

**ANALISIS KEPUASAN PELANGGAN PADA PELAYANAN
E-COMMERCE TOKOPEDIA DAN SHOPEE PADA TWITTER
MENGUNAKAN *RECURRENT NEURAL NETWORK***

(Skripsi)

Oleh

**AJIJAH
1617051136**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022**

ABSTRAK

ANALISIS KEPUASAN PELANGGAN PADA PELAYANAN E-COMMERCE TOKOPEDIA DAN SHOPEE PADA TWITTER MENGUNAKAN *RECURRENT NEURAL NETWORK*

Oleh

Ajjah

Data di Twitter dapat digunakan untuk melihat kepuasan pelayanan suatu *e-commerce* karena para pengguna memberi suatu tulisan atau *tweet* mengenai pelayanan *e-commerce*. Analisis yang akan dilakukan dapat bermanfaat untuk mengetahui pendapat pengguna Twitter terkait kepuasan pelanggan pada suatu *e-commerce*. Objek yang dipilih dalam penelitian ini yaitu *platform e-commerce* perusahaan teknologi Indonesia Tokopedia dan perusahaan teknologi asing Shopee. Penelitian ini membandingkan akurasi, presisi, *recall*, dan *f1-score* antara dua algoritme, yaitu RNN dan *Naïve Bayes*. Data yang digunakan berbahasa Indonesia, dengan jumlah sebanyak 1002 data Tokopedia dan 1000 data Shopee, dibagi menjadi dua jenis data sebanyak 80% data latih dan 20% data uji. Diperoleh hasil dari analisis dengan metode RNN pada data Tokopedia menggunakan algoritme LSTM akurasi sebesar 78%, presisi 78%, *recall* 78%, dan *f1-score* 78%. Pada data Shopee menggunakan *Simple RNN* menghasilkan nilai akurasi 58%, presisi 58%, *recall* 58%, dan *f1-score* 58%.

Kata kunci: Analisis Sentimen, Klasifikasi, RNN, Shopee, Teks *mining*, Tokopedia, Twitter.

ABSTRACT

ANALYSIS OF CUSTOMER SATISFACTION WITH TOKOPEDIA AND SHOPEE E-COMMERCE SERVICES ON TWITTER USING RECURRENT NEURAL NETWORK

By

Ajjah

Data on Twitter can be used to see the satisfaction of an e-commerce service because users give an article or tweet about e-commerce services. The analysis that will be carried out can be useful for knowing the opinions of Twitter users regarding customer satisfaction in an e-commerce. The objects chosen in this study are the e-commerce platform of the Indonesian technology company Tokopedia and the foreign technology company Shopee. This study compares the accuracy, precision, recall, and f1-score between the two RNN algorithms and the Naïve Bayes algorithm. The data is in Indonesian language, totaling 1002 Tokopedia data and 1000 Shopee data, were divided into two types of data, 80% training data and 20% test data. The results obtained from the analysis using the RNN method on Tokopedia data using the LSTM algorithm is 78% accuracy, 78% precision, 78% recall, and 78% f1-score, where the Shopee data using Simple RNN resulted 58% accuracy, 58% precision, 58% recall, and an f1-score of 58%.

Keywords: Clasification, RNN, Shopee, Sentiment analysis, Text mining Tokopedia, Twitter.

**ANALISIS KEPUASAN PELANGGAN PADA PELAYANAN
E-COMMERCE TOKOPEDIA DAN SHOPEE PADA TWITTER
MENGUNAKAN *RECURRENT NEURAL NETWORK***

Oleh

AJIJAH

Skripsi

**Sebagai Salah Satu Syarat Untuk Mencapai Gelar
SARJANA KOMPUTER**

Pada

**Jurusan Ilmu Komputer
Fakultas Matematika dan Ilmu Pengetahuan Alam**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022**

Judul Skripsi : ANALISIS KEPUASAN PELANGGAN PADA PELAYANAN E-COMMERCE TOKOPEDIA DAN SHOPEE PADA TWITTER MENGGUNAKAN RECURRENT NEURAL NETWORK.

Nama Mahasiswa : Ajjah

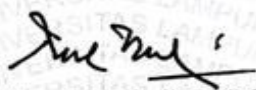
Nomor Pokok Mahasiswa : 1617051136

Program Studi : Ilmu Komputer

Fakultas : Matematika dan Ilmu Pengetahuan Alam

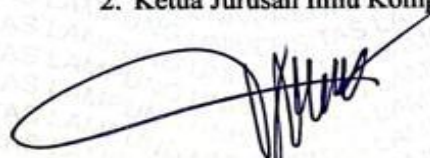


1. Komisi Pembimbing


Dr. Ir. Kurnia Muludi, M.S.Sc.
NIP 19640616 198902 1 001


Dewi Asiah Shofiana, S.Komp., M.Kom.
NIP 19950929 202012 2 030

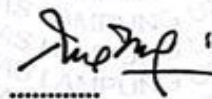
2. Ketua Jurusan Ilmu Komputer


Didik Kurniawan, S.Si., M.T.
NIP 19800419 200501 1 004

MENGESAHKAN

1. Tim Penguji

Ketua Penguji : Dr. Ir. Kurnia Muludi, M.S.Sc.



Sekretaris : Dewi Asiah Shofiana, S.Komp., M.Kom.



Anggota : Favorisen Rosyking Lumbanraja, Ph.D.



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Satripto Dwi Yuwono, S.Si., M.T.
NIP 197407052000031001



Tanggal Lulus Ujian : 20 Juli 2022

PERNYATAAN

Saya yang bertanda tangan di bawah ini, menyatakan bahwa skripsi saya yang berjudul “Analisis Kepuasan Pelanggan Pada Pelayanan *E-commerce* Tokopedia dan Shopee Pada Twitter Menggunakan *Recurrent Neural Network*” merupakan karya saya sendiri dan bukan karya orang lain. Semua tulisan yang tertuang di skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila di kemudian hari terbukti skripsi saya merupakan hasil penjiplakan atau dibuat orang lain, maka saya bersedia menerima sanksi berupa pencabutan gelar yang telah saya terima.

Bandar Lampung, 20 Juli 2022



Ajijah
NPM 1617051136

RIWAYAT HIDUP



Penulis dilahirkan pada tanggal 11 September 1997 di Bandar Lampung, dengan Ibu bernama Sutira dan Ayah bernama Kusno. Penulis menyelesaikan pendidikan formal Sekolah Dasar di SDN 2 Rajabasa Jaya Bandar Lampung tahun 2010, menyelesaikan Sekolah Menengah Pertama di SMPN 3 Natar Lampung Selatan tahun 2013, kemudian melanjutkan jenjang Sekolah Menengah Kejuruan di SMKN 2 Bandar Lampung diselesaikan pada tahun 2016.

Pada tahun 2016, penulis terdaftar sebagai mahasiswi Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam melalui jalur PMPAP. Selama menjadi mahasiswi, penulis aktif di Organisasi Himpunan Mahasiswa Ilmu Komputer (HIMAKOM) FMIPA Universitas Lampung sebagai anggota bidang Kewirausahaan. Pada tahun 2019, penulis melakukan Kerja Praktik di Sekolah Polisi Negara (SPN) Kemiling Polda Lampung dan melakukan kegiatan Kuliah Kerja Nyata (KKN) selama 40 hari di Pekon Lombok Selatan, Kecamatan Lombok Seminung, Kabupaten Lampung Barat.

PERSEMBAHAN

Puji dan syukur saya panjatkan kepada Allah SWT atas segala berkah-Nya sehingga skripsi ini dapat diselesaikan.

Teruntuk yang tercinta kupersembahkan skripsi ini, hasil dukungan dan semangat dari kedua orang tua Mamak dan Bapak ku.

Terima kasih untuk cinta kasihmu untuk ku, anakmu..

Terima kasih telah menyayangi, menjaga, mengasahi, melindungiku sampai detik ini dengan kesabaran, tetes keringat serta air mata.

Kakak-kakak, Mbak, dan Adik ku yang telah memberikan semangat untuk menyelesaikan pendidikan ini.

Keluarga Ilmu Komputer 2016

Serta Almamater tercinta,

Universitas Lampung.

MOTTO

“Sesungguhnya sesudah kesulitan itu ada kemudahan”

Q.S. Al Insyirah : 6

“Life is tough, and things don't always work out well, but we should be brave and go on with our lives”

-Suga

“Skripsi yang baik adalah skripsi yang selesai”

SANWACANA

Assalamualaikum wr,wb.

Alhamdulillah, segala puji bagi Allah SWT yang telah melimpahkan rahmat, hidayah, kesehatan, dan segala karunia-Nya sehingga penulis dapat menyelesaikan penulisan skripsi yang berjudul “Analisis Kepuasan Pelanggan Pada Pelayanan *E-commerce* Tokopedia dan Shopee Pada Twitter Menggunakan *Recurrent Neural Network*” dengan baik. Terima kasih penulis ucapkan kepada semua pihak yang telah membantu dan berperan dalam menyusun skripsi ini, antara lain:

1. Kedua orang tua tercinta, Mamak dan Bapak yang telah memberikan doa, kasih sayang, dukungan, dan semangat yang tak terhingga serta memfasilitasi kebutuhan untuk menyelesaikan skripsi ini.
2. Bapak Dr. Ir. Kurnia Muludi, M.S.Sc. sebagai pembimbing utama yang telah meluangkan waktu membimbing, memotivasi serta memberikan ide, kritik, dan saran selama masa perkuliahan dan penyusunan skripsi sehingga penulis bisa sampai di tahap ini.
3. Ibu Dewi Asiah Shofiana, S.Komp., M.Kom. sebagai pembimbing kedua yang telah membimbing dan memberikan bantuan ide, kritik serta saran dalam penyusunan skripsi ini sehingga penulis bisa sampai di tahap ini.
4. Bapak Favorisen R. Lumbanraja, Ph.D. sebagai pembahas, yang telah memberikan komentar dan masukan yang bermanfaat untuk perbaikan dalam penyusunan skripsi ini.
5. Bapak Dr. Eng. Suropto Dwi Yuwono, S.Si., M.T. sebagai Dekan FMIPA Universitas Lampung.

6. Bapak Didik Kurniawan, S.Si., M.T. selaku Ketua Jurusan Ilmu Komputer FMIPA Universitas Lampung yang telah banyak membantu selama perkuliahan.
7. Bapak Dr. rer. nat. Akmal Junaidi, M.Sc. sebagai Sekretaris Jurusan Ilmu Komputer FMIPA Universitas Lampung yang telah banyak membantu selama perkuliahan.
8. Bapak dan Ibu Dosen Jurusan Ilmu Komputer yang telah memberikan ilmu dan pengalaman hidup selama penulis menjadi mahasiswa Ilmu Komputer 2016.
9. Sahabat tercinta semasa SMK Iffa, Revily, Silfia, dan Wulan yang tak henti menemani, memberikan semangat dan menghibur penulis dalam sedih dan bahagia.
10. Sahabat seperjuangan di Ilkomp Amrina, Friska, Anes, Anggie, Dewi, Elva, Eggi, Dara, Akbar, dan Ilham yang selalu memberikan kegembiraan dan semangat di saat penulis merasa jenuh dan sedih dalam penyusunan skripsi ini.
11. Keluarga Ilmu Komputer 2016 yang tidak bisa disebutkan satu per satu, terima kasih atas kebersamaan selama ini.
12. Ibu Ade Nora Maela yang telah membantu segala urusan administrasi penulis di Jurusan Ilmu Komputer.
13. Mas Naufal yang telah menjaga MIPA terpadu dan membantu untuk menyiapkan ruang seminar.
14. Semua pihak yang telah membantu penulis yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, akan tetapi sedikit harapan semoga skripsi ini bermanfaat bagi perkembangan ilmu pengetahuan terutama bagi adik tingkat Ilmu Komputer.

Bandar Lampung, 20 Juli 2022



Ajjah
NPM 1617051136

DAFTAR ISI

Halaman

DAFTAR TABEL	xv
---------------------------	-----------

DAFTAR GAMBAR.....	xv
---------------------------	-----------

I. PENDAHULUAN	1
-----------------------------	----------

1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5

II. TINJAUAN PUSTAKA	6
-----------------------------------	----------

2.1 Analisis.....	6
2.2 Analisis Sentimen.....	8
2.3 Kepuasan Pelanggan	8
2.4 Twitter	8
2.5 <i>E-commerce</i>	9
2.6 Tokopedia.....	9
2.7 Shopee	10
2.8 <i>Recurrent Neural Network (RNN)</i>	10
2.9 <i>Long Short-Term Memory (LSTM)</i>	12
2.10 <i>Naïve Bayes</i>	13
2.11 <i>Confusion Matrix</i>	13
2.12 <i>Word Embedding</i>	15
2.13 <i>Word2vec</i>	15
2.14 Bahasa Pemrograman Python.....	16
2.15 <i>Deep Learning</i>	16
2.16 <i>Machine Learning</i>	16

III. METODE PENELITIAN.....	17
------------------------------------	-----------

3.1 Tempat dan Waktu Penelitian	17
---------------------------------------	----

3.2	Alat dan Bahan	17
1.	Perangkat keras (<i>Hardware</i>)	17
2.	Perangkat lunak (<i>Software</i>)	17
3.3	Tahapan Penelitian	20
V.	SIMPULAN DAN SARAN.....	39
5.1	Simpulan.....	39
5.2	Saran.....	39
	DAFTAR PUSTAKA	41

DAFTAR TABEL

Tabel	Halaman
Tabel 1. Penelitian terdahulu terkait analisis sentimen.....	6
Tabel 2. <i>Confusion matrix</i>	14

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1. Metode <i>Simple Recurrent Neural Network</i>	12
Gambar 2. Alur tahapan penelitian analisis pelayanan Tokopedia dan Shopee ...	20

I. PENDAHULUAN

1.1 Latar Belakang

Semakin berkembang internet pada zaman yang sudah *modern* ini, manusia dituntut untuk mampu mengikuti perkembangan teknologi. Dalam perkembangannya, internet bukan hanya digunakan sebagai media informasi dan komunikasi saja, adapun dampak pengguna internet, yaitu pengguna media sosial yang terus meningkat (Kautsar, 2018). Perkembangan dunia internet juga mempengaruhi proses transaksi jual dan beli yang awalnya bersifat konvensional berubah menjadi *modern* dengan bantuan internet, yang disebut juga sebagai *e-commerce* (Syah et al., 2017). Salah satu media sosial yang mempengaruhi transaksi *e-commerce* tersebut adalah Twitter, karena salah satu media sosial yang sangat populer di Indonesia. Menurut Kementerian Komunikasi dan Informasi (Kemenkominfo), Indonesia menempati peringkat 5 terbanyak pengguna Twitter terbesar di dunia. Posisi Indonesia hanya kalah dari USA, Brazil, Jepang, dan Inggris (Kominfo, 2019).

Data di Twitter dapat digunakan untuk melihat kepuasan pelayanan suatu *e-commerce* karena para pengguna memberi suatu tulisan atau *tweet* mengenai pelayanan *e-commerce*. Data pengguna Twitter tersebut dapat dimanfaatkan oleh *e-commerce* untuk menganalisis pengembangan sistem yang lebih baik. Analisis yang akan dilakukan dapat bermanfaat untuk mengetahui pendapat pengguna Twitter terkait kepuasan pelanggan pada suatu *e-commerce*. Ulasan tersebut dapat dianalisis untuk mencari kelemahan pelayanan *e-commerce* dan jika terdapat keunggulan pelayanan dapat dipertahankan oleh *e-commerce* tersebut. Sosial media seperti Twitter memiliki peran besar dalam mempengaruhi aspek dari segi sosial, komunikasi, psikologi,

marketing, dan politik, oleh karena itu data yang didapat untuk penelitian ini dari Twitter.

Saat ini ada beberapa *e-commerce* yang berkembang di Indonesia seperti Shopee, Lazada, Tokopedia, dan Bukalapak. Objek yang dipilih dalam penelitian ini, yaitu *platform e-commerce* perusahaan teknologi Indonesia Tokopedia dan perusahaan teknologi asing Shopee. Tokopedia merupakan toko pertama dalam *website* yang menyediakan peluang bisnis *online*. Tokopedia berdiri pada 6 Februari 2009 oleh Wiliam Tanuwijaya (Lupi dan Nurdin, 2016). Shopee adalah *platform e-commerce* yang berkantor pusat di Singapura di bawah SEA Group. Shopee pertama kali diluncurkan di Singapura pada tahun 2015, dan sejak itu memperluas jangkauannya ke Malaysia, Thailand, Taiwan, Indonesia, Vietnam, dan Filipina (Indraloka dan Santosa, 2017). Tokopedia dan Shopee sangat bersaing yang ditunjukkan dengan adanya banyak promo pada kedua *e-commerce*. Promo tersebut mencakup program gratis ongkos kirim barang dan adanya *event* yang disiarkan secara langsung di televisi sebagai cara memperkenalkan pada khalayak ramai tentang Tokopedia dan Shopee di Indonesia. Kedua *e-commerce* menggunakan *brand ambassador* artis yang terkenal dalam negeri maupun luar negeri sebagai cara menarik perhatian pengguna kedua *e-commerce*.

Deep Learning merupakan salah satu implementasi dari *Machine Learning* adalah ilmu yang mempelajari tentang beberapa tingkat representasi dan abstraksi yang membantu memahami data seperti gambar, suara, dan teks (Zulfa dan Winarko, 2017). Sedangkan *Machine Learning* adalah bagian *artificial intelligence* yang banyak digunakan untuk memecahkan berbagai masalah (Roihan et al., 2020). Metode *Deep Learning* dan *Machine Learning* digunakan untuk dilakukan analisis sentimen data Twitter kepuasan pelanggan antara Tokopedia dan Shopee. Sumber data pada Twitter dipilih untuk mengetahui bagaimana kepuasan pelanggan terhadap pelayanan *e-commerce* di Indonesia guna mendapatkan hasil penilaian kepuasan pelanggan pada *e-commerce* Tokopedia dan Shopee. Analisis sentimen pada penelitian ini menggunakan model *Word2vec* sebagai fitur

ekstraksi dan dipadukan dengan metode RNN dan *Naïve Bayes*. *Deep Learning* digunakan untuk analisis sentimen, yaitu *Recurrent Neural Network* (RNN) sedangkan algoritme *Machine Learning* yang digunakan yaitu *Naïve Bayes*. RNN merupakan salah satu bagian dari *Neural Network* untuk memproses data yang tersambung (*sequential data*). Analisis ini menerapkan algoritme RNN yang menunjukkan performa yang baik dalam memproses data seperti teks. Algoritme *Naïve Bayes* adalah algoritme yang mempelajari probabilitas suatu objek dengan ciri-ciri tertentu termasuk dalam kelompok atau kelas tertentu. Metode *Naïve Bayes* yang hanya membutuhkan jumlah data pelatihan (*training data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian (Putro et al., 2021).

Pada penelitian sebelumnya, bahasan studi tentang RNN dan *Naïve Bayes* dan kepuasan pelanggan terhadap pelayanan *e-commerce* telah banyak dilakukan. Tetapi terdapat perbedaan hasil penelitian yang akan dilakukan dari penelitian sebelumnya seperti pada penelitian (Rahayu et al., 2018) menyimpulkan bahwa Shopee ID menggunakan metode *Naïve Bayes Classifier* menghasilkan tingkat akurasi sebesar 87%. Akurasi sistem salah satunya dipengaruhi oleh jumlah data latih yang digunakan. Selain itu, penelitian yang dilakukan oleh (Ivanedra dan Mustikasari, 2019) menyimpulkan bahwa perbedaan banyaknya jumlah *dataset* digunakan sebagai data *training* sangat berpengaruh, karena kata yang lebih banyak tentunya menjadikan program lebih mengerti dan lebih akurat dalam menciptakan ringkasan. Nilai rata-rata *recall* artikel berita non-stemming adalah sebesar 41%, *precision* sebesar 81%, dan *F-Measure* sebesar 54,27%. Sedangkan nilai rata-rata *recall* artikel berita dengan teknik *stemming* sebesar 44%, *precision* sebesar 81%, dan *F-Measure* sebesar 58,20%. Lalu terdapat penelitian oleh (Putro et al., 2021) yang dijadikan data latih, telah berhasil mengklasifikasikan 23 data dari 15 data yang diuji. Sehingga berhasil memprediksi pelanggannya dengan nilai *precision* mencapai 100%, *recall* 91%, dan nilai *accuracy* mencapai 92%.

Berdasarkan pemaparan tersebut, penelitian ini mengklasifikasikan data dengan topik *tweet* mengenai pelayanan *e-commerce* Tokopedia dan Shopee dari media sosial Twitter. Didalam penelitian ini, dilakukan tahapan yang akan dimulai dari tahap *preprocessing* sampai tahap analisis mengukur kualitas akurasi, presisi, *recall*, dan *f1-score* antara dua algoritme, yaitu algoritme RNN dan *Naïve Bayes*.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah analisis sentimen terhadap pelayanan Tokopedia dan Shopee, maka rumusan masalah pada penelitian ini adalah bagaimana hasil klasifikasi data berdasarkan sentimen positif dan negatif, digunakan untuk membandingkan pelayanan antara kedua *e-commerce* yang lebih baik menggunakan algoritme *Long Short-Term Memory*, *Simple Recurrent Neural Network*, dan *Naïve Bayes*.

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini, yaitu:

1. Melakukan penelitian pada *e-commerce* Tokopedia dan Shopee.
2. Data yang digunakan dalam penelitian ini bersumber dari Twitter berbahasa Indonesia.
3. Metode yang diterapkan dalam pengolahan data adalah *Recurrent Neural Network* (RNN) dan *Naïve Bayes*.

1.4 Tujuan Penelitian

Tujuan penelitian yang akan dicapai adalah sebagai berikut:

1. Mengetahui nilai perbandingan kepuasan pelanggan terhadap pelayanan *e-commerce* Tokopedia dan Shopee berdasarkan data Twitter menggunakan metode RNN dan *Naïve Bayes*.

2. Mengetahui apakah nilai pelayanan perusahaan teknologi Indonesia Tokopedia lebih unggul daripada perusahaan teknologi asing Shopee.

1.5 Manfaat Penelitian

Manfaat yang ingin dicapai dalam penelitian ini, yaitu:

1. Memberikan penilaian kepuasan pelayanan pelanggan pada *e-commerce* Tokopedia dan Shopee dari data Twitter berbahasa Indonesia.
2. Mengetahui nilai perbandingan kualitas pelayanan dari *e-commerce* Tokopedia dan Shopee menggunakan metode RNN dan *Naïve Bayes*.

II. TINJAUAN PUSTAKA

2.1 Analisis

Analisis adalah usaha untuk mengamati secara detail sesuatu hal atau benda dengan cara menguraikan komponen pembentuknya atau penyusunnya untuk dikaji lebih lanjut. Jadi, analisis adalah aktivitas yang membuat sejumlah kegiatan seperti menguraikan, membedakan, dan memilih sesuatu untuk digolongkan dan dikelompokkan kembali menurut kriteria tertentu, lalu dicari kaitannya dan ditafsirkan maknanya (Lupi dan Nurdin, 2016).

Beberapa penelitian mengenai analisis terhadap *e-commerce* menggunakan teknik analisis sentimen dan penjelasan mengenai perbedaan antara penelitian tersebut diuraikan pada Tabel 1.

Tabel 1. Penelitian terdahulu terkait analisis sentimen.

Penulis	Judul	Metode	Hasil
Meidita et al., (2016)	Pengaruh kualitas pelayanan terhadap kepuasan, kepercayaan, dan loyalitas pada <i>e-commerce</i> (studi kasus: Shopee)	PLS-SEM dengan dua metode pengujian yaitu <i>Measurement</i> dan <i>Structural</i>	Penelitian ini menunjukkan bahwa kepuasan pelanggan tidak mempengaruhi pelanggan untuk merekomendasikan kepada pengguna lain.
Nugroho dan Sari (2016)	Analisis pengaruh kualitas <i>website</i> Tokopedia terhadap kepuasan pengguna menggunakan metode Webqual 4.0	Webqual 4.0 (<i>Usability, Information, Quality, Service interaction</i>)	Dari hasil kusioner 26 pertanyaan yang dibagikan kepada responden, <i>website</i> Tokopedia memperoleh persentase sebesar 73.46%. Berdasarkan nilai persentase

Penulis	Judul	Metode	Hasil
			tersebut, kualitas <i>website</i> Tokopedia secara keseluruhan termasuk dalam kategori baik menurut pengguna.
Indraloka dan Santosa (2017)	Penerapan <i>text mining</i> untuk melakukan <i>clustering</i> data Twitter Shopee Indonesia	<i>K-means</i>	Hasil yang didapatkan menunjukkan bahwa konsumen dari pelaku bisnis Shopee Indonesia lebih tertarik terhadap program kuis dibandingkan dengan promo atau acara yang memberikan potongan harga.
Putro, Vuldari, dan Saptomo (2021)	Penerapan metode <i>Naïve Bayes</i> untuk klasifikasi pelanggan	<i>Naïve Bayes</i>	Berdasarkan penelitian tentang pelanggan pada data latih yang digunakan, telah berhasil mengklasifikasikan 23 data dari 25 data yang diuji. Sehingga berhasil memprediksi pelanggannya dengan nilai <i>precision</i> mencapai 100%, nilai <i>recall</i> mencapai 91%, nilai <i>accuracy</i> mencapai 92%.

Pada Tabel 1 menjabarkan beberapa penelitian terdahulu dengan metode yang berbeda-beda yang terkait dengan penelitian yang dilakukan yaitu analisis sentimen. Berbeda dengan penelitin terdahulu yang telah dijabarkan pada Tabel 1, penelitian ini akan menggunakan metode *Deep Learning* model *Recurrent Neural Network* (RNN) menggunakan algoritme *Long Short-Term Memory* (LSTM) dan *Simple Recurrent Neural Network* (*Simple RNN*). Dan juga menggunakan metode *Machine Learning* yaitu algoritme *Naïve Bayes* dengan *Word2vec* sebagai fitur ekstraksi kata untuk mencari kedekatan antara satu kata dengan kata lainnya. Data yang

digunakan diperoleh dari media sosial Twitter karena pengguna dapat dengan mudah menyatakan pendapat, ekspresi, opini, informasi bahkan menyampaikan bagaimana perasaan kepuasan pengguna terhadap suatu *e-commerce* sehingga dapat dilakukan penilaian perbandingan kualitas pelayanan. Opini tersebut dapat digunakan untuk menilai baik atau buruk pelayanan pada *e-commerce* Tokopedia dan Shopee karena pengguna cenderung bersikap jujur dalam mengungkapkan perasaan terhadap sesuatu pada media sosial. Analisis yang dilakukan berfokus pada dua *e-commerce* yaitu Tokopedia dan Shopee.

2.2 Analisis Sentimen

Analisis sentimen merupakan cabang penelitian dari *Text Mining* yang berguna untuk mengklasifikasikan dokumen teks berupa opini berdasarkan sentimen. Tugas dasar analisis sentimen adalah mengklasifikasikan beberapa teks dokumen, kalimat atau fitur, kalimat dari fitur tersebut bisa bersifat positif, negatif atau netral (Rusli et al., 2019). Analisis sentimen merupakan suatu aktivitas menganalisa suatu perasaan, emosi, dan opini seseorang yang diekspresikan dalam teks (Mohansyah dan Parani, 2018).

2.3 Kepuasan Pelanggan

Kepuasan merupakan salah satu faktor kunci bagi konsumen dalam melakukan pembelian secara *continue*. Selain itu, pelanggan yang merasa puas akan merekomendasikan produk atau jasa yang digunakannya kepada oranglain, sehingga merupakan salah satu faktor meningkatnya *volume* penjualan perusahaan tersebut (Rorin dan Johan, 2020).

2.4 Twitter

Twitter adalah pelayanan jejaring sosial dan mikrobolog dalam jaringan yang memungkinkan penggunanya untuk mengirim dan membaca pesan

berbasis teks hingga 140 karakter. Pada tanggal 7 November 2017, jumlah karakter bertambah hingga 280 karakter yang dikenal dengan sebutan kicauan (*tweet*). Twitter didirikan pada bulan Maret 2006 oleh Jack Dorsey yang kemudian diluncurkan pada bulan Juli 2006. Sejak diluncurkan, Twitter telah menjadi salah satu dari sepuluh situs yang paling sering dikunjungi di internet dan dijuluki dengan “pesan singkat dari internet”. Salah satu media sosial yang paling sering digunakan oleh para warganet adalah Twitter. Berbeda dengan media sosial lain, Twitter membatasi jumlah penulisannya, yaitu 280 karakter. Hal ini membuat tulisan atau *tweet* yang ditampilkan oleh pengguna Twitter, berisikan pesan yang lebih singkat, padat, dan jelas sehingga para pengguna lebih ringkas untuk memberikan suatu informasi atau opininya (Paputungan dan Jacobus, 2021). Tidak seperti Facebook, Instagram dan media sosial lainnya, Twitter merupakan jejaring sosial yang dapat digunakan sebagai sebuah *graph* berarah yang artinya pengguna dapat mengikuti pengguna lain tetapi pengguna kedua tidak perlu untuk mengikutinya kembali. Biasanya akun yang berstatus publik sehingga dapat diikuti tanpa persetujuan pemilik.

2.5 E-commerce

Fungsi dari internet adalah untuk menyimpan berbagai jenis informasi dengan jumlah besar sebagai media komunikasi dan transaksi sehingga bermanfaat untuk melakukan pembelian kapan saja. Berbagai macam informasi dalam internet memudahkan pengguna untuk mencari pelayanan, seperti penjual dan pembeli tidak harus bertemu tatap muka untuk melakukan transaksi secara langsung, yang dikenal sebagai *Electronic Commerce (E-commerce)* (Meidita et al., 2016).

2.6 Tokopedia

Tokopedia resmi diluncurkan ke publik pada 17 Agustus 2009 di bawah naungan PT Tokopedia, yang didirikan oleh William Tanuwijaya dan

Leotinus Alpha Edison pada 6 Februari 2009 (Malik dan Rohandi, 2017). Tokopedia merupakan salah satu *e-commerce* di Indonesia yang menyediakan pengguna internet dan para wirausaha untuk bertransaksi jual beli secara *online* melalui *smartphone* atau *website*.

2.7 Shopee

Shopee merupakan salah satu *e-commerce* yang sedang berkembang pesat di Indonesia saat ini. Shopee merupakan *e-commerce* yang menyediakan beragam macam produk penjualan. *E-commerce* ini secara resmi masuk ke Indonesia pada tahun 2015. Shopee Indonesia merupakan salah satu peluang bisnis di Indonesia yang bergerak di bidang *e-commerce* dan menggunakan media sosial Twitter sebagai sarana untuk melakukan *advertising* (Indraloka dan Santosa, 2017). Shopee membangun kepercayaan pada pelanggan dengan strategi garansi, gratis biaya ongkos kirim, peminjaman uang atau *Shopeepay later*, dan harga termurah dari *e-commerce* lain.

2.8 Recurrent Neural Network (RNN)

Deep Learning memiliki beberapa metode pembelajaran salah satunya *Recurrent Neural Network* (RNN). RNN adalah salah satu bagian dari keluarga *Neural Network*. RNN dapat digunakan untuk memproses data yang tersambung (*sequential data*) seperti teks. RNN ini merupakan pengembangan dari jaringan syaraf tiruan dan arsitekturnya hampir sama dengan *multi layer perception*, cara kerja dari otak manusia sehingga dapat mengirimkan informasi dari sebuah neuron lain (Suyudi et al., 2018). Jaringan RNN memiliki proses perulangan, setiap *input* yang masuk dan menghasilkan *output*. Lalu akan masuk kembali menjadi *input* untuk diproses di dalam *hidden layer* (Graves et al., 2009).

RNN menggunakan arsitektur yang disebut *simple recurrent neural network*. Arsitektur ini mudah diimplementasikan dan dilatih. *Network* ini memiliki lapisan *input* x , lapisan *context* s , dan lapisan *output* y . *Input network* dalam

timestep t adalah $x(t)$, *output network* dilambangkan dengan $y(t)$, dan $s(t)$ adalah *state network (hidden layer)*. Vektor input $x(t)$ dibentuk oleh vektor w yang mewakili kalimat atau kata yang dilatih, dan *output* dari *neuron* di lapisan *context* s pada saat $t - 1$. Lapisan *input*, *context*, dan *output* dihitung dengan persamaan sebagai berikut (Mikolov et al., 2010).

$$x(t) = w(t) + s(t - 1) \quad (1)$$

$$s_j(t) = f \left(\sum_i x_i(t) u_{ji} \right) \quad (2)$$

$$y_k(t) = g \left(\sum_j s_j(t) u_{kj} \right) \quad (3)$$

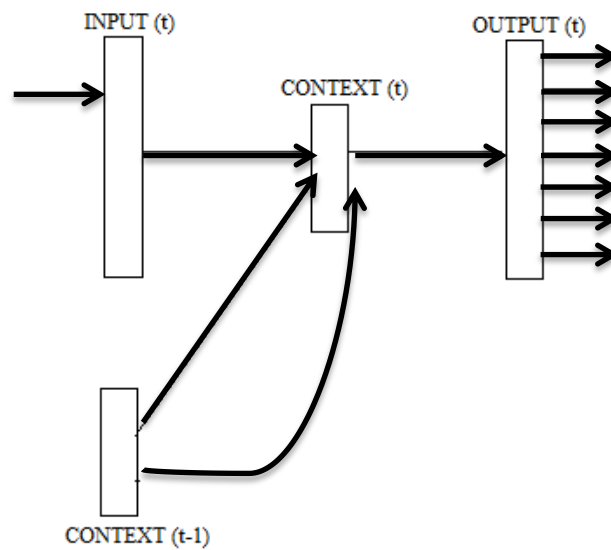
dengan $f(z)$ adalah fungsi aktivasi *sigmoid*:

$$f(z) = \frac{1}{1 + e^{-z}} \quad (4)$$

dan $g(z)$ adalah fungsi *softmax*:

$$g(z_m) = \frac{e^{z_m}}{\sum_k e^{z_k}} \quad (5)$$

Persamaan (1),(2),dan (3) maka cara kerja dari *Simple Recurrent Neural Network* diilustrasikan pada Gambar 2.



Gambar 1. Metode *Simple Recurrent Neural Network* (Mikolov et al., 2010)

Jumlah *context* atau *hidden layer* harus sama dengan jumlah data latih. Untuk data dalam jumlah besar, *hidden layer* dalam jumlah besar sangat diperlukan. *Output layer* $y(t)$ merupakan distribusi probabilitas kata berikutnya dari kata sebelumnya $w(t)$ dan *context* $s(t-1)$. *Softmax* memastikan bahwa distribusi probabilitas ini valid, yaitu $y_m(t) > 0$ untuk kata apapun dari m dan $\sum_k y_k(t) = 1$ (Mikolov et al., 2010).

2.9 Long Short-Term Memory (LSTM)

RNN menunjukkan performa yang baik untuk menyelesaikan tugas *Natural Language Processing* (NLP) mayoritas pengguna RNN menggunakan tipe LSTM, hal itu dikarenakan algoritme LSTM dinilai cukup baik dalam merekam ketergantungan jangka panjang (*long-term dependencies*) jika dibandingkan dengan RNN tipe lain. Pertama kali LSTM diperkenalkan oleh Hochreiter dan Schmidhuber pada tahun. Metode ini dapat memproses data secara sekuensial seperti teks, suara, dan video. RNN memiliki kesulitan dalam mempelajari ketergantungan jarak jauh yang merupakan interaksi antar kata yang terpisah beberapa bagian. Peran LSTM dalam hal

ini adalah menentukan bagian penting dari teks yang akan disimpan dalam sel (Cahyadi et al., 2020).

2.10 *Naïve Bayes*

Thomas Bayes menemukan *Naïve Bayes* di abad ke-18. Algoritme ini didasari oleh teorema bayes yang menggunakan perhitungan berupa probabilitas dan statistic untuk menyelesaikan berbagai masalah dalam *Machine Learning* (Thenata, 2021). *Naïve Bayes* memiliki kelebihan, yaitu cepat dalam perhitngan, algoritme yang sederhana dan berakurasi tinggi dan hanya membutuhkan jumlah data pelatihan yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian (Putro et al., 2021). Berikut ini rumus yang digunakan dalam perhitungan *Naïve Bayes* (Kautsar dan L, 2018) :

$$P(C|X) = \frac{P(X|C)P(C)}{P(C)} \quad (6)$$

Keterangan:

X = peubah acak

C = kelas

$P(C/X)$ = probabilitas C didasarkan pada kondisi X

$P(X/C)$ = probabilitas X di bawah kondisi C

$P(C)$ = probabilitas C

$P(X)$ = probabilitas X

2.11 *Confusion Matrix*

Confusion matrix adalah tabel yang menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah (Normawati dan Prayogi, 2021). Evaluasi menggunakan *confusion matrix* menghasilkan nilai akurasi, presisi, *recal*, dan *f1-score*. Akurasi dalam klasifikasi merupakan presentasi ketepatan *record* data diklasifikasikan secara benar setelah dilakukan pengujian pada hasil klasifikasi. Presisi merupakan proposikasi yang diprediksi positif yang juga positif benar pada data sebenarnya. *Recall*

merupakan proporsi kasus positif yang sebenarnya diprediksi positif secara benar.

Tabel *confusion matrix* digunakan untuk mengukur kinerja suatu metode klasifikasi dengan menghitung nilai akurasi, presisi, *recall*, dan *f1-score*. Perhitungan kinerja sistem klasifikasi ini dihitung menggunakan data dalam tabel *confusion matrix* pada Tabel 2.

Tabel 2. *Confusion matrix*

Aktual	Prediksi	
	Positive	Negative
Positive	TP (True Positive)	FP (False Positive)
Negative	FN (False Negative)	TN (True Negative)

True Positive (TP) merupakan jumlah *record* positif dalam dataset yang diklasifikasikan positif. *True Negative* (TN) merupakan jumlah *record* negatif dalam dataset yang diklasifikasikan positif. *False Negative* (FN) merupakan jumlah *record* positif dalam dataset yang diklasifikasikan negatif. Berikut adalah persamaan model *confusion matrix* (Putra dan Wibowo, 2020).

- a. Akurasi (*accuracy*) presentase dari jumlah total prediksi yang benar.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (7)$$

- b. Presisi (*precision*) adalah porsi kasus positif yang teridentifikasi dengan benar.

$$precision = \frac{TP}{TP + FP} \times 100\% \quad (8)$$

- c. *Recall* merupakan presentase kasus positif yang teridentifikasi benar.

$$recall = \frac{TP}{TP + FN} \times 100\% \quad (9)$$

- d. *F1-Score* merupakan perbandingan rata-rata presisi dan *recall* yang dibobotkan.

$$f1\ score = 2 \times \frac{(recall \times precision)}{(recall + precision)} \times 100\% \quad (10)$$

2.12 *Word Embedding*

Word embedding merupakan vektor bernilai *real* yang mempresentasikan suatu kata yang dapat mewakili konteks di mana kata tersebut muncul. Pada vektor ini kata akan digambarkan dalam bentuk numerik yang setiap katanya dapat dipetakan ke dalam ruang ruang vektor. Salah satu tipe *word embedding* adalah *word2vec* (Wijayanto et al., 2021). *Word embedding* digunakan dalam penelitian ini karena *word embedding* dapat memperluas *query* yang digunakan dengan mencari hubungan kata terdekat sehingga harapannya dapat memperluas *query* yang dapat menambah performa dari *Word2vec* dan *word embedding* menghasilkan vektor yang bernilai *real* mempresentasikan suatu kata yang mewakili konteks pada kemunculannya, kata ini dipresentasikan ke bentuk numerik yang memungkinkan untuk memetakan setiap kata ke suatu titik ruang vektor (Sari et al., 2021)

2.13 *Word2vec*

Word2vec merupakan algoritme *word embedding* yaitu pemetaan dari kata menjadi *vector*. *Vector* ini digunakan untuk berbagi tugas *Natural Language Processing* (NLP). Pada *Word2vec* dalam pembobotan kata digunakan nilai rata-rata *vector* yang mewakili kata tersebut (Rusli et al., 2019). Metode *Word2vec* bertujuan untuk mencari kedekatan antara satu kata dengan kata lainnya. Metode ini memiliki dua model arsitektur yaitu *Continuous Bag-of-words* (CBOW) dan *Skip-gram* (Niasita et al., 2019). Model CBOW bertujuan untuk memprediksi target kata dari konteks kata, sedangkan model Skip-Gram bertujuan untuk memprediksi probabilitas kata yang dapat menjadi konteks kata dari target kata yang ada (Mikolov et al., 2013). Keunggulan dari *Word2vec* adalah memiliki ekstensi untuk menyimpan kata-kata dari seluruh kalimat yang disebut dengan *paragraf2vec* atau *doc2vec*. *Word2vec* juga telah diimplementasikan dalam beberapa bahasa pemrograman yaitu C, Python, dan Java yang dapat mendukung inferensi penyimpanan dokumen pada dokumen tak terlihat (Le dan Mikolov, 2014).

2.14 Bahasa Pemrograman Python

Python merupakan bahasa pemrograman yang populer saat ini. Python digunakan oleh perusahaan besar dan juga para pengembang perangkat lunak untuk membuat aplikasi berbasis desktop, berbasis web, dan aplikasi *mobile* (Romzi dan Kurniawan, 2020). Python merupakan bahasa pemrograman yang *freeware* atau perangkat bebas dalam arti sebenarnya. Lengkap dengan *source code*, *debugger*, dan *profiler*, antarmuka yang terkandung di dalamnya untuk pelayanan antarmuka, fungsi sistem, GUI (antarmuka pengguna grafis), dan basis datanya (Clinton dan Sengkey, 2019)

2.15 Deep Learning

Deep Learning adalah cabang ilmu *Machine Learning* berbasis jaringan syaraf tiruan atau perkembangan dari syaraf tiruan jaringan. Dalam *Deep Learning*, sebuah komputer belajar mengklasifikasi secara langsung dari gambar atau suara dan teks (Ilahiyah dan Nilogiri, 2018). *Deep Learning* mempelajari beberapa tingkat representasi dan abstraksi yang membantu memahami data seperti gambar, suara, dan teks. *Deep Learning* mencakup algoritma *unsupervised* dan *supervised learning* sehingga dapat memanfaatkan data yang sudah berlabel maupun tidak berlabel (Zulfa dan Winarko, 2017).

2.16 Machine Learning

Machine Learning merupakan salah satu cabang dari disiplin ilmu kecerdasan buatan yang membahas mengenai pembangunan sistem yang berdasarkan pada data. Banyak hal yang dipelajari, akan tetapi pada dasar ada empat hal pokok yang perlu dipelajari dalam *Machine Learning* yaitu, *supervised learning*, *unsupervised learning*, *semi-supervised learning*, dan *reinforcement learning* (Zailani et al., 2020).

III. METODE PENELITIAN

3.1 Tempat dan Waktu Penelitian

Penelitian ini dilakukan di Jurusan Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung. Penelitian mulai dilaksanakan pada semester ganjil tahun ajaran 2020/2021.

3.2 Alat dan Bahan

Penelitian ini dilakukan dengan menggunakan alat untuk mendukung dan menunjang pelaksanaan penelitian, yaitu sebagai berikut:

1. Perangkat keras (*Hardware*)

Perangkat keras yang digunakan dalam penelitian analisis ini adalah satu unit laptop dengan spesifikasi Prosesor: Intel(R) Celeron(R) CPU N3060 @ 1.60 GHz (CPUs), ~1.6 GHz.

2. Perangkat lunak (*Software*)

Perangkat lunak yang digunakan dalam analisis ini sebagai berikut:

- *Operating System*: Windows 10
- *Web Browser* Google Chrome
- *IDE* untuk melakukan pemrograman Python yaitu Jupyter.

Jupyter digunakan dalam penelitian ini untuk menghasilkan analisis sentimen. Jupyter didukung oleh berbagai fitur yang mudah dipahami untuk perhitungan data, visualisasi data teks menjadi diagram, dan dapat menjelaskan makna dari data.

- *IDE* untuk melakukan pemograman Python yaitu Spyder.
Spyder digunakan untuk menampilkan sebuah tabel seperti tabel pada normalisasi teks, tokenisasi teks, penghilangan *stopword*, dan *word2vec*.
- *Library* pada Python yang digunakan pada penelitian.
Library pada Python merupakan gabungan dari sekumpulan *package* dan *module* dengan fungsi yang sama dengan tujuan untuk memudahkan membuat suatu sistem. *Library* pada Python merupakan sebutan untuk kode program tambahan yang digunakan dalam kebutuhan kode program, dengan menggunakan *library* pada Python dapat menghasilkan kode secara efisien. Pada penelitian ini ada beberapa *library* yang digunakan antara lain:
 - a. Pandas 1.3.2
Pandas merupakan *library* yang digunakan untuk analisis data karena memiliki struktur data yang diperlukan untuk membersihkan data mentah ke dalam sebuah bentuk yang sesuai untuk tahap analisis.
 - b. Matplotlib 3.4.2
Merupakan *library* yang digunakan untuk membuat visualisasi data yang akan ditampilkan dalam bentuk diagram.
 - c. Skicit-learn 0.24.2
Skicit-learn atau sklearn merupakan sebuah *module* dari bahasa pemograman Python. Fungsi dari *module* ini untuk membantu melakukan *processing* data ataupun melakukan *training* data untuk kebutuhan *data science*.
 - d. Keras 2.6.0
Library yang digunakan untuk menyederhanakan model *Deep Learning*. Keras merupakan *library* untuk pembelajaran mesin

terbaik. Keras juga menyediakan beberapa utilitas terbaik dalam menyusun model, memproses *dataset*, dan memvisualisasikan grafik

e. NLTK 3.6.2

Natural Language Tool Kit (NLTK) adalah sebuah *library* yang digunakan untuk memudahkan memproses teks seperti penghapusan *stopword*, *tokenization*, dan normalisasi teks dalam analisis data penelitian ini.

f. Gensim 4.1.0

Gensim adalah *library* untuk *Word2vec* dan *document tagging*.

g. Numpy 1.19.5

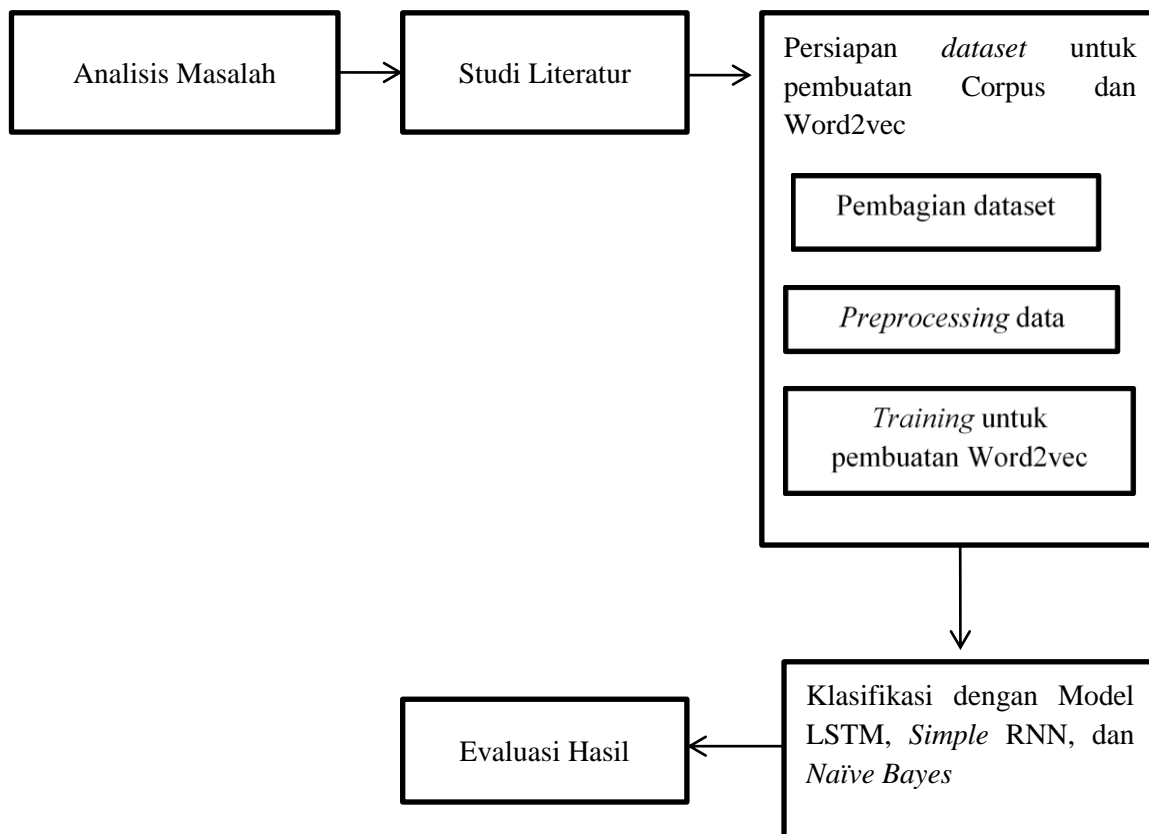
Numpy merupakan *library* standar untuk perhitungan angka.

h. Sastrawi 1.0.1

Sastrawi merupakan *library* Python untuk melakukan *stemming* dalam bahasa Indonesia.

3.3 Tahapan Penelitian

Dalam penelitian ini melalui beberapa tahap, sebagaimana yang ditunjukkan pada Gambar 2.



Gambar 2. Alur tahapan penelitian analisis pelayanan Tokopedia dan Shopee

Pada Gambar 2 merupakan tahap penelitian yang dilakukan. Tahapan penelitian ini mencakup analisis masalah yang terjadi, studi literatur yang berkaitan dengan penelitian, persiapan *dataset* untuk *processing* data dan *Word2vec*, klasifikasi dengan metode yang digunakan, yaitu algoritme LSTM, Simple RNN, dan *Naïve Bayes*, dan tahap terakhir evaluasi hasil dari nilai akurasi, presisi, *recall*, dan *f1-score*.

Berikut ini merupakan penjabaran dari tahapan penelitian pada Gambar 2.

1. Analisis masalah

Analisis permasalahan yang ada yaitu persaingan antara *e-commerce* Tokopedia dan Shopee yang terdapat pada ulasan/*tweets* pengguna di sosial media Twitter. Membandingkan kepuasan pelayanan yang

dirasakan pelanggan terhadap kedua *e-commerce* menggunakan *Recurrent Neural Network* (RNN) dan *Naïve Bayes* untuk mendapatkan nilai akurasi yang terbaik.

2. Studi literatur

Tahap ini dilakukan pengumpulan informasi pada penelitian sebelumnya yang berhubungan dengan topik permasalahan yang sama dan mempelajari beberapa metode penelitian.

3. Persiapan *dataset*

Tahap persiapan *dataset* yang berupa data ulasan kepuasan pelayanan *e-commerce* Tokopedia dan Shopee pada media sosial Twitter dengan mengumpulkan *tweet* mengenai bagaimana pelayanan kedua *e-commerce* tersebut. Lalu pemabagian *dataset* untuk data latih dan data uji. Selanjutnya tahap *Preprocessing data*, beberapa tahapan *preprocessing* yaitu normalisasi teks, tokenisasi teks, dan penghilangan *stopword*. Bahasa pemrograman yang akan digunakan yaitu Python. Klasifikasi sentimen dari perbandingan kepuasan pelayanan antara *e-commerce* Tokopedia dan Shopee yang diperoleh dari Twitter yang telah melewati tahap *preprocessing* dengan menggunakan metode *Recurrent Neural Network* (RNN) dan *Naïve Bayes* yang bertujuan untuk menerapkan analisis data pada Twitter untuk isu *marketing*.

Fungsi dari tahapan *preprocessing* adalah sebagai berikut:

a. Normalisasi Teks

Normalisasi teks berfungsi untuk melakukan penormalan kata atau kalimat dalam suatu data, normalisasi akan dilakukan pada seluruh data teks penelitian yang digunakan. Proses normalisasi teks akan menghilangkan sekumpulan simbol-simbol seperti seperti tanda titik, tanda koma, tanda tanya, tanda seru dan simbol lainnya atau *URL* (Muludi et al., 2021).

b. Tokenisasi Teks

Tahap ini merupakan tahap pemotongan *string* atau kalimat input berdasarkan tiap kata yang menyusun kalimat dalam teks. Tokenisasi secara garis besar memecah sekumpulan karakter dalam suatu teks ke dalam satuan kata, bekerja dengan membedakan karakter-karakter tertentu yang dapat diperlakukan sebagai pemisah kata atau bukan sebuah kata (Muludi et al., 2021).

c. *Stopword*

Stopword adalah sekumpulan kata-kata yang sering muncul tetapi jika dihapus tidak mengubah makna dari kalimat didalam *dataset* yang diperoleh dari tweet tersebut. Pembuangan *stopword* dilakukan untuk mengetahui suatu kata yang tidak memiliki arti. Kata yang telah diproses melalui tahap tokenisasi diperiksa dalam *stopword*, apabila ada sebuah kata masuk di dalam *stopword* maka kata itu tidak akan diproses lebih lanjut (Muludi et al., 2021).

4. Evaluasi hasil

Pada tahap evaluasi hasil data dilakukan perbandingan nilai sentimen dari pelabelan manual dan setelah itu dilakukan perbandingan hasil akurasi dari klasifikasi data Tokopedia dan Shopee menggunakan algoritme RNN dan algoritme *Naïve Bayes*. Dalam proses pengujian klasifikasi dari hasil perhitungan *confusion matrix* yang menjadi acuan untuk prediksi dan aktual, yang terdiri dari akurasi, presisi, *recall*, dan *f1-score*.

V. SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan pembahasan dan hasil analisis, bahwa ulasan *tweet* yang berjumlah 1002 data Tokopedia dan 1000 data Shopee. Data yang dibagi menjadi dua kelas yaitu kelas negatif yang memberi opini keluhan dan kecewa dari *user* Twitter atas pelayanan kedua *e-commerce* dan kelas positif menunjukkan opini kepuasan dan kesenangan atas pelayanan kedua *e-commerce*. Dalam proses pengujian klasifikasi dari hasil perhitungan *confusion matrix* yang menjadi acuan untuk prediksi dan aktual, lalu menggunakan tiga model algoritme yaitu *Simple RNN*, *LSTM*, dan *Naïve Bayes*. Pada ulasan pelanggan lebih banyak menyukai pelayanan Tokopedia dari *user* Twitter dapat dilihat dengan data uji yang positif sebesar 120 data sedangkan data uji yang positif pada pelayanan Shopee sebesar 97 data. Hasil akurasi yang memperoleh performa lebih baik yaitu Tokopedia menggunakan algoritme *LSTM* dengan nilai akurasi mencapai 78%, presisi mencapai 78%, *recall* mencapai 78%, dan *f1-score* mencapai 78%. Jika dibandingkan nilai klasifikasi terbaik dari data Shopee dari algoritme *Simple RNN* menunjukkan nilai klasifikasi dengan hasil akurasi sebesar 58%, presisi mencapai 58%, *recall* sebesar 58%, dan *f1-score* mencapai 58%.

5.2 Saran

Saran yang dapat diberikan untuk penelitian selanjutnya adalah mengembangkan analisis sentimen ini dengan menambah jumlah data ulasan Tokopedia dan Shopee dan menggunakan metode lain seperti *CNN* dan *SVM*, sehingga dapat diketahui hasil nilai perbandingan yang lebih baik

lagi dan bagi pihak *e-commerce*, hasil klasifikasi dari ulasan yang telah diberikan pengguna Twitter khususnya ulasan pada pelayanan yang membuat kecewa dapat dijadikan bahan evaluasi dalam peningkatan kepuasan pelanggan dan memberikan pelayanan yang lebih baik.

DAFTAR PUSTAKA

- Badieah, Gernowo, R. dan Surarso, B. (2016) “Metode Jaringan Syaraf Tiruan Untuk Prediksi Performa Mahasiswa Pada Pembelajaran Berbasis Problem Based Learning (PBL),” *Jurnal Sistem Informasi Bisnis*, 6(1), hal. 46. doi: 10.21456/vol6iss1pp46-58.
- Cahyadi, R., Damayanti, A. dan Aryadani, D. (2020) “Recurrent Neural Network (RNN dengan Long Short Term Memory (LSTM) untuk Analisis Sentimen Data Instagram,” *Jurnal Informatika dan Komputer*, 5(1), hal. 1–9.
- Clinton, R. M. R. dan Sengkey, S. (2019) “Purwarupa Sistem Daftar Pelanggaran Lalulintas Berbasis Mini-Komputer Raspberry Pi,” *Jurnal Teknik Elektro dan Komputer*, 8(3), hal. 181–192.
- Graves, A., Liwicki, M. dan Fernandez, S. (2009) “A Novel Connectionist System for Unconstrained Handwriting Recognition,” *IEEE Trans. Pattern Anal. Mach. Intell*, 31(5), hal. 855–868. doi: 10.1109/TPAMI.2008.137.
- Ilahiyah, S. dan Nilogiri, A. (2018) “Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network,” hal. 49–56.
- Indraloka, D. S. dan Santosa, B. (2017) “Penerapan Text Mining untuk Melakukan Clustering Data Tweet Shopee Indonesia,” *Jurnal Sains dan Seni ITS*, 6(2), hal. 6–11. doi: 10.12962/j23373520.v6i2.24419.
- Ivanedra, K. dan Mustikasari, M. (2019) “Implementasi Metode Recurrent Neural Network Pada Text Summarization Dengan Teknik Abstraktif the Implementation of Text Summarization With Abstractive Techniques Using Recurrent Neural Network Method,” 6(4), hal. 377–382. doi: 10.25126/jtiik.201961067.
- Kautsar, R. dan L, K. M. (2018) “Analisis Sentimen terhadap Toko Online menggunakan Naïve Bayes pada Media Sosial Twitter,” 5(3), hal. 8141–8151.
- Le, Q. dan Mikolov, T. (2014) “Distributed Representations of Sentences and Documents Quoc,” *Proceedings of the 24th International Conference on World Wide Web - WWW '15 Companion*, 32, hal. 29–30. doi: 10.1145/2740908.2742760.

- Lupi, F. R. dan Nurdin (2016) “Analisis strategi pemasaran dan penjualan,” *Jurnal Elektronik Sistim Informasi Dan Komputer (Jesik)*, 2(1), hal. 20–29.
- Malik, M. dan Rohandi, A. (2017) “Effectiveness C2C E-Commerce Media In Bandung (Case study at Tokopedia.com and Bukalapak.com),” *Jurnal Ekonomi dan Bisnis-Manajemen*, 1(3 pp), hal. 177–197.
- Meidita, Y., Suprpto dan Rokhmawati, R. I. (2016) “Pengaruh Kualitas Layanan Terhadap Kepuasan, Kepercayaan dan Loyalitas Pelanggan Pada E-Commerce (Studi Kasus : Berrybenka),” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(11), hal. 5682–5690.
- Mikolov, T., Karafiat, M. dan Burget, L. (2010) “Recurrent neural network based language model,” *Proceedings of the 11th Annual Conference of the International Speech Communication Association, INTERSPEECH 2010*, (26-30 September 2010), hal. 1045–1048.
- Mikolov, T., Karafiat, M. dan Burget, L. (2013) “Efficient Estimation of Word Representations in Vector Space,” *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings*, (16 Januari 2013), hal. 1–12.
- Muludi, K., Akbar, M., Shofiana, D. dan Syarif, A. (2021) “Sentiment Analysis Of Energy Independence Tweets Using Simple Recurrent Neural Network,” *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 15(4), hal. 339. doi: 10.22146/ijccs.66016.
- Niasita, A. F., Adikara, P. P. dan Adinugroho, S. (2019) “Analisis Sentimen Pembangunan Infrastruktur di Indonesia dengan Automated Lexicon Word2Vec dan Naive-Bayes,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 3(3), hal. 2673–2679.
- Normawati, D. dan Prayogi, S. A. (2021) “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” *Jurnal Sains Komputer dan Informatika (J-SAKTI)*, 5(2), hal. 697–711.
- Nugroho, A. K. dan Sari, P. K. (2016) “Analisis Pengaruh Kualitas Website Tokopedia Terhadap Kepuasan Pengguna Menggunakan Metode Webqual 4.0,” *e-Proceeding of Management*, 3(Desember 2016), hal. 2930–2937.
- Paputungan, C. K. N. dan Jacobus, A. (2021) “Sentiment Analysis of Social Media Users Using Long-Short Term Memory Method,” *Jurnal Teknik Elektro dan Komputer*, 10(2), hal. 99–106.
- Putra, D. dan Wibowo, A. (2020) “Prediksi Keputusan Minat Penjurusan Siswa SMA Yadika 5 Menggunakan Algoritma Naïve Bayes,” *Prosiding Seminar Nasional Riset Dan Information Science (SENARIS)*, 2(30 Juli 2020), hal. 84–92. doi: DOI:10.30645/SENARIS.V2I0.147.
- Putro, H. F., Vulandari, R. T. dan Saptomo, W. L. (2021) “Penerapan Metode

- Naive Bayes Untuk Klasifikasi Pelanggan,” *Jurnal TIKomSIN*, 8(2), hal. 1–7. doi: 10.30646/tikomsin.v8i2.500.
- Rahayu, S. dan Sismoro, H. (2018) “Sentimen Analisis Review Pengguna Marketplace Online Menggunakan Naïve Bayes Classifier,” *Jurnal Informasi Interaktif*, 3(3), hal. 154–209.
- Roihan, A., Sunarya, P. A. dan Rafika, A. S. (2020) “Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper,” *IJCIT (Indonesian Journal on Computer and Information Technology)*, 5(1), hal. 75–82. doi: 10.31294/ijcit.v5i1.7951.
- Romzi, M. dan Kurniawan, B. (2020) “Pembelajaran Pemrograman Python Dengan Pendekatan Logika Algoritma,” *JTIM: Jurnal Teknik Informatika Mahakarya*, 03(2), hal. 37–44.
- Rorin, D., Insana, M. dan Johan, R. S. (2020) “Peningkatan Kepuasan Konsumen Melalui Penggunaan E-Commerce,” *sosio e-kons*, 12(2), hal. 125–137. doi: 10.30998/sosioekons.v12i02.6451.
- Rusli, M., Faisal, M. dan Budiman, I. (2019) “Ekstraksi Fitur Menggunakan Model Word2Vec Untuk Analisis Sentimen Pada Kolom Komentar Kuisisioner Evaluasi Dosen Oleh Mahasiswa,” *Kumpulan Jurnal Komputer (KLIK)*, 2(1), hal. 104–109. doi: 10.20527/klik.v7i1.296.
- Suyudi, M. A. D., Djamal, E. c dan Maspupah, A. (2019) “Prediksi Harga Saham menggunakan Recurrent Neural Network,” *Seminar Nasional Aplikasi Teknologi Informasi (sNATi)*, (Yogyakarta, 03 Agustus 2019), hal. A33-37. doi: 10.1007/978-1-4842-3564-5_10.
- Syah, A. P., Adiwijaya dan Faraby, S. Al (2017) “Analisis Sentimen Pada Data Ulasan Produk Toko Online Dengan Metode Maximum Entropy,” *e-Proceeding of Engineering*, 4(3)(Desember 2017), hal. 4632–4640.
- Thenata, A. P. (2021) “Text Mining Literature Review on Indonesian Social Media,” *Jurnal Edukasi dan Penelitian Informatika*, 7(2), hal. 226–232.
- Wijayanto, I. R., Cholissodin, I. dan Sari, Y. A. (2021) “Pengaruh Metode Word Embedding dalam Vector Space Model pada Pemerolehan Informasi Materi IPA Siswa SMP,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 5(3), hal. 950–959.
- Zailani, A. U., Perdananto, A. dan Nurjaya. (2020) “Pengenalan Sejak Dini Siswa Smp Tentang Machine Learning Untuk Klasifikasi Gambar Dalam Menghadapi Revolusi 4.0,” *KOMMAS: Jurnal Pengabdian Kepada Masyarakat*, 1(1), hal. 7–15.
- Zulfa, I. dan Winarko, E. (2017) “Sentimen Analisis Tweet Berbahasa Indonesia Dengan Deep Belief Network,” *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 11(2), hal. 187. doi: 10.22146/ijccs.24716.