Classifier	48

Tabel 4.8. 1 Tabel Pembagian Split data pengujian K-Nearest Neighbor49

Tabel 4.8. 2 Tabel Confussion Matrix Percobaan pertama K-Nearest Neighbor50

Tabel 4.8. 3 Tabel Confussion Matrix Percobaan kedua K-Nearest Neighbor51

Tabel 4.8. 4 Tabel Confussion Matrix Percobaan ketiga K-Nearest Neighbor52



DAFTAR RUMUS

Persamaan 1 Probabilitas 1.....	18
Persamaan 2 Prior Probability.....	19
Persamaan 3 Conditional Probability 1.....	19
Persamaan 4 Conditional Probabilities 2.....	19
Persamaan 5 Euclidean Distance.....	21
Persamaan 6 Menghitung Akurasi Nilai k.....	22
Persamaan 7 Pembobotan TF-IDF.....	22
Persamaan 8 Nilai Tingkat Akurasi Confussion Matrix.....	24



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

BAB I

PENDAHULUAN

1.1. Latar Belakang

Saat ini masyarakat di Indonesia bahkan dunia lebih senang berbelanja barang-barang yang mereka inginkan menggunakan onlineshop. Salah satu onlineshop yang sering digunakan oleh masyarakat Indonesia adalah *shopee*. Hal tersebut terjadi dikarenakan masyarakat menilai menggunakan online shop akan menjadi lebih mudah dan praktis serta tidak membuang-buang waktu untuk pergi ke toko secara langsung. Selain itu, ada juga faktor yang mempengaruhi hal tersebut adalah barang yang dibeli pelanggan lebih murah karena banyaknya promo yang diberikan oleh toko online. Sehingga akan lebih banyak menarik pelanggan online shop untuk mengunjungi toko online tersebut.

Media social adalah sebuah media untuk bersosialisasi satu sama lain dan dilakukan secara online yang memungkinkan masyarakat untuk saling berinteraksi satu sama lain tanpa adanya batasan ruang dan waktu. Dengan adanya media social masyarakat akan dipermudah untuk banyak hal, seperti mengirim pesan singkat, bertukar pendapat, berkomentar, menjalin pertemanan, berkirin foto atau video, dan lain sebagainya. Salah satu media social yang banyak digunakan oleh masyarakat adalah Twitter.

Twitter adalah salah satu aplikasi yang cukup banyak digunakan masyarakat di Indonesia Twitter sering digunakan oleh masyarakat untuk mempublikasikan kegiatan yang dilakukan oleh pengguna dalam kehidupan sehari-hari. Biasanya masyarakat mempublikasikannya dalam bentuk teks, foto, atau video.

Penelitian ini menggunakan Twitter sebagai media untuk menganalisis kepuasan pelanggan online shop melalui komentar yang ada di akun twitter resmi “Shopeecare.

Dalam proses menganalisis kepuasan pelanggan online shop dibutuhkan metodologi yang tepat untuk mendapatkan hasil yang akurat. Data yang di analisis adalah tweet berupa komentar-komentar para pelanggan online shop yang kemudian diklasifikasi menjadi dua kelas sentiment, diantaranya positif dan negative. Metode yang digunakan dalam penelitian ini adalah *Naïve Bayes Classifier* dan *K-Nearest Neighbor*. Kedua metode ini dipilih karena sederhana dan memberikan kemudahan dalam proses pengolahan data serta memberikan tingkat akurasi yang cukup akurat. Dengan adanya penelitian ini, diharapkan dapat memberikan wawasan baru yang berupa hasil penelitian dan selanjutnya bisa digunakan pihak-pihak terkait yang membutuhkan analisis informasi mengenai kepuasan pelanggan online shop “Shopeecare” pada platform twitter baik itu digunakan untuk penelitian ataupun kebijakan.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah bagaimana melakukan sentiment analisis pada komentar di akun Twitter online shop berbahasa Indonesia dengan menggunakan algoritma *Naïve Bayes Classifier* dan *K-Nearest Neighbor* sehingga dapat membandingkan hasil sentiment dari online shop dengan kedua metode tersebut.

1.3. Batasan Masalah

Dalam penelitian ini terdapat beberapa batasan masalah yang dibahas agar penyusunan dan pembahasan penelitian dapat dilakukan secara terarah dan tercapai sesuai dengan yang diharapkan. Antara lain sebagai berikut :

1. Data komentar kepuasan pelanggan pada akun twitter “Shopeecare”.

2. Metode klasifikasi yang digunakan adalah *Naïve Bayes Classifier* dan *K-Nearest Neighbor*.
3. Klasifikasi Sentimen positif dan negative.
4. Bahasa pemrograman yang digunakan adalah Python.

1.4. Tujuan Penelitian

Tujuan penelitian ini adalah mencari model agar bisa dilakukan untuk melakukan analisis sentiment terhadap data komentar kepuasan di akun twitter “Shopeecare” dengan membandingkan algoritma *Naïve Bayes Classifier* dan *K-Nearest Neighbor*.

1.5. Manfaat Penelitian

Dengan tercapainya tujuan penelitian diatas, maka diharapkan akan bermanfaat bagi semua pihak manfaat penelitian ini adalah :

1. Dapat mengetahui perbandingan hasil sentiment analisis kepuasan pelanggan online shop dengan menggunakan model algoritma *Naïve Bayes Classifier* dan *K-Nearest Neighbor*.
2. Bagi UIN Sunan Kalijaga, dapat digunakan sebagai tambahan referensi yang terkait penelitian dengan tema “*Text Mining*”.
3. Bagi penulis, dapat memberikan pengetahuan yang baru, bahwa kepuasan pelanggan online shop dapat dilihat dari komentar di Twitter.
4. Bagi pembaca, dengan adanya penulisan ini dapat dimanfaatkan sebagai sumber pembelajaran untuk penelitian selanjutnya serta untuk menambah wawasan.

1.6.Keaslian Penelitian

Penelitian mengenai analisis sentiment dengan metode *Naïve Bayes Classifier* dan *K-Nearest Neighbor*, samapai saat ini sudah dilakukan oleh beberapa peneliti sebelumnya. Namun berdasarkan referensi dan tinjauan pustaka, penelitian Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta mengenai analisis sentiment menggunakan social media twitter dengan metode *Naïve Bayes Classifier* dan *K-Nearest Neighbor* mengenai kepuasan pelanggan online shop pada akun twitter resmi “Shopeecare” belum pernah ditemukan oleh peneliti.

1.7.Sistematika Penulisan

Sebagai gambaran dan kerangka yang jelas mengenai pokok bahasan setiap bab dalam penelitian ini, maka diperukan sistematika penulisan. Berikut adalah penjelasan pada tiap-tiap dalam laporan peneitian ini :

BAB I PENDAHULUAN

Bab pendahuluan berisikan penjelasan mengenai latar belakang dilakukannya penelitian, rumusan masalah penelitian, batasan masalah, tujuan penelitian, manfaat penelitian, keaslian penelitian, dan sistematika penulisan penelitian.

BAB II TINJAUAN PUSTAKAN DAN LANDASAN TEORI

Bab tinjauna pustaka dan landasan teori berisikan mengenai tinjauan pustaka dan landasan teori yang berhubungan dengan topik yang akan dibahas dalam penelitian ini.

BAB III METODE PENELITIAN

Bab metode penelitian ini berisi tentang metode ataupun algoritma yang digunakan serta tahapan-tahapan yang akan dilakukan dalam penelitian ini.

BAB IV HASIL DAN PEMBAHASAN

Bab hasil dan pembahasan berisi tentang proses penelitian sampai hasil dari penelitian yang telah dilakukan.

BAB V PENUTUP

Bab penutup menjelaskan tentang kesimpulan dari hasil penelitian yang telah dilakukan. Kekurangan yang ada pada penelitian ini dituliskan pada saran sehingga dapat berguna untuk perkembangan penelitian selanjutnya.

BAB V PENUTUP

5.1. Kesimpulan

Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa hasil dari model sentiment analisis yang digunakan menggunakan data latih sebanyak 1015 pada resmi “Shopeecare” dengan membandingkan metode *Naïve Bayes Classifier* dan *K-Nearest Neighbor* yang diproses menggunakan *Confussion Matrix* menghasilkan nilai akurasi lebih tinggi yang didapat menggunakan *Naïve Bayes Classifier*. Dengan hasil sebagai berikut :

1. Pada metode *Naïve Bayes Classifier* menghasilkan nilai akurasi sebanyak 86,27% dengan rasio *split* 90:10, nilai akurasi sebanyak 88,17% dengan rasio *split* 80:20, dan nilai akurasi 85,90% dengan rasio *split* 70:30.
2. Pada metode *K-Nearest Neighbor* menghasilkan nilai akurasi sebanyak 84,31% dengan rasio *split* 90:10, nilai akurasi sebanyak 87,68% dengan rasio *split* 80:20, dan nilai akurasi 87,21% dengan rasio *split* 70:30.

Sedangkan pada hasil implementasi pada data tweet baru sejumlah 480 yang didapatkan pada *Naïve Bayes Classifier* adalah 25,6% data positif dan 74,4% data negative. Sedangkan pada implementasi *K-Nearest Neighbor* adalah 40,2% data positif dan 59,8% data negativ

5.2. Saran

Pada penelitian ini penulis tentu tidak jauh dari kekurangan. Oleh karena itu, terdapat beberapa saran untuk dijadikan perbaikan untuk penelitian selanjutnya.

Saran-saran tersebut diantaranya :

- 1) Membandingkan kembali menggunakan metode pembelajaran mesin lainnya agar mendapatkan hasil yang lebih baik.
- 2) Menggunakan data latih yang diseleksi lebih baik dari sebelumnya dan menggunakan data yang lebih banyak supaya mendapatkan hasil yang lebih akurat.
- 3) Merancang system analisis yang dapat digunakan secara praktis.



DAFTAR PUSTAKA

- Azha., Masruroh, Ummi Masruroh., Wardhani, Luh Kesuma., & Okfalisa, (2019). Kinerja Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor Pendekatan Lexicon pada Analisis Sentimen di Media Twitter
- Chalida, Muftia, (2019). Analisis Sentimen Ujaran Kebencian Pemilihan Presiden 2019 Menggunakan Algoritma Naïve Bayes
- Fitriani, Risma Putri., Kurniawati, Ana., & Agustin, Dina, (2019). Implementasi Algoritma K Nearest Neighbor Terhadap Analisis Sentimen Review Restoran dengan Teks Bahasa Indonesia
- Baharuddin, Mus Mulyadi., Hasanudin, Tasrif., & Aziz, Huzain, (2019). Analisis Performa Metode K-Nearest Neighbor untuk Identifikasi Jenis Kaca
- Harsehanto, Ire Icca Agustiorini (2018). Analisis Karakter Kepribadian Seseorang Menggunakan Algoritma Naïve Bayes Classifier (studi Kasus Twitter Resmi Basuki Tjahaja Purnama dan Anies Baswedan)
- <https://medium.com/@ksnugroho/dasar-text-preprocessing-dengan-python-a4fa52608ffe>
- <https://yunusmuhammad007.medium.com/basic-text-preprocessing-menggunakan-nltk-86ba3e65a1dc>
- Nurintan, Tesya, (2019). Implementasi Algoritma Naïve Bayes Classifier untuk Menganalisis

Kepribadian Pengguna Twitter di Indonesia Berdasarkan Pendapatnya Terhadap Pasangan Calon Presiden dan Wakil Presiden Tahun 2019

Sasti, Laksmi, (2017). Analisis Sentimen Komentar pada Sistem Penelitian Kinerja Instruktur Training ICT (Information and Communication Technology) UIN Sunan Kalijaga Menggunakan *Naïve Bayes Classifier*

Satrio, Halim Satrio., Fauzi, Mochammad Ali., & indriati (2019). Klasifikasi *Tweets* Pada Twitter Menggunakan Metode *K-Nearest Neighbor* (K-NN) dengan Pembobotan TF-IDF.

Septian, Jeremy Andre., Fahrudin, Tresna Maulana, & Nugroho, Aryo. Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF IDF dan K Nearest Neighbor

Sevsa, Bama Abpama, (2018). Analisis Sentimen Komentar pada Indeks Kinerja Dosen Fakultas Sains dan Teknologi UIN Sunan Kalijaga Menggunakan *Naïve Bayes Classifier*

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA