

**SENTIMENT ANALYSIS PENILAIAN TOKO *ONLINE*
MENGUNAKAN *GAUSSIAN NAIVE BAYES* DAN *NEURAL NETWORK***



LAPORAN PENELITIAN

Sebagai Salah Satu Bentuk Pengamalan Tri Dharma Perguruan Tinggi

Oleh:

Dyah Apriliani, S.T., M.Kom.	NIPY. 09.015.225
Hepatika Zidny Ilmadina	NIPY. 08.017.340
Nurlaela	NIM. 19090029

**PROGRAM STUDI DIV TEKNIK INFORMATIKA
POLITEKNIK HARAPAN BERSAMA
AGUSTUS 2021**

SK Direktur Nomor : 098.05/PHB/V/2021 Tanggal 31 Mei 2021

Surat Perjanjian Pelaksanaan Penelitian

Nomor : 050.16/P3M.PHB/V/2021

**HALAMAN PERSETUJUAN
LAPORAN PENELITIAN**

**SENTIMENT ANALYSIS PENILAIAN TOKO *ONLINE*
MENGUNAKAN *NAIVE BAYES* DAN *NEURAL NETWORK***

Sebagai Salah Satu Bentuk Pengamalan Tri Dharma Perguruan Tinggi

Oleh:

Dyah Apriliani, S.T., M.Kom

NIPY: 09.015.225

Hepatika Zidny Ilmadina

NIPY. 08.017.340

Nurlaela

NIM. 19090029

Tegal, Agustus 2021

Menyetujui,

Ketua Program Studi DIV Teknik Informatika Ketua P3M

POLITEKNIK HARAPAN BERSAMA

POLITEKNIK HARAPAN BERSAMA



Slamet Wiyono, S.Pd., M.Eng.

NIPY. 08.015.222



Kusnadi, M.Pd.

NIPY. 04.015.217

**HALAMAN PENGESAHAN
PROPOSAL PENELITIAN**

1. Judul : SENTIMENT ANALYSIS PENILAIAN TOKO ONLINE
MENGUNAKAN NAIVE BAYES DAN NEURAL NETWORK

2. Ketua Peneliti

a. Nama Lengkap : Dyah Apriliani, ST, M.Kom
b. NIDN : 0614049002
c. NIPY : 09.015.225
d. Jabatan Fungsional : -
e. Program Studi : Sarjana Terapan Teknik Informatika
f. Alamat e-mail :

3. Jumlah Anggota : 3
Nama Anggota 1 : Hepatika Zidny Ilmadina, S. Pd, M. Kom
Nama Mahasiswa 1 : Nurlaela
Nama Mahasiswa 2 : Nurlaela

Biaya Penelitian : Rp. 3,314,000

Tegal, April 2021

Reviewer 1


Arif Rakhman, SE, S.Pd, M.Kom

NIPY. 05.016.291

Menyetujui,

Ketua Prodi Sarjana Terapan Teknik
Informatika

Politeknik Harapan Bersama


SLAMET WIYONO, S. Pd, M. Eng

NIPY. 08.015.222

Mengetahui,

Wakil Direktur 1

Politeknik Harapan Bersama


Heru Nurcahyo, S.Farm., M.Sc

NIPY. 10.007.038

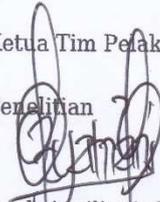
Reviewer 2


M. FIKRI HIDAYATTULLAH,
M.Kom

NIPY. 09.016.307

Ketua Tim Pelaksana

Penelitian


Dyah Apriliani, ST, M.Kom

NIPY. 09.015.225

Mengesahkan,

Ketua P3M

Politeknik Harapan Bersama


Kusnadi, M.Pd

NIPY. 04.015.217

HALAMAN PERNYATAAN

Dengan ini kami menyatakan bahwa:

1. Penelitian ini tidak pernah dibuat oleh peneliti lain dengan tema, judul, isi, metode, objek penelitian yang sama.
2. Penelitian ini bukan merupakan karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi
3. Dalam penelitian ini juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka

Tegal, Agustus 2021

Ketua Tim Peneliti



Dyah Apriliani, S.T., M.Kom

NIPY. 09.015.225

Anggota Tim Peneliti

Hepatika Zidny I, S.Pd., M.Kom

NIPY. 08.017.340

Anggota Tim Peneliti

Nurlaela

NIM. 19090029

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Illahi Robbi, karena atas rahmat dan karuniaNya, kami dapat menyelesaikan kegiatan penelitian dengan judul *entiment Analysis Penilaian Toko Online Menggunakan Naive Bayes Dan Neural Network* dengan lancar.

Penelitian ini merupakan perwujudan salah satu Tri Dharma Perguruan tinggi yang dilaksanakan oleh program studi D IV Teknik Informatika Politeknik Harapan Bersama Tegal. Tema penelitian dipilih berdasarkan kebutuhan akan tantangan kehidupan di jaman modern.

Dalam kesempatan ini, kami mengucapkan terima kasih yang sebesar – besarnya kepada :

1. Direktur Politeknik Harapan Bersama Tegal.
2. Kepala P3M (Penelitian dan Pengabdian Pada Masyarakat) Politeknik Harapan Bersam Tegal yang telah memberikan kemudahan dalam pelaksanaan pengabdian.
3. Semua pihak yang telah membantu berjalannya kegiatan penelitian ini.

Akhir kata semoga kegiatan penelitian ini dapat bermanfaat bagi peningkatan pengetahuan.

Tegal, Agustus 2021

Dyah Apriliani, S.T., M.Kom

DAFTAR ISI

HALAMAN JUDUL	Error! Bookmark not defined.
HALAMAN PERSETUJUAN LAPORAN PENELITIAN	Error! Bookmark not defined.
HALAMAN PENGESAHAN LAPORAN PENELITIAN	ii
HALAMAN PERNYATAAN	Error! Bookmark not defined.
KATA PENGANTAR	v
DAFTAR ISI	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL	ix
ABSTRAK	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	2
1.5 Manfaat Penelitian	2
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	4
2.1 Tinjauan Pustaka	4
2.2 Dasar Teori	5
BAB III METODE PENELITIAN	8
1.1 Kerangka Penelitian.....	8
3.2 Bahan dan Alat Penelitian	8
3.3 Prosedur Penelitian.....	9
3.4 Target Luaran	10
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	11
4.1 Hasil Penelitian.....	11

BAB V KESIMPULAN DAN SARAN	16
5.1 Kesimpulan	16
5.2 Saran	16
LAMPIRAN	19

DAFTAR GAMBAR

Gambar 2.1 Skema Dasar Neural Networks	7
Gambar 3.1 Kerangka Metodologi Penelitian	8
Gambar 3.2 Prosedur Penelitian	9
Gambar 4.1 Grafik Tiap Kelas pada Dataset	11

DAFTAR TABEL

Tabel 4.1 Contoh Unigram dan Bigram Komentar Pelanggan	12
Tabel 4.2 Evaluasi Performa Kedua Metode Klasifikasi	13
Tabel 4.3 Perbandingan Hasil Evaluasi Rata-rata	14

ABSTRAK

Perkembangan teknologi yang semakin pesat mengakibatkan semakin berkembangnya kegiatan jual beli yang dilakukan secara *online*. Hal tersebut memicu munculnya berbagai toko *online* dalam berbagai *marketplace*. Sistem pendukung keputusan berperan penting dalam proses penilaian terhadap suatu pelayanan toko *online* bagi pembeli. Tujuan dari penelitian ini adalah melakukan sentiment analysis untuk penilaian toko online dengan menggunakan dan *Gaussian Naive Bayes* Neural Network. Metode penelitian dilakukan dalam beberapa tahapan penelitian. Tahap awal dimulai dengan proses pengambilan data set yang akan digunakan dalam penelitian berupa teks penilaian toko *online*, *praprocessing* data, pembobotan dan eksperimen model dengan menggunakan *Naive Bayes* Dan *Neural Network* serta evaluasi terhadap hasil eksperimen. Hasil penelitian mneghasilkan performa model *Gaussian Naive Bayes* pada *sentiment analysis* berupa klasifikasi penilaian pelanggan lebih baik, yaitu akurasi sebesar 69,01% pada dataset yang terbatas.

Kata kunci: Sentiment analysis; toko *online*; *Naive Bayes*, *Neural Networks*

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi yang sangat pesat mengakibatkan semakin berkembangnya dunia digital. Perkembangan tersebut juga dirasakan dalam proses jual beli barang. Transaksi yang dulunya dilakukan dengan cara manual, saat ini sudah berubah ke arah digital yaitu dengan menggunakan toko *online*. Dalam proses jual beli seseorang melakukan pemilihan toko *online* tidak hanya dilihat dari harga yang ditawarkan saja, tetapi juga dilihat dari berbagai penilaian yang diberikan oleh pembeli lain terhadap toko tersebut. Permasalahan terkait dengan rekomendasi penilaian terhadap toko online saat ini banyak diinginkan, hal ini mengingat informasi yang ada dapat memberikan sebuah pendukung keputusan dalam melakukan pencarian toko online terbaik dan yang diinginkan.

Proses menjual dan membeli produk dilakukan secara *online* dengan menggunakan situs *online* atau biasa dikenal dengan istilah *marketplace*, dimana dalam toko *online* tersebut semua produk dapat di jual dan dibeli kapanpun dan dimanapun dengan berbagai kelebihan dan kekurangannya[1]. Semakin sering digunakan, tidak sedikit pembeli yang juga mengalami beberapa keluhan atau pun komentar terkait dengan produk yang dibeli. Beberapa orang memberikan komentar berupa komentar positif maupun komentar negatif tentang suatu produk atau suatu toko nya[2]. Sentimen analisis atau biasa dikenal dengan istilah analisis sentimen atau opinion mining merupakan salah satu proses untuk menemukan pendapat untuk berbagai topik atau teks yang diampaiakan oleh pengguna [3]. Dalam penelitian ini adalah pengguna toko *online*. *Sentiment Analysis* (SA) bekerja dengan cara memproses, memahami, meng-ekstrak, dan mengolah data secara otomatis yang sifatnya berupa tekstual dimana nantinya akan menghasilkan suatu informasi yang terkandung didalamnya. Salah satu kelebihan yang dimiliki SA ini akan menghasilkan sistem yang digunakan sebagai pendukung keputusan yang dapat berupa klasifikasi sentiment, *review* sebuah objek, mendeteksi sebuah *spam* dan lainnya [4].

Ada beberapa metode sentiment analisis yang sering digunakan, diantaranya adalah *Support Vector Machine*, *Naïve Bayes*, *Neural Network*, *Decision Tree*, *Bayesian Network* dan *Maximum Entropy* [4]. Dari beberapa tersebut, dalam penelitian ini akan digunakan metode *Gaussian Naive Bayes* Dan *Neural Network*. Tujuan dari penelitian ini adalah didapatkan model terbaik untuk sentiment analisis pada penilaian toko online dengan menggunakan metode *Gaussian Naive Bayes* Dan *Neural Network* untuk mengkalsifikasikan hasil *review* pelanggan.

1.2 Rumusan Masalah

Berdasarkan permasalahan di atas, maka rumusan masalah dalam penelitian ini adalah “Bagaimana menghasilkan rekomendasi dan mengevaluasi penilaian terhadap toko *online* saat ini masih sulit dan banyak diinginkan untuk dapat memberikan sebuah pendukung keputusan dalam melakukan pencarian toko *online* yang terbaik dan diinginkan oleh pembeli”

1.3 Batasan Masalah

Agar lingkup penelitian tidak melebar, maka usulan penelitian ini dibatasi pada beberapa poin berikut:

1. Data yang akan digunakan diambil dari komentar pelanggan dari situs: <https://shopee.co.id/elhijab.official>.
2. Metode yang digunakan adalah *Naive Bayes* dan *Neural Network*.

1.4 Tujuan Penelitian

Tujuan dari usulan penelitian ini adalah memperoleh rekomendasi terbaik untuk sentiment analisis penilaian toko online dengan menggunakan metode *Gaussian Naive Bayes* Dan *Neural Network*.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Membantu memberikan rekomendasi hasil *review* toko *online* dengan mengimplementasikan metode *Gaussian Naive Bayes* Dan *Neural Network*.

2. Menjadi referensi bagi peneliti yang akan melakukan penelitian dengan tema dan/atau metode yang serupa.
3. Terealisasinya Tri Dharma perguruan tinggi.

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Penelitian dengan judul “Analisa Sentimen Analisa Untuk penilaian Pelayanan Situs Belanja Online Menggunakan Metode Naïve Bayes”. Proses analisis terhadap hasil review marketplace yang diungkapkan ke dalam media sosial, yaitu dengan menggunakan twitter. Dalam penelitian tersebut dilakukan beberapa tahapan penelitian, diantaranya : koleksi data penelitian, pre processing data yang meliputi *leaning data, case folding, tokenizing, case normalization, stop word, convert negation dan stemming*. Metode *Naïve Bayes* digunakan untuk melakukan klasifikasi data yang telah dihasilkan. Dan dilakukan evaluasi terhadap model tersebut. Hasil akurasi rata-rata yang dihasilkan adalah sebesar 93,33% [5].

Penelitian selanjutnya yang dilakukan oleh Elik, dkk dengan judul Analisis Sentiment pada Ulasan Pembelian Produk di *Marketplace Shopee* Menggunakan Pendekatan *Natural Language Processing*” dalam penelitian ini dilakukan analisis ulasan produk yang digunakan dalam evaluasi produk dan juga layanan dari Shopee. Kegiatan analisis dilakukan dengan melihat komentar dari para pelanggan. Penelitian dilakukan dengan menggunakan metode KNN dan TF-IDF dengan pendekatan NLP untuk melakukan klasifikasi produk “hijab instan” ke dalam dua klasifikasi yaitu positif dan negatif. Proses klasifikasi dengan menggunakan NLP mendapatkan akurasi sebesar 76,92%, presisi 80,00% dan recall 74,07%, sedangkan tanpa NLP hanya mendapat akurasi sebesar 69,23%, presisi 80,00% dan recall 64,52% [6].

Penelitian lain dengan judul “ Analisis Sentimen *Tweet* Menggunakan Backpropagation Neural Network”. Pada penelitian tersebut mengkategorikan komentar di *tweet* ke dalam tiga kelas yaitu (positif, netral dan negatif). Dalam penelitian ini digunakan metode *Backpropagation Neural Network*. Dalam penelitian ini digunakan input jaringan yang dirangking dengan menggunakan TF-

IDF. Hasil pengujian dari metode tersebut menunjukkan hasil akurasi sebesar 78.34% dan presisi 84.2 [7].

Penelitian lain dengan judul “*Sentiment analysis for Indonesia hotel services review using optimized neural network*”. Dalam penelitian ini model sentiment diimplementasikan untuk melakukan klasifikasi terhadap hasil *review* terhadap pelayanan hotel. Dalam penelitian ini digunakan model *neural network* dalam melakukan klasifikasi. Penelitian dilakukan dalam beberapa tahapan diantaranya: *praprocessing* data, pemilihan atribut, dan optimasi parameter untuk menghasilkan akurasi terbaik. Dalam penelitian tersebut diperoleh akurasi sebesar 88,99% [8].

2.2 Dasar Teori

1. Sentimen Analisis

Sentimen analisis atau yang dikenal dengan istilah analisis sentimen atau *opinion mining* merupakan salah satu bidang ilmu yang menganalisa pendapat, sentiment, evaluasi, penilaian, sikap, emosi public terhadap suatu produk, berita, informasi ataupun jasa organisasi maupun individu [9]. Analisis sentiment biasanya fokus pada komentar atau pendapat yang mengungkapkan suatu ekspresi, dimana ekspresi tersebut bisa berupa ekspresi positif maupun negatif. Sentiment analisis ini mengolah data tekstual secara otomatis untuk melihat sentiment yang terkandung di dalam suatu opini atau pendapat [11].

2. Text Mining

Text mining juga dikenal dengan text data mining atau pencarian pengetahuan di basis data tekstual adalah proses yang semi otomatis melakukan ekstraksi dari pola data[12]. Tipe pekerjaan *text mining* meliputi kategorisasi, text clustering, ekstraksi konsep/entitas, analisis sentimen, document summarization, dan entity-relation modeling (yaitu, hubungan pembelajaran antara entitas) [13]. Sumber data yang digunakan pada *text mining* adalah kumpulan teks yang memiliki format yang tidak terstruktur

atau minimal semi terstruktur. Tujuan dari *text mining* adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen.

3. *Gaussian Naive Bayes*

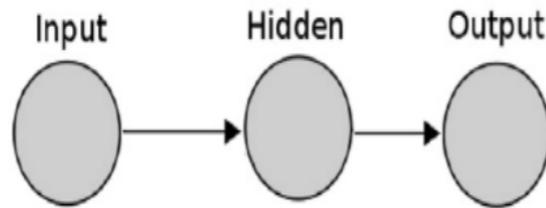
Ketika dihadapkan dengan data kontinu maka akan dilakukan pendistribusian Gaussian Naive Bayes dibawah bentuk distribusi Gaussian Naive Bayes[14]:

1. Baca data latih
2. Hitung nilai probabilitas pada data kontinu dengan cara
 - a. Cari nilai mean dan standar deviasi dari masing-masing parameter yang merupakan data kontinu.
 - b. Cari nilai probabilitik dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.
3. Mendapatkan nilai mean, standar deviasi dan probabilitas

4. *Neural Network*

Neural Networks (jaringan syaraf tiruan) merupakan sistem prosesor paralel terhubung satu sama lain dalam bentuk grafik diarahkan. menurut bagan setiap neuron dari jaringan direpresentasikan sebagai node. koneksi ini memberikan struktur hirarkis yang mencoba untuk meniru fisiologi otak, mencari model-model baru pengolahan untuk memecahkan masalah tertentu di dunia nyata. Permasalahan penting dalam pengembangan teknik jaringan syaraf tiruan adalah perilaku yang berguna harus dipelajari, diakui dan diterapkan antara obyek dan frame benda di dunia nyata [15]

Neural Networks terdiri dari 3 lapisan, yaitu Input Layer, Hidden Layer dan Output Layer, seperti ditunjukkan pada gambar 2.1. Setiap lapisan bertanggung jawab untuk melakukan fungsi yang sama melengkapi sistem. Struktur ini didasarkan pada modifikasi dari model tiga lapisan arsitektur perangkat lunak, ini adalah: Data lapisan, lapisan layanan (Web, sensor) dan lapisan presentasi (Web, Windows dan Agen) [16].



Gambar 2.1. Skema Dasar Neural Networks

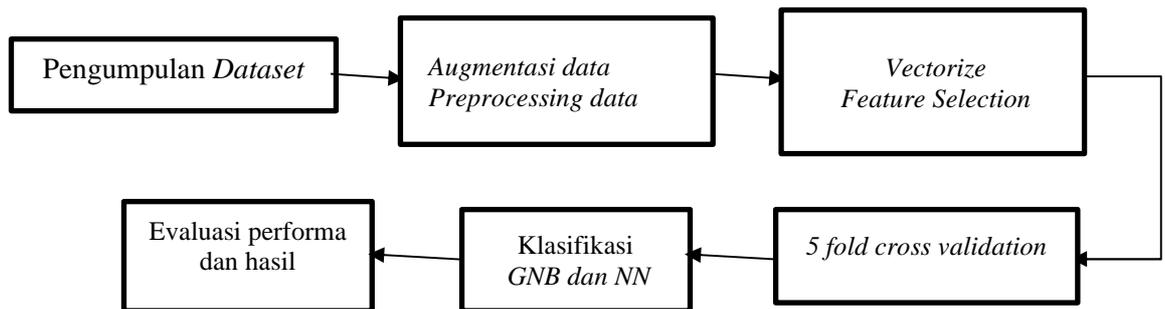
Neural Network ditentukan oleh 3 hal yang sangat penting yaitu : Pola hubungan antar neuron, metode untuk menentukan bobot penghubung (training learning) dan fungsi aktivasi.

BAB III METODE PENELITIAN

1.1 Kerangka Penelitian

Penelitian ini dimulai dengan mengidentifikasi dan merumuskan permasalahan berdasarkan latar belakang penelitian akan dilakukan. Melakukan studi literatur untuk mendapat metode terbaik untuk memecahkan permasalahan yang diangkat. Penelitian kemudian dilanjutkan dengan pengumpulan dataset secara manual berdasarkan masalah yang telah dirumuskan. Dataset ini dikumpulkan komentar dan rating pelanggan dari perusahaan *e-commerce Shopee* Indonesia, pada toko *Elzatta Official*.

Pengolahan dataset kemudian dilanjutkan ke proses augmentasi dan preprocessing data. Selanjutnya untuk melakukan sebuah *sentiment analysis* dilakukan vektorisasi kalimat menggunakan *TF-IDF Vectorizer*. Setelah mendapatkan vector untuk masing-masing komentar pelanggan, peneliti mengklasifikasikan rating yang diberikan pelanggan berdasarkan dari hasil komentarnya. Kerangka metodologi penelitian secara detail dijelaskan pada Gambar 3.1 berikut ini.



Gambar 3. 1. Kerangka Metodologi Penelitian

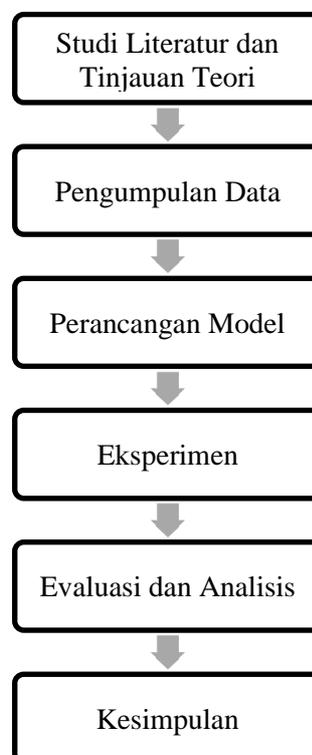
3.2 Bahan dan Alat Penelitian

Dataset yang digunakan pada penelitian ini dikumpulkan secara manual berupa komentar dan rating pelanggan dari perusahaan *e-commerce Shopee* Indonesia, pada toko *Elzatta Official*. Pengolahan data selanjutnya

diimplementasikan menggunakan *Python* dengan memanfaatkan *Keras library* dengan *Tensorflow backend* serta diimplementasikan menggunakan *Google Colab* dengan *single 12GB NVIDIA Tesla K80 GPU*. Implementasi model juga menggunakan *library* pendukung seperti *numpy, sklearn, matplotlib, panda, nltk, bs4, spacy, re, dan textblob*. Penelitian dibagi menjadi 2 skenario utama, yaitu membandingkan hasil klasifikasi model *Gaussian Naïve Bayes (GNB)* dengan *Neural Network (NN)*.

3.3 Prosedur Penelitian

Prosedur penelitian yang digunakan terdiri atas beberapa tahapan yang meliputi identifikasi masalah, studi literatur yang terkait dengan permasalahan dan pengetahuan terkait penelitian *sentiment analysis*, tinjauan teori atas penelitian-penelitian yang serupa, merancang model yang akan diusulkan berdasarkan studi literatur dan teori, melakukan eksperimen percobaan, evaluasi dan analisis hasil penelitian. Secara detail, prosedur penelitian yang akan dilakukan ditampilkan pada Gambar 3.2.



Gambar 3. 2. Prosedur penelitian

Studi Literatur dan Tinjauan Teori

Kajian teori yang digunakan sebagai landasan dalam penelitian. Selain itu, dijelaskan juga penelitian-penelitian terdahulu yang menunjang pengembangan dan pengimplementasian metode yang diajukan dalam penelitian. Terdapat penjelasan tentang kendala dan alasan pemilihan metode yang menunjukkan posisi penelitian sebagai kontribusi penelitian terhadap penelitian-penelitian yang sudah ada.

Pengumpulan Data

Penelitian yang akan dilakukan menggunakan *private* dataset yaitu rating dan komentar pelanggan toko Elzatta Official pada *e-commerce Shopee* Indonesia.

Perancangan Model

Pada tahap ini dilakukan perancangan model yang akan digunakan dalam *sentiment analysis*, dengan menggunakan berbagai teknik *text processing*.

Eksperimen

Pada tahap ini dilakukan pengujian/eksperimen pada metode GNB dan NN.

Evaluasi dan Analisa Hasil Uji

Evaluasi dilakukan untuk mengetahui performa model yang digunakan serta dilakukan analisa untuk mengetahui parameter-parameter yang mempengaruhi performa model.

Kesimpulan

Kesimpulan dari serangkaian penelitian yang dilakukan serta saran yang dapat diimplementasikan untuk mengembangkan penelitian selanjutnya.

3.4 Target Luaran

Target luaran dari penelitian ini berupa:

- a. Publikasi pada Jurnal Teknologi Informasi dan Ilmu Komputer Fakultas Ilmu Komputer Universitas Brawijaya. Yang Terakreditasi Sinta 2. JTIK Terbit di bulan Februari, April, Juni, Agustus, Oktober dan Desember, masih dalam bentuk draft yang selanjutnya akan disubmit.
- b. Materi atau bahan ajar untuk mata kuliah Data Mining.
- c. Akan di daftarkan dalam HKI (jika memungkinkan).

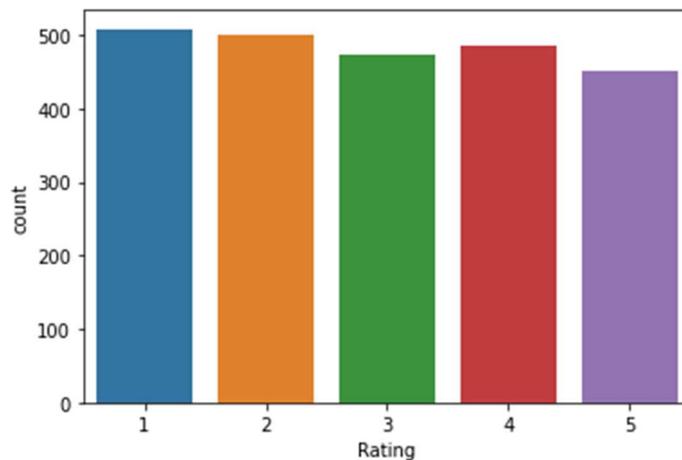
BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

4.1 Hasil Penelitian

a. Pelatihan dan Evaluasi Model

Sentiment Analysis ini dilakukan untuk mengklasifikasikan rating yang diberikan pelanggan berdasarkan komentarnya. Rating dibagi menjadi 5 kelas, yaitu rating 1, rating 2, rating 3, rating 4, dan rating 5. Dataset asli berjumlah 604 komentar. Sehingga untuk menunjang pelatihan data dengan *deep learning*, dilakukan *data augmentation* dengan cara menduplikasi data sehingga dataset total adalah 2416 komentar, digambarkan pada grafik Gambar 4.1. Sebelum menuju tahap klasifikasi dilakukan langkah *text processing* seperti menghilangkan tautan html, menghapus karakter di luar teks, menghilangkan angka dan simbol, menghilangkan *multiple* spasi, dan menghapuskan *emoticon*.



Gambar 4. 1. Grafik Tiap Kelas pada Dataset

Untuk melakukan klasifikasi pada teks, selanjutnya dilakukan vektorisasi dengan menggunakan algoritma TF-IDF. Untuk mengoptimasi metode klasifikasi yang akan digunakan, maka diimplementasikan *feature selection chi2*. Berdasarkan hasil seleksi fitur tersebut diperoleh beberapa

contoh unigram dan bigram komentar yang berkorelasi dengan rating pelanggan. Lebih lengkap dijelaskan pada Tabel 4.1 berikut ini.

Tabel 4. 1. Contoh Unigram dan Bigram Komentar Pelanggan

Rating	Unigram	Bigram
1	<ul style="list-style-type: none"> ● buruk ● tidak ● dikirim ● chat ● pcs 	<ul style="list-style-type: none"> ● dikirim tidak ● kecewa banget ● tdk sesuai ● pengiriman lama ● di chat
2	<ul style="list-style-type: none"> ● lain ● padahal ● tipis ● pengiriman ● model 	<ul style="list-style-type: none"> ● tipis banget ● ga sesuai ● warna tidak ● sangat lama ● ga ada
3	<ul style="list-style-type: none"> ● harusnya ● walau ● ukurannya ● sayings ● si 	<ul style="list-style-type: none"> ● sedikit kecewa ● pengemasan lama ● bahannya bagus ● tidak ada ● bahannya tipis
4	<ul style="list-style-type: none"> ● rapih ● oke ● enak ● cepat ● lumayan 	<ul style="list-style-type: none"> ● sesuai pesanan ● harga segitu ● produk baik ● mudah dibentuk ● pengiriman cepat
5	<ul style="list-style-type: none"> ● lembut ● nyaman ● makasih ● terimakasih ● bagus 	<ul style="list-style-type: none"> ● sudah sampai ● terima kasih ● barang sesuai ● barangnya bagus ● paket sudah

Dari dataset yang telah menjadi vector kemudian dibagi menjadi data menjadi 80% data latih dan 20% data uji dengan penerapan *5-fold cross validation*. Klasifikasi dilakukan dengan menggunakan metode *Gaussian Naïve Bayes* (GNB) dan *Neural Network* (NN). Hasil evaluasi performa kedua model kemudian dibandingkan untuk mengetahui metode terbaik dalam klasifikasi rating berdasarkan komentar pelanggan toko.

b. Perbandingan Metode GNB dengan NN

Implementasi metode GNB menggunakan modul *sklearn* dengan melakukan *min-max scaling* agar data normal. Sedangkan pada NN dilakukan pembentukan *layer* menggunakan *embedding layer*. Model diterapkan *categorical cross entropy* sebagai *loss function* dan *softmax* sebagai *activation function* dikarenakan merupakan eksperimen klasifikasi *multiclass*. Model kemudian dikompilasikan dengan *adam optimizer* dilatih sebanyak 20 *epoch* dengan *early-stopping*. Pelatihan juga mengimplementasikan metode pembagian data latih dengan *5-fold validation* pada seluruh skenario percobaan. Secara keseluruhan perbandingan evaluasi performa kedua model tersebut dijelaskan pada Tabel 4.2.

Tabel 4. 2. Evaluasi Performa Kedua Metode Klasifikasi

Rating	<i>Gaussian Naïve Bayes</i>			<i>Neural Network</i>		
	<i>Precisio n</i>	<i>Recall</i>	<i>f-1 Score</i>	<i>Precisio n</i>	<i>Recall</i>	<i>f-1 Score</i>
1	0,65	0,81	0,72	0,59	0,52	0,64
2	0,80	0,53	0,64	0,59	0,53	0,64
3	0,96	0,52	0,68	0,80	0,52	0,53
4	0,67	0,61	0,64	0,67	0,61	0,52
5	0,59	1,00	0,75	0,59	0,53	0,65

Secara keseluruhan dapat terlihat bahwa beberapa kali GNB maupun NN akan lebih baik pada salah satu aspek. Hal ini juga ditunjukkan dengan

perbedaan akurasi yang tidak terlalu jauh, yaitu 69,01% pada metode GNB dan 63,52% pada NN.

Tabel 4. 3. Perbandingan Hasil Evaluasi Rata-rata *accuracy*, *precision*, *recall*, dan *f-score* pada kedua metode

Metode	Accuracy (%)	macro avg		
		<i>precision</i>	<i>recall</i>	<i>f-score</i>
<i>Gaussian Naïve Bayes</i>	69,01	0,74	0,69	0,68
<i>Neural Network</i>	63,52	0,65	0,54	0,60

Secara lengkap diperoleh bahwa hasil evaluasi performa menggunakan GNB lebih besar dibandingkan dengan menggunakan NN dilihat dari nilai *accuracy*, *precision*, *recall*, maupun *f-score*.

c. Pembahasan

Dataset *private* berupa kumpulan komentar dan rating pelanggan toko *online* Elzatta Official diimplementasikan dengan metode Gaussian Naive Bayes (GNB) dan Neural Network (NN). Implementasi kedua metode tersebut menghasilkan hasil akurasi masing-masing 69,01% untuk GNB dan 63,52% untuk NN. Metode GNB diimplementasikan dengan menggunakan modul *sklearn*, sedangkan NN dengan membangun *layer embedded* dengan melakukan vektorisasi pada komentar pelanggan. Pre-processing data juga dilakukan untuk membuat data dapat diolah secara maksimal. Penghilangan simbol, karakter, stemming kalimat juga dilakukan. Akan tetapi stemming dilakukan dengan metode biasa, tidak dikhususkan pada modul bahasa Indonesia. *Stop word* dalam bahasa Indonesia juga belum dilakukan.

Implementasi pada NN yang merupakan algoritma *deep learning* seharusnya dapat memberikan performa yang lebih baik dibandingkan dengan GNB. Hal tersebut tidak dapat dicapai dalam penelitian ini dikarenakan ketersediaan variasi dataset yang terbatas, yaitu hanya 604 komentar untuk 5 kelas data rating 1-5. Dalam hal ini, *layer* hanya mempelajari fitur yang sama dan dilakukan berulang untuk masing-masing

dataset karena proses augmentasi. Sehingga pada penelitian ini performa NN tidak lebih bagus dibandingkan dengan GNB pada klasifikasi rating pelanggan berdasarkan komentar yang diberikan.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan, beberapa simpulan yang dapat diperoleh yaitu:

1. Identifikasi kepuasan pelanggan yang menjadi sebuah penilaian (rating) dapat dilakukan dengan metode *Gaussian Naive Bayes* dan *Neural Network*.
2. Performa model *Gaussian Naive Bayes* pada *sentiment analysis* berupa klasifikasi penilaian pelanggan lebih baik, yaitu akurasi sebesar 69,01% pada dataset yang terbatas.

5.2 Saran

Adapun saran untuk penelitian lanjutan ini adalah dengan mengimplementasikan modul “Sastrawi” untuk melakukan *preprocessing* yang lebih tepat dapat dalam mengolah data *sentiment* dalam bahasa Indonesia.

DAFTAR PUSTAKA

- [1] A. Ishak, “Analisis Kepuasan Pelanggan Dalam Belanja Online : Sebuah Studi Tentang Penyebab (Antecedents) dan Konsekuensi (Consequents),” *Jurnal Siasat Bisnis*, Vol. 16 No. 2, , Hal : 141-154, Juli 2012.
- [2] D. G. Nugroho, Y. H. Chrisnanto, and A. Wahana, “Analisis Sentimen pada Jasa Ojek Online Menggunakan Metode Naive Bayes,” *Jur. Inform. Fak. Mat. dan Ilmu Pengetah. Alam Univ. Jenderal Achmad Yani*, pp. 156–161, 2015.
- [3] B. S. D and P. D. Gore, “Sentiment Analysis on Twitter Data Using Support Vector Machine,” *International Journal of Computer Science Trends and Technology (IJCST) – Volume 4 Issue 3*, pp. 365–370, May - Jun 2016
- [4] A. Tripathy, A. Agrawal, and S. K. Rath, “Classification of Sentimental Reviews Using Machine Learning Techniques,” *Procedia Comput. Sci.*, vol. 57, pp. 821–829, 2015.
- [5] Muljono, D.P Artanti, A. Syukur, dkk “Analisis Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naïve Bayes,” *Stimik Admaluhur Pangkalpinang*, 8-9 Maret 2018
- [6] E.H. Muktafin, Kusri, E.T Luthfi, “Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing,” *Jurnal Eksplora Informatika*. Vol 1 no 10, pp. 32-42, 2020
- [7] M. A. Asuja, Saniati, “Analisis Sentimen Tweet Menggunakan Backpropagation Neural Network,” *Jurnal Teknoinfo*. Vol 11 No 2, pp 23-28, 2016
- [8] D. Apriliani, T. Abdidin, E. Sutanta, A. Hamzah, O. Somantri “ Sentiment analysis for Indonesia hotel services review using optimized neural network,” *ICCGANT 2019*
- [9] F. Z. Tala, “A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia,” *M.Sc. Thesis, Append. D*, vol. pp, pp. 39–46, 2003.
- [10] K. Chen, Z. Zhang, J. Long, and H. Zhang, “Turning from TF-IDF to TF-

- IGM for term weighting in text classification,” *Expert Syst. Appl.*, vol. 66, pp. 1339–1351, 2016.
- [11] Liu, Bing. (2012). *Sentiment Analysis and Opinion Mining*. USA : Morgan & Claypool Publishers.
- [12] Ma, Jian dkk. (2012). An Ontology-Based Text-Mining Method to Cluster Proposals for Research Project Selection. *IEEE Trans. Syst., Man, Cybern. Part A, Syst & Humans* (vol. 42).
- [13] Han, Jiawei., Kamber, Micheline., & Pei, Jian. (2012). *Data Mining Concepts and Techniques* (3rd ed.). USA : Simon Frase University
- [14] Blachier, M., Leleu, H., Peck-Radosavljevic, M., Valla, D. C., & Roudot-Thoraval, F. (2013). The burden of liver disease in Europe: A review of available epidemiological data. In *Journal of Hepatology*. <https://doi.org/10.1016/j.jhep.2012.12.005>

LAMPIRAN

Lampiran 1. Justifikasi Realisasi Anggaran Biaya Penelitian

1. Honorarium Pembantu Peneliti					
No.	Justifikasi Pemakaian				Jumlah Honor
1	Pembantu Pengumpulan Data				Rp 500,000
2	Pembantu Analisis Data				Rp 1.000,000
Sub Total					Rp 1.500,000
2. Bahan Habis Pakai dan Peralatan					
No	Material	Justifikasi Pemakaian	Satuan	Harga Satuan	Jumlah Harga
1	Material 1	Kertas A4 Bola Dunia 80 gram	7 Rim	Rp 50.000	Rp 350.000
2	Material 2	ATK (Pulpen, Spidol, Penjepit Kertas,dll)	Pack, eceran		Rp 50.000
3	Material 3	Tinta Print hitam	2 Buah	Rp 60.000	Rp 120.000
4	Material 3	Tinta Print warna	3 Buah	RP 70.000	Rp 210.000
5	Material 4	Fotocopy dan Penjilidan	6 Jilid		Rp 284.000
Sub Total (Rp)					Rp. 1.014.000
3. Lain-Lain					
No.	Justifikasi Pemakaian				Jumlah Honor
1	Biaya Publikasi				Rp 800,000
Sub Total					Rp 800,000
Total					Rp 3,314,000

Tegal, Agustus 2021

Menyetujui,
Ketua P3M

Mengusulkan,
Ketua Peneliti

Kusnadi, M.pd
NIPY. 04.015.217

Dyah Apriliani, S.T., M.Kom
NIPY. 09.015.225

Lampiran 2. Susunan Organisasi Tim Peneliti Dan Pembagian Tugas

No.	Nama/Status	Jabatan	Uraian Tugas
1	Dyah Apriliani, S.T., M.Kom /Dosen	Ketua Peneliti	<p>a. Mengkoordinasi proses pengambilan data, pengumpulan data, preprocessing data, pemodelan, validasi dan evaluasi.</p> <p>b. Mengkoordinasi persiapan instrumen penelitian, perlengkapan penelitian, dan instrumen penunjang.</p> <p>c. Bertanggungjawab dalam pengembangan model.</p> <p>d. Mengkoordinasi penyusunan laporan akhir penelitian, publikasi hasil penelitian jurnal nasional/internasional (terakreditasi/ bereputasi) dan HKI.</p> <p>e. Bertanggung jawab terhadap hasil pelaporan penelitian.</p>
2	Hepatika Zidny Ilmadina, S.Pd., M.Kom / Dosen	Anggota Peneliti (1)	<p>a. Turut bertanggung jawab dalam proses pengambilan data, pengumpulan data, preprocessing data, pemodelan, validasi dan evaluasi.</p> <p>b. Turut bertanggung jawab terhadap hasil pelaporan penelitian</p>
3	Nurlaela / Mahasiswa	Anggota Peneliti (2)	Membantu pengumpulan data, analisis data, dan implementasi <i>metode</i> .

Lampiran 3. Organisasi pengusul penelitian

1. Ketua

Nama : Dyah Apriliani, S.T., M.Kom
NIPY : 09.015.225
NIDN : 0614049002
Pangkat/Golongan : III/c
Jabatan Fungsional : Lektor
Jabatan Struktural : Koordinator Penjamin Mutu
Bidang Ilmu : Sistem Informasi
Unit Kerja : D IV Teknik Informatika

Pengalaman Penelitian:

- a. Implementasi BACK-END Sistem E-ABSENSI Mahasiswa Politeknik Harapan Bersama (Studi Kasus : Program Studi DIV Teknik Informatika) (2017)
- b. Sentiment Analysis Kepuasan Pelanggan Warung Dan Restoran Kuliner Kota Tegal Berbasis Text Mining Sebagai Upaya Peningkatan Nilai Jual Produk Kuliner Lokal (2017)
- c. Sistem Penilaian Raport Dosen Menggunakan Metode Promethee
- d. Sentiment Analysis Tingkat Kepuasan Pelanggan Terhadap Pelayanan Warung dan Restoran Kuliner Di Kota Tegal Menggunakan Support Vector Machine (SVM) Berbasis Feature Selection (2018)
- e. Perbandingan Metode Promethee Dan Topsis Dalam Sistem Pendukung Keputusan Penilaian Kinerja Dosen (Studi Kasus : Politeknik Harapan Bersama)
- f. Optimasi Parameter Pada Sentiment Analysis Penilaian Pelayanan Hotel Menggunakan Algoritma Genetika Berbasis Hybrid Feature Selection

2. Anggota (1)

Nama : Hepatika Zidny Ilmadina, S.Pd., M.Kom
NIPY : 08.017.340
NIDN : -
Pangkat/Golongan : -
Jabatan Fungsional : Tenaga Pengajar
Jabatan Struktural : -
Bidang Ilmu : Teknik Informatika
Unit Kerja : D IV Teknik Informatika

3. Anggota (2)

Nama : Nurlaela
NIM : 19090029

Lampiran 4. Surat Keputusan



Yayasan Pendidikan Harapan Bersama
PoliTeknik Harapan Bersama

Kampus I : Jl. Mataram No.9 Tegal 52142 Telp. 0283-352000 Fax. 0283-353353
Kampus II : Jl. Dewi Sartika No. 71 Tegal 52117 Telp. 0283-350567
Website : www.poltektegal.ac.id | Email : sekretariat@poltektegal.ac.id

SURAT KEPUTUSAN
DIREKTUR POLITEKNIK HARAPAN BERSAMA
NOMOR: 098 .05/PHB/V/2021
TENTANG
PENERIMA PENDANAAN HIBAH KOMPETITIF PENELITIAN DAN
PENGABDIAN MASYARAKAT OLEH INSTITUSI
BAGI DOSEN POLITEKNIK HARAPAN BERSAMA
TAHUN ANGGARAN 2020/2021 SEMESTER GENAP

- DIREKTUR POLITEKNIK HARAPAN BERSAMA,**
- Menimbang** : a. bahwa untuk meningkatkan kualitas dan kuantitas pelaksanaan penelitian dan pengabdian kepada masyarakat bagi Dosen di Politeknik Harapan Bersama, maka perlu menetapkan kebijakan dalam bidang pendanaan penelitian dan pengabdian kepada masyarakat;
- b. bahwa untuk tertib administrasi keuangan dalam pendanaan penelitian dan pengabdian kepada masyarakat, maka perlu ditetapkan tahapan penyerahan pendanaan oleh institusi untuk hibah kompetitif penelitian dan pengabdian masyarakat kepada Dosen Politeknik harapan Bersama;
- c. bahwa nama-nama yang tercantum dalam lampiran telah lolos kualifikasi untuk menerima pendanaan hibah kompetitif dari Institusi;
- d. berdasarkan pertimbangan sebagaimana dimaksud pada huruf a dan b, dipandang perlu menetapkan Surat Keputusan Direktur Politeknik Harapan Bersama;
- Mengingat** : 1. Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional (Lembaran Negara Republik Indonesia Tahun 2003 Nomor 78, Tambahan Lembaran Negara Republik Indonesia Tahun 2003 Nomor 4301);
2. Undang-Undang Republik Indonesia Nomor 28 Tahun 2004 tentang Perubahan Undang-Undang Nomor 16 Tahun 2001 tentang Yayasan (Lembaran Negara Republik Indonesia Tahun 2004 Nomor 115, Tambahan Lembaran Negara Republik Indonesia Tahun 2004 Nomor 4430);
3. Undang-Undang Republik Indonesia Nomor 14 Tahun 2005 tentang Guru dan Dosen (Lembaran Negara Republik Indonesia Tahun 2005 Nomor 157, Tambahan Lembaran Negara Republik Indonesia Tahun 2005 Nomor 4586);
4. Undang-Undang Republik Indonesia Nomor 12 Tahun 2012 tentang Pendidikan Tinggi (Lembaran Negara Republik Indonesia Tahun 2012 Nomor 158, Tambahan Lembaran Negara Republik Indoneisa Tahun 2012 Nomor 5336);

5. Peraturan Pemerintah..

Dipindai dengan CamScanner

5. Peraturan Pemerintah Republik Indonesia Nomor 4 Tahun 2014 tentang Penyelenggaraan Pendidikan Tinggi dan Pengelolaan Perguruan Tinggi (Lembaran Negara Republik Indonesia Tahun 2014 Nomor 16, Tambahan Lembaran Negara Republik Indonesia Tahun 2014 Nomor 5500);
 6. Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 3 Tahun 2020 tentang Standar Nasional Pendidikan Tinggi (Berita Negara Republik Indonesia Tahun 2020 Nomor 47);
 7. Keputusan Menteri Pendidikan Nasional Republik Indonesia Nomor: 128/D/0/2002 tentang Pemberian Ijin Penyelenggaraan Program-Program Studi dan Pendirian Politeknik Harapan Bersama di Tegal yang Diselenggarakan oleh Yayasan Pendidikan Harapan Bersama di Tegal;
 8. Keputusan Menteri Hukum dan Hak Asasi Manusia Republik Indonesia Nomor: AHU-2674.AH.01.04 Tahun 2012 tentang pengesahan Yayasan Pendidikan Harapan Bersama (Tambahan Berita Negara Republik Indonesia Tanggal 20/6-2014 No. 49);
 9. Keputusan Menteri Riset, Teknologi, dan Pendidikan Tinggi Nomor: 231/KPT/I/2018 tentang Yayasan Pendidikan Harapan Bersama sebagai Badan Penyelenggara Politeknik Harapan Bersama;
 10. Surat Keputusan Yayasan Pendidikan Harapan Bersama Nomor 114.05/YPHB/XII/2020 tentang Statuta Politeknik Harapan Bersama;
- Memperhatikan : Surat Pemberitahuan Pusat Penelitian dan Pengabdian Masyarakat (P3M) Nomor: 064.03/P3M.PHB/III/2021 tentang pengajuan dan penerimaan proposal Penelitian dan Pengabdian Kepada Masyarakat Politeknik Harapan Bersama Semester Genap Tahun Akademik 2020/2021.

MEMUTUSKAN:

- Menetapkan : Surat Keputusan Direktur Politeknik Harapan Bersama tentang Penerima Pendanaan Oleh Institusi Untuk Hibah Kompetitif Penelitian dan Pengabdian Masyarakat Bagi Dosen Politeknik Harapan Bersama Tahun Anggaran 2020/2021.
- Pertama : Menetapkan nama yang tercantum dalam lampiran Keputusan ini sebagai Penerima Pendanaan Oleh Institusi Untuk Hibah Kompetitif Penelitian dan Pengabdian Masyarakat Bagi Dosen Politeknik Harapan Bersama Tahun Anggaran 2020/2021.
- Kedua :
 1. Pemberian bantuan dana penelitian minimal Rp. 2.000.000,- (Dua juta rupiah) per judul;
 2. Pemberian bantuan dana pengabdian kepada masyarakat minimal Rp. 2.000.000,- (Dua juta rupiah) per judul);
 3. Pembayaran dilakukan dengan 2 (dua) tahap, yaitu:
 - a. Pembayaran tahap I sebesar 60% dari total dana yang didapatkan setelah menyerahkan proposal dan perjanjian yang telah ditandatangani oleh Direktur Politeknik Harapan Bersama;
 - b. Pembayaran Tahap II sebesar 30% dari total dana yang didapatkan setelah menyerahkan laporan hasil; dan
 - c. 10% dari total dana yang didapatkan diserahkan kepada P3M.

- Ketiga : Dosen yang melaksanakan Penelitian dan/atau Pengabdian Kepada Masyarakat wajib menyerahkan laporan hasil kepada Direktur dan Wakil Direktur I melalui Pusat Penelitian dan Pengabdian Masyarakat (P3M), meliputi:
- a. Laporan penelitian sebanyak 2 (dua) eksemplar;
 - b. Softcopy Jurnal;
 - c. Softcopy.
- Keempat : Semua produk hasil penelitian dan pengabdian masyarakat termasuk Paten menjadi hak milik Politeknik Harapan Bersama.
- Kelima : Surat Keputusan ini berlaku sejak tanggal ditetapkan dan apabila di kemudian hari terdapat kekeliruan akan diadakan perbaikan sebagaimana mestinya.

Ditetapkan di: Tegal

Pada tanggal: 31 Mei 2021

Direktur,



Nizar Suhendra, S.E., MPP

NIPY.08.020.008

Lampiran: Surat Keputusan Direktur Politeknik
Harapan Bersama
Tentang : Penerima Pendanaan Oleh Institusi
Untuk Hibah Kompetitif Penelitian dan
Pengabdian Masyarakat Bagi Dosen
Politeknik Harapan Bersama Tahun
Anggaran 2020/2021 Semester Genap
Nomor : 038 .05/PHB/V/2021
Tanggal : 31 Mei 2021

**Daftar Penerima Bantuan Biaya Penelitian dan Pengabdian Kepada Masyarakat
Bagi Dosen Program Studi Sarjana Terapan Politeknik Harapan Bersama
Tahun Akademik 2020/2021 Semester Genap**

NO	KETUA	JUDUL	PRODI	SKEMA	NOMINAL
1	Ghea Dwi Rahmadiane, S.E., M.Si. Arifany Ferida, S.E., M.Si. Nizar Fahrezi Majid	Upaya Inovasi Dan Kreativitas Pelaku UMKM Kota Tegal Di Masa Pandemi Covid-19	Sarjana Terapan Akuntansi Sektor Publik	Penelitian	Rp. 3,271,000
2	Nurul Mahmudah, S.E., M. Si., AK, CA Yusri Anis Faidah, S.E., M. Si. Laelatul Maghfiroh	Transparansi Dan Akuntabilitas Dalam Pengelolaan Anggaran Pendapatan Dan Belanja Desa (Apbdes) (Studi Kasus: Desa Banjaranyar Kecamatan Brebek Kabupaten Brebes)	Sarjana Terapan Akuntansi Sektor Publik	Penelitian	Rp. 3,228,500
3	Dwi Intan Af'idah, S.T., M. Kom, Dairoh, M.Sc. Susi Nurindah sari	<i>Bidirectional Long Short Term Memory</i> Dan <i>Word2vec</i> Untuk Analisis Sentimen Terhadap Ulasan Destinasi Wisata Pulau Bali	Sarjana Terapan Teknik Informatika	Penelitian	Rp. 3,457,000
4	Dega Surono Wibowo, ST, M.Kom. Ardi Susanto, S.Kom., M.Cs. Khibar Pusaka	Analisis Dan Pengujian Celah Keamanan Pada Website DIV Teknik Informatika Politeknik Harapan Bersama	Sarjana Terapan Teknik Informatika	Penelitian	Rp. 2,814,000
5	Sharfina Febbi Handayani M.Kom. Riszi Wijayatun Pratiwi, S.Kom., M.Cs. Mulyana Putriyani	Analisis Sentimen Pada Data Ulasan <i>Twitter</i> Dengan Menggunakan <i>Long Short Term Memory</i>	Sarjana Terapan Teknik Informatika	Penelitian	Rp. 3,342,500
6	Dyah Apriliani, ST, M.Kom. Hepatika Zidny Ilmadina, S. Pd., M. Kom. Nurlaela	Sentiment Analysis Penilaian Toko Online Menggunakan <i>Naive Bayes</i> Dan <i>Neural Network</i>	Sarjana Terapan Teknik Informatika	Penelitian	Rp. 3,314,000
7	Taufiq Abidin, S.Pd, M. Kom. Slamet Wiyono, S.Pd., M. Eng. Agung Iswanto.	Implementasi Algoritma Nrf Dalam <i>Recommender System</i> <i>Berbasis Content</i> Dan <i>Collaborative Filtering</i> Sebagai Strategi Bisnis UMKM	Sarjana Terapan Teknik Informatika	Penelitian	Rp. 3,242,500

Handwritten signature

Lampiran: Surat Keputusan Direktur Politeknik
Harapan Bersama

Tentang : Penerima Pendanaan Oleh Institusi
Untuk Hibah Kompetitif Penelitian dan
Pengabdian Masyarakat Bagi Dosen
Politeknik Harapan Bersama Tahun
Anggaran 2020/2021 Semester Genap

Nomor : 098 .05/PHB/V/2021

Tanggal : 31 Mei 2021

13	Slamet Wiyono, S. Pd., M. Eng Dega Surono Wibowo, S.T., M. Kom, Riszi Wijayatun Pratiwi, S.Kom., M.Cs. Naimatul Maulidiyah Getar Dewantara Agung Iswanto	Pemanfaatan Teknik <i>Scraping</i> Data Untuk Perencanaan Usaha Jualan Online Menggunakan <i>Marketplace</i>	Sarjana Terapan Teknik Informatika	PKM	Rp. 2,900,000
14	Muhammad Fikri Hidayattullah, S.T., M.Kom. Dega Surono Wibowo, S.T., M. Kom. Ardi Susanto, S.Kom., M.Cs. Alfin Auzikri Wildan Sania Alfiansyah	Pengenalan <i>Software</i> Al- Mausu'ah Al-Hadits Bagi Santri Madrasah Fiqih Sumber Ilmu Dalam Melakukan Studi Takhrij Hadits	Sarjana Terapan Teknik Informatika	PKM	Rp. 2,787,500

Direktur,

Nizar Suhendra, S.E., MPP
NIPY.08.020.008

Lampiran 5. Draft Luaran Penelitian



Sentiment Analysis: Klasifikasi Rating Pelanggan Toko Online Menggunakan Metode Gaussian Naive Bayes dan Neural Network

Nama Penulis1^{a,*}, Nama Penulis2^b, Nama Penulis3^c

^a Sarjana Terapan Teknik Informatika, Politeknik Harapan Bersama

^b Sarjana Terapan Teknik Informatika, Politeknik Harapan Bersama

Naskah Diterima :xxx ; Diterima Publikasi : xxx

DOI : xxx

Abstract

The rapid development of technology has resulted in the development of buying and selling activities carried out online. This has triggered the emergence of various online stores in various marketplaces. Decision support systems play an essential role in the process of evaluating an online store service for buyers. This study aims to conduct sentiment analysis for online shop assessments using the Gaussian Naive Bayes Neural Network. The research method was carried out in several stages of research. The initial stage begins with collecting data sets that will be used in research in the form of online store assessment texts, data preprocessing, weighting, and model experiments using Naive Bayes and Neural Networks and evaluation of experimental results. The results of the study show that the performance of the Gaussian Naive Bayes model on sentiment analysis in the form of better classification of customer ratings, namely an accuracy of 69.01% on a limited dataset.

Keywords : *Sentiment analysis; online stores; Naive Bayes, Neural Networks*

Abstrak

Perkembangan teknologi yang semakin pesat mengakibatkan semakin berkembangnya kegiatan jual beli yang dilakukan secara *online*. Hal tersebut memicu munculnya berbagai toko *online* dalam berbagai *marketplace*. Sistem pendukung keputusan berperan penting dalam proses penilaian terhadap suatu pelayanan toko *online* bagi pembeli. Tujuan dari penelitian ini adalah melakukan sentiment analisis untuk penilaian toko online dengan menggunakan dan *Gaussian Naive Bayes* Neural Network. Metode penelitian dilakukan dalam beberapa tahapan penelitian. Tahap awal dimulai dengan proses pengambilan data set yang akan digunakan dalam penelitian berupa teks penilaian toko *online*, *praprocessing* data, pembobotan dan eksperimen model dengan menggunakan *Naive Bayes* Dan *Neural Network* serta evaluasi terhadap hasil eksperimen. Hasil penelitian menghasilkan performa model *Gaussian Naive Bayes* pada *sentiment analysis* berupa klasifikasi penilaian pelanggan lebih baik, yaitu akurasi sebesar 69,01% pada dataset yang terbatas.

Keywords: *Sentiment analysis; toko online; Naive Bayes, Neural Networks*

1. Pendahuluan

Teknologi memberikan dampak pada kehidupan manusia yang terkait di dalamnya menjadi lebih cepat dan mudah, serta terlihat sangat signifikan pada kegiatan bisnis. Salah satu sektor bisnis yang merasakan pesatnya perkembangan teknologi dan informasi adalah kegiatan jual beli barang (Moon, Talha, & Salehin, 2021). Transaksi yang dulunya hanya dilakukan dengan cara manual, saat ini sudah beralih ke transaksi digital dengan menggunakan toko *online*. Proses jual beli barang secara *online*

dengan menggunakan situs *online* biasa dikenal dengan istilah *marketplace*, dimana dalam toko *online* tersebut semua produk dapat di jual dan dibeli kapanpun dan dimanapun dengan berbagai kelebihan dan kekurangannya (Ishak, 2012).

Pelanggan diuntungkan dalam proses transaksi jual beli digital ini dikarenakan mereka dengan bebas dapat memilih dan membandingkan beberapa toko yang menjual barang serupa dengan memilih barang dengan harga yang paling *affordable*. Keuntungan lainnya pelanggan bebas memilih toko

dengan melihat reputasi toko *online* tersebut. Reputasi dari toko yang mereka kunjungi dihitung berdasarkan rekapitulasi penilaian yang diberikan oleh pembeli lain terhadap toko tersebut.

Semakin banyak pelanggan pada toko *online*, tidak sedikit pembeli yang juga mengalami beberapa keluhan atau pun beberapa komentar miring terkait dengan produk yang dibeli. Meskipun terdapat komentar bernada positif, hal negatif yang disampaikan akan mempengaruhi *rating* dari toko *online* tersebut (Nugroho, Chrisnanto, & Wahana, 2015).

Sentiment Analysis atau biasa dikenal dengan istilah analisis sentimen atau *opinion mining* merupakan salah satu proses untuk menemukan pendapat untuk berbagai topik atau teks yang disampaikan oleh pengguna (Gore, 2019). *Sentiment Analysis* bekerja dengan cara memproses, memahami, meng-ekstrak, dan mengolah data secara otomatis yang sifatnya berupa tekstual dimana nantinya akan menghasilkan suatu informasi yang terkandung di dalamnya. Salah satu kelebihan yang dimiliki *Sentiment Analysis* ini adalah dapat menghasilkan sistem yang digunakan sebagai pendukung keputusan yang dapat berupa klasifikasi sentimen, *review* sebuah objek, mendeteksi sebuah *spam* dan lainnya (Tripathy, Agrawal, & Rath, 2015). Dimana rekomendasi penilaian terhadap toko *online* saat ini banyak diinginkan, agar dapat memberikan sebuah pendukung keputusan terutama dalam melakukan pencarian toko *online* terbaik dan yang diinginkan.

Ada beberapa metode *Sentiment Analysis* yang sering digunakan, diantaranya adalah *Support Vector Machine*, *Naïve Bayes*, *Neural Network*, *Decision Tree*, *Bayesian Network* dan *Maximum Entropy* (Tripathy, Agrawal, & Rath, 2015). Dari beberapa tersebut, dalam penelitian ini akan digunakan metode *Gaussian Naïve Bayes* Dan *Neural Network*. Dari penelitian ini nantinya akan didapatkan model terbaik dalam mengklasifikasikan *rating* toko *online* berdasarkan hasil *review* pelanggan.

2. Kerangka Teori

2.1. Penelitian Terdahulu

Sistem rekomendasi bekerja dengan mengumpulkan informasi tentang preferensi pengguna terhadap produk baik secara eksplisit maupun implisit [4]. Sistem rekomendasi diklasifikasikan menjadi tiga kategori: *content-based*, *collaborative filtering (CF)*, dan *hybrid* [5]. Metode *content-based* menghasilkan rekomendasi berdasarkan profil user dan kesamaan dari deskripsi produk. CF menghasilkan rekomendasi untuk user berdasarkan riwayat transaksi (penilaian) user yang lainnya. Hal inilah yang menyebabkan proses rekomendasi yang lebih kompleks karena

membutuhkan data user lainnya, bukan berdasarkan deskripsi dari produk. Sedangkan metode hybrid adalah metode gabungan antara CF dan content-based. Metode CF menjadi metode yang populer dalam pembuatan sistem rekomendasi [10]. Namun demikian, walaupun CF merupakan metode yang populer, CF memiliki permasalahan utama yaitu *cold start* dan *sparsity* [11]. *Cold start* adalah keadaan kekurangan data tentang entitas baru baik produk maupun user baru, sedangkan *sparsity* adalah keadaan dimana user menilai sedikit produk.

Penelitian untuk mengatasi permasalahan *cold start* CF dilakukan oleh Lika dkk [12] dan Uyangoda dkk [13] dengan menggunakan pendekatan berbasis *rating*. Lika dkk menggunakan C4.5 dan Naïve Bayes untuk melakukan klasifikasi dan Nearest Neighbors untuk mencari similaritas. Metode yang ditawarkan Lika dkk menggunakan informasi demografi untuk melakukan prediksi sehingga tidak ada penentuan bobot. Uyangoda dkk menggunakan pendekatan profil user yaitu dengan membuat profil setiap user dari data yang tersimpan. Profil tersebut kemudian digunakan untuk mencari similaritas. Metode yang ditawarkan mirip seperti Lika dkk yaitu menggunakan informasi demografi, belum ada penentuan bobot.

Penelitian untuk mengatasai *sparsity* pada CF dilakukan oleh Tang dan Tong dengan menggunakan pendekatan berbasis ranking [14], yaitu dengan menawarkan metode BordaRank. Metode yang ditawarkan Tang dan Tong menggabungkan item CF dan Borda. Adapun penelitian lainnya yang menggunakan pendekatan berbasis ranking pernah dilakukan dengan menggunakan metode agregasi *Copeland Score* [15]. Penelitian yang dilakukan tidak melibatkan data *rating* dalam agresasi. Selain itu, proses rekomendasi membutuhkan waktu yang cukup lama.

Penelitian berbasis ranking yang telah dilakukan dengan metode Borda dan Copeland belum memanfaatkan data *rating* dengan maksimal. Penelitian yang dilakukan hanya menggabungkan beberapa metode dan data *rating* belum termanfaatkan secara maksimal. Kedua metode tersebut menggunakan data *rating* hanya untuk penyusunan user preference profile dengan mengurutkan *rating* terbesar ke *rating* terkecil. Salah satu penelitian yang berusaha menggunakan data *rating* secara maksimal telah dilakukan oleh Lestari dengan menawarkan metode NRF (*normalized rating frequency*) [16]. Metode NRF dikembangkan dalam rangka mengatasi kelemahan dari metode dasar Borda dan Copeland. Proses kerja metode NRF dilakukan dengan melakukan normalisasi frekuensi *rating* dalam proses agresasi untuk mendapatkan ranking dari produk. Namun demikian, proses normalisasi yang dilakukan

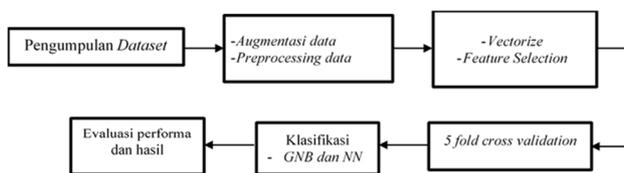
berdasarkan total frekuensi tiap level rating, bukan berdasarkan total frekuensi tiap produk. Dengan demikian, diperlukan peninjauan kembali mengenai metode yang ditawarkan tersebut, apakah penggunaan total frekuensi tiap level rating untuk normalisasi sudah tepat atau belum.

3. Metode

Penelitian ini dimulai dengan mengidentifikasi dan merumuskan permasalahan berdasarkan latar belakang penelitian akan dilakukan. Melakukan studi literatur untuk mendapat metode terbaik untuk memecahkan permasalahan yang diangkat. Penelitian kemudian dilanjutkan dengan pengumpulan dataset secara manual berdasarkan masalah yang telah dirumuskan. Dataset ini dikumpulkan komentar dan rating pelanggan dari perusahaan *e-commerce Shopee* Indonesia, pada toko *Elzatta Official*.

Pengolahan data selanjutnya diimplementasikan menggunakan *Python* dengan memanfaatkan *Keras library* dengan *Tensorflow backend* serta diimplementasikan menggunakan *Google Colab* dengan *single 12GB NVIDIA Tesla K80 GPU*. Implementasi model juga menggunakan *library* pendukung seperti *numpy, sklearn, matplotlib, panda, nltk, bs4, spacy, re, dan textblob*.

Pengolahan dataset kemudian dilanjutkan ke proses augmentasi dan preprocessing data. Selanjutnya untuk melakukan sebuah *sentiment analysis* dilakukan vektorisasi kalimat menggunakan *TF-IDF Vectorizer*. Setelah mendapatkan vector untuk masing-masing komentar pelanggan, peneliti mengklasifikasikan rating yang diberikan pelanggan berdasarkan dari hasil komentarnya. Penelitian dibagi menjadi 2 skenario utama, yaitu membandingkan hasil klasifikasi model *Gaussian Naive Bayes (GNB)* dengan *Neural Network (NN)*. Kerangka metodologi penelitian secara detail dijelaskan pada Gambar 1 berikut ini.



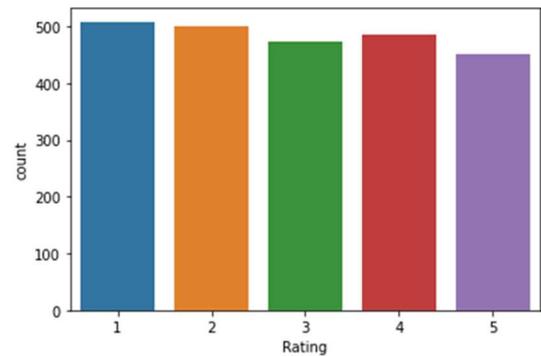
Gambar 1. Kerangka Metodologi Penelitian

4. Hasil dan Pembahasan

4.1. Pelatihan dan Evaluasi Model

Sentiment Analysis ini dilakukan untuk mengklasifikasikan rating yang diberikan pelanggan berdasarkan komentarnya. Rating dibagi menjadi 5 kelas, yaitu rating 1, rating 2, rating 3, rating 4, dan rating 5. Dataset asli berjumlah 604 komentar.

Sehingga untuk menunjang pelatihan data dengan *deep learning*, dilakukan *data augmentation* dengan cara menduplikasi data sehingga dataset total adalah 2416 komentar, digambarkan pada grafik Gambar 2. Sebelum menuju tahap klasifikasi dilakukan langkah *text processing* seperti menghilangkan tautan html, menghapus karakter di luar teks, menghilangkan angka dan symbol, menghilangkan *multiple* spasi, dan menghapuskan *emoticon*.



Gambar 2. Grafik Tiap Kelas pada Dataset

Untuk melakukan klasifikasi pada teks, selanjutnya dilakukan vektorisasi dengan menggunakan algoritma TF-IDF. Optimasi fitur yang dipilih dilakukan dengan mengimplementasikan *feature selection chi2*. Hal tersebut dilakukan untuk mengurangi waktu komputasi dengan hanya memilih fitur yang merepresentasikan data saja. Berdasarkan hasil seleksi fitur tersebut diperoleh beberapa contoh *unigram* dan *bigram* komentar yang berkorelasi dengan rating yang diberikan oleh pelanggan. Lebih lengkap dijelaskan pada Tabel 1.

Tabel 1. Contoh *Unigram* dan *Bigram* Komentar Pelanggan

Rating	Unigram	Bigram
1	<ul style="list-style-type: none"> • buruk • tidak • dikirim • chat • pcs 	<ul style="list-style-type: none"> • dikirim tidak • kecewa banget • tdk sesuai • pengiriman lama • di chat
2	<ul style="list-style-type: none"> • lain • padahal • tipis • pengiriman • model 	<ul style="list-style-type: none"> • tipis banget • ga sesuai • warna tidak • sangat lama • ga ada
3	<ul style="list-style-type: none"> • harusnya • walau • ukurannya • sayang • si 	<ul style="list-style-type: none"> • sedikit kecewa • pengemasan lama • bahannya bagus • tidak ada • bahannya tipis

4	<ul style="list-style-type: none"> • rapih • oke • enak • cepat • lumayan 	<ul style="list-style-type: none"> • sesuai pesanan • harga segitu • produk baik • mudah dibentuk • pengiriman cepat
5	<ul style="list-style-type: none"> • lembut • nyaman • makasih • terimakasih • bagus 	<ul style="list-style-type: none"> • sudah sampai • terima kasih • barang sesuai • barangnya bagus • paket sudah

Berdasarkan hasil *unigram* dan *bigram* dapat dilihat beberapa kata yang sangat berkorelasi dengan pemberian rating pelanggan. Contohnya pada kalimat “kecewa banget” diidentifikasi bahwa pelanggan memberikan rating 1 pada toko. Performa toko dinilai dengan rating 2 ditunjukkan dengan adanya komentar “ga sesuai” maupun “sangat lama”. Penilaian pada rating 3 dianggap yang paling netral dengan komentar pelanggan berupa “sedikit kecewa”; “pengemasan lama”; “ga ada” dan komentar yang hampir menuju komentar positif yaitu “bahannya bagus”. Pemberian rating tidak tinggi tersebut memperlihatkan bahwa pelanggan sebenarnya kecewa dengan pelayanan toko akan tetapi barang yang diterima sudah sesuai ataupun sebaliknya. Komentar positif akan terlihat di *unigram* dan *bigram* rating 4 dan 5, contohnya ada beberapa apresiasi pada toko yang menunjukkan kepuasan pelanggan seperti “sesuai pesanan”; “pengiriman cepat”; “terima kasih”; dan “barangnya bagus”. Komentar tersebut menunjukkan kepuasan pelanggan akan pelayanan dan barang yang dijual oleh toko tersebut.

Setelah melihat beberapa kata unik yang menjadi penentu pada klasifikasi komentar dilanjutkan ke proses pelatihan data. Pada proses ini dataset yang telah menjadi vector (melalui algoritma TF-IDF) kemudian dibagi menjadi data latih dan data uji masing-masing 80% dan 20%. Pelatihan kemudian menerapkan *5-fold cross validation* pada seluruh skenario metode *Gaussian Naïve Bayes* (GNB) dan *Neural Network* (NN).

Implementasi metode GNB menggunakan modul *sklearn* dengan melakukan *min-max scaling* agar data normal. Sedangkan pada NN dilakukan pembentukan *layer* menggunakan *embedding layer*. Model diterapkan *categorical cross entropy* sebagai *loss function* dan *softmax* sebagai *activation function* dikarenakan merupakan eksperimen klasifikasi *multiclass*. Model kemudian dikompilasikan dengan *adam optimizer* pada 20 *epoch* dengan *early-stopping*. Hasil evaluasi performa kedua model kemudian dibandingkan untuk mengetahui metode terbaik dalam klasifikasi rating berdasarkan komentar pelanggan toko.

4.1. Perbandingan Metode GNB dengan NN

Pada proses Pelatihan pada implementasi kedua model klasifikasi diperoleh hasil *precision*, *recall*, dan *f-score*. Secara keseluruhan perbandingan

evaluasi performa kedua model jika ditinjau dari hasil klasifikasi tiap kelas dapat dijelaskan pada Tabel 2.

Tabel 2. Evaluasi Performa Kedua Metode Klasifikasi

Rating	Gaussian Naïve Bayes			Neural Netwo	
	Precision	Recall	f-1 Score	Precision	Recall
1	0,65	0,81	0,72	0,59	0,52
2	0,80	0,53	0,64	0,59	0,53
3	0,96	0,52	0,68	0,80	0,52
4	0,67	0,61	0,64	0,67	0,61
5	0,59	1,00	0,75	0,59	0,53

Secara keseluruhan dapat terlihat bahwa beberapa kali GNB maupun NN akan lebih baik pada salah satu aspek. Hal ini juga ditunjukkan dengan perbedaan akurasi yang tidak terlalu jauh, yaitu 69,01% pada metode GNB dan 63,52% pada NN meskipun GNB terlihat lebih unggul. Sedangkan untuk melihat performa model secara keseluruhan, dapat dilihat dari hasil nilai rata-rata *accuracy*, *precision*, *recall*, dan *f-score* yang disajikan pada Tabel 3.

Tabel 3. Perbandingan Hasil Evaluasi Rata-rata *accuracy*, *precision*, *recall*, dan *f-score* pada kedua metode

Metode	Accuracy (%)	macro avg		
		precision	recall	f-sco
Gaussian Naïve Bayes	69,01	0,74	0,69	0,6
Neural Network	63,52	0,65	0,54	0,6

Secara performa keseluruhan model diperoleh bahwa hasil evaluasi performa menggunakan GNB lebih besar dibandingkan dengan menggunakan NN. Hal tersebut dilihat dari nilai *accuracy*, *precision*, *recall*, maupun *f-score* yang lebih tinggi.

4.2. Pembahasan

Dataset *private* yang digunakan pada penelitian ini adalah kumpulan komentar dan rating pelanggan toko *online Elzatta Official* yang diakuisisi secara manual pada *marketplace Shopee* Indonesia. Dataset kemudian dianalisa dengan menggunakan metode *Gaussian Naive Bayes* (GNB) dan *Neural Network* (NN). Implementasi kedua metode tersebut menghasilkan hasil akurasi masing-masing 69,01% untuk GNB dan 63,52% untuk NN. Metode GNB diimplementasikan dengan menggunakan modul *sklearn*, sedangkan NN dengan membangun *layer embedded* dengan melakukan vektorisasi pada komentar pelanggan. *Pre-processing* data juga dilakukan untuk membuat data dapat diolah secara maksimal. Penghilangan simbol, karakter,

stemming kalimat juga dilakukan. Akan tetapi *stemming* dilakukan dengan metode biasa, tidak dikhususkan pada modul bahasa Indonesia. *Stop word* yang dilakukan adalah dengan metode biasa tanpa mengkhususkan pada kata Bahasa Indonesia.

Implementasi pada NN yang merupakan algoritma *deep learning* seharusnya dapat memberikan performa yang lebih baik dibandingkan dengan GNB. Hal tersebut tidak dapat dicapai dalam penelitian ini dikarenakan ketersediaan variasi dataset yang terbatas, yaitu hanya 604 komentar untuk 5 kelas data rating 1-5. Meskipun pada proses pelatihan telah kita terapkan augmentasi data, hal tersebut tetap menjadi keterbatasan fitur pada pengolahan data. Dalam hal ini, *layer* hanya akan mempelajari fitur yang sama dan dilakukan berulang untuk masing-masing dataset karena proses augmentasi tersebut. Oleh karena itu pada penelitian ini performa NN tidak lebih bagus dibandingkan dengan GNB pada klasifikasi rating pelanggan berdasarkan komentar yang diberikan. GNB dapat memberikan hasil yang maksimal dikarenakan metode ini

4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan, beberapa simpulan yang dapat diperoleh yaitu:

- a. Identifikasi kepuasan pelanggan yang menjadi sebuah penilaian (rating) dapat dilakukan dengan metode *Gaussian Naive Bayes* dan *Neural Network*.
- b. Performa model *Gaussian Naive Bayes* pada *sentiment analysis* berupa klasifikasi penilaian pelanggan lebih baik, yaitu akurasi sebesar 69,01% pada dataset yang terbatas.

Daftar Pustaka

- A. Ishak, "Analisis Kepuasan Pelanggan Dalam Belanja Online: Sebuah Studi Tentang Penyebab (Antecedents) dan Konsekuensi (Consequents)," *Jurnal Siasat Bisnis*, Vol. 16 No. 2, , Hal : 141-154, Juli 2012.
- D. G. Nugroho, Y. H. Chrisnanto, and A. Wahana, "Analisis Sentimen pada Jasa Ojek Online Menggunakan Metode Naive Bayes," *Jur. Inform. Fak. Mat. dan Ilmu Pengetah. Alam Univ. Jenderal Achmad Yani*, pp. 156–161, 2015.
- B. S. D and P. D. Gore, "Sentiment Analysis on Twitter Data Using Support Vector Machine," *International Journal of Computer Science Trends and Technology (IJCSST) – Volume 4 Issue 3*, pp. 365–370, May - Jun 2016
- A. Tripathy, A. Agrawal, and S. K. Rath, "Classification of Sentimental Reviews Using Machine Learning Techniques," *Procedia Comput. Sci.*, vol. 57, pp. 821–829, 2015.

- Muljono, D.P Artanti, A. Syukur, dkk "Analisis Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naive Bayes," *Stimik Admaluhur Pangkalpinang*, 8-9 Maret 2018
- E.H. Muktafin, Kusriani, E.T Luthfi, "Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing," *Jurnal Eksplora Informatika*. Vol 1 no 10, pp. 32-42, 2020
- M. A. Asuja, Saniati, "Analisis Sentimen Tweet Menggunakan Backpropagation Neural Network," *Jurnal Teknoinfo*. Vol 11 No 2, pp 23-28, 2016
- D. Apriliani, T. Abdidin, E. Sutanta, A. Hamzah, O. Somantri " Sentiment analysis for Indonesia hotel services review using optimized neural network," *ICCGANT 2019*
- F. Z. Tala, "A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia," *M.Sc. Thesis, Append. D*, vol. pp, pp. 39–46, 2003.
- Chen, Z. Zhang, J. Long, and H. Zhang, "Turning from TF-IDF to TF-IGM for term weighting in text classification," *Expert Syst. Appl.*, vol. 66, pp. 1339–1351, 2016.
- Liu, Bing. (2012). *Sentiment Analysis and Opinion Mining*. USA : Morgan & Claypool Publishers.
- Ma, Jian dkk. (2012). *An Ontology-Based Text-Mining Method to Cluster Proposals for Research Project Selection*. *IEEE Trans. Syst., Man, Cybern. Part A, Syst & Humans* (vol. 42).
- Han, Jiawei., Kamber, Micheline., & Pei, Jian. (2012). *Data Mining Concepts and Techniques* (3rd ed.). USA : Simon Frase University
- Blachier, M., Leleu, H., Peck-Radosavljevic, M., Valla, D. C., & Roudot-Thoraval, F. (2013). *The burden of liver disease in Europe: A review of available epidemiological data*. In *Journal of Hepatology*. <https://doi.org/10.1016/j.jhep.2012.12.005>