



***SENTIMENT ANALYSIS CUSTOMERS REVIEW PADA PLATFORM  
E-COMMERCE TOKOPEDIA***

**KERJA PRAKTIK**



**Program Studi  
S1 Teknik Komputer**

**Oleh :**

**SALGA AMAR RUSYDA**

**21410200009**

---

**FAKULTAS TEKNOLOGI DAN INFORMATIKA**

**UNIVERSITAS DINAMIKA**

**2024**

***SENTIMENT ANALYSIS CUSTOMERS REVIEW PADA PLATFORM***

***E - COMMERCE TOKOPEDIA***

Diajukan Sebagai salah satu syarat untuk menyelesaikan

Program Sarjana



**Disusun Oleh :**

**Nama** : Salga Amar Rusyda

**NIM** : 21410200009

**Program** : S1 (Strata Satu)

**Jurusan** : Teknik Komputer

**FAKULTAS TEKNOLOGI DAN INFORMATIKA**

**UNIVERSITAS DINAMIKA**

**2024**



UNIVERSITAS  
Dinamika

Semua bisa karena terbiasa  
Salga Amar Rusyda

## LEMBAR PENGESAHAN

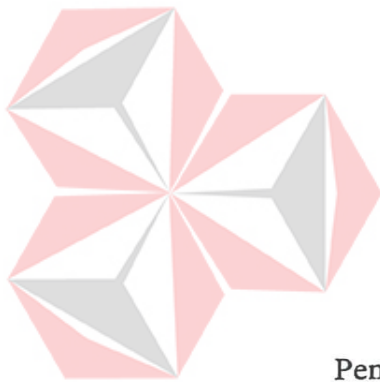
### SENTIMENT ANALYSIS CUSTOMERS REVIEW PADA PLATFORM E - COMMERCE TOKOPEDIA

Laporan Kerja Praktik oleh

**Salga Amar Rusyda**

NIM: 21410200009

Telah diperiksa, diuji, dan disetujui



UNIVERSITAS

Surabaya, 31 Juli 2024

Dinamika

Disetujui:

Pembimbing

Digitally signed by  
Heri Pratikno, M.T.  
Date: 2024.07.31  
14:50:47 +07'00'

Heri Pratikno, M.T.  
NIDN. 0716117302

Penyelia

Lutfi Dwimulya

Mengetahui,

Ketua Program Studi S1 Teknik Komputer

Pauladie Susanto, S.Kom., M.T.  
Ketua Program Studi S1 Teknik Komputer  
2024.07.31 14:34:00 +07'00'

Pauladie Susanto, S.Kom., M.T.  
NIDN. 0729047501

**PERNYATAAN**  
**PERSETUJUAN PUBLIKASI DAN KEASLIAN KARYA ILMIAH**

Sebagai mahasiswa **Universitas Dinamika**, Saya :

Nama : **Salga Amar Rusyda**  
NIM : **2141020009**  
Program Studi : **Teknik Komputer**  
Fakultas : **Teknologi dan Informatika**  
Jenis Karya : **Laporan Kerja Praktik**  
Judul Karya : **SENTIMENT ANALYSIS CUSTOMERS REVIEW  
PADA PLATFORM E-COMMERCE TOKOPEDIA**

Menyatakan dengan sesungguhnya bahwa :

1. Demi pengembangan Ilmu Pengetahuan, Teknologi dan Seni, Saya menyetujui memberikan kepada **Universitas Dinamika** Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty Free Right*) atas seluruh isi/sebagian karya ilmiah Saya tersebut diatas untuk disimpan, dialihmediakan, dan dikelola dalam bentuk pangkalan data (*database*) untuk selanjutnya didistribusikan atau dipublikasikan demi kepentingan akademis dengan tetap mencantumkan nama Saya sebagai penulis atau pencipta dan sebagai pemilik Hak Cipta.
2. Karya tersebut diatas adalah hasil karya asli Saya, bukan plagiat baik sebagian maupun keseluruhan. Kutipan, karya, atau pendapat orang lain yang ada dalam karya ilmiah ini semata-mata hanya sebagai rujukan yang dicantumkan dalam Daftar Pustaka Saya.
3. Apabila dikemudian hari ditemukan dan terbukti terdapat tindakan plagiasi pada karya ilmiah ini, maka Saya bersedia untuk menerima pencabutan terhadap gelar keserjanaan yang telah diberikan kepada Saya.

Demikian surat pernyataan ini Saya buat dengan sebenar-benarnya.

Surabaya, 29 Juli 2024



Salga Amar Rusyda  
NIM : 2141020009



UNIVERSITAS  
Dinamika

*Laporan Kerja Praktik ini  
Saya persembahkan untuk  
Keluarga tersayang, Dosen Pembimbing, serta  
Teman-teman yang saya cintai*

## ABSTRAK

Proyek ini membahas penerapan analisis sentimen terhadap ulasan pelanggan pada *platform e-commerce* Tokopedia menggunakan metode *Natural Language Processing (NLP)*. Dalam era digital saat ini, ulasan pelanggan menjadi sumber informasi yang sangat berharga bagi perusahaan untuk memahami persepsi dan kepuasan pelanggan. Analisis sentimen digunakan untuk mengklasifikasikan ulasan-ulasan tersebut ke dalam kategori sentimen positif dan negatif. Proyek ini menggunakan berbagai teknik *NLP* seperti *tokenisasi*, *stemming*, *lemmatization*, dan penghapusan *stop words* untuk memproses teks ulasan. Selain itu, model machine learning seperti *Long Short-Term Memory (LSTM)* diterapkan untuk membangun model klasifikasi *sentiment*.

Data yang digunakan dalam proyek ini dikumpulkan dari dataset ulasan pelanggan Tokopedia yang tersedia di platform *Kaggle*. Hasil analisis menunjukkan bahwa metode *NLP* yang digunakan mampu mengklasifikasikan sentimen dengan akurasi yang cukup tinggi 89 %, dengan model *LSTM* menunjukkan kinerja yang lebih baik dibandingkan dengan model tradisional. Model yang dibangun memberikan wawasan mendalam tentang persepsi pelanggan terhadap produk dan layanan yang ditawarkan di Tokopedia. Temuan ini diharapkan dapat membantu perusahaan dalam meningkatkan kualitas layanan dan strategi pemasaran mereka.

**Kata Kunci :** Analisis Sentimen, *Natural Language Processing*, *Machine Learning*, *LSTM*

## KATA PENGANTAR

Saya mengucapkan syukur kepada Tuhan Yang Maha Esa atas segala rahmat-Nya, yang memungkinkan saya menyelesaikan laporan kerja praktik ini dengan judul “*SENTIMENT ANALYSIS CUSTOMERS REVIEW PADA PLATFORM E-COMMERCE TOKOPEDIA*” Laporan ini merupakan salah satu syarat untuk menyelesaikan studi sarjana saya. Dalam kesempatan ini, saya ingin menyampaikan penghargaan dan terima kasih yang tulus kepada :

1. Ayah dan Ibu, atas doa dan dukungan yang tiada henti
2. Bapak Heri Pratikno, M.T. selaku dosen pembimbing, atas bimbingan dan arahan yang sangat berarti.
3. Kak M. Riziq Maulana, mentor dari Hactiv8, atas panduan teknis dan wawasan yang mendalam.
4. Sahabat dan teman – teman atas dukungan dan semangat dalam proses penyelesaian kerja praktik.

Saya menyadari bahwa laporan ini mungkin masih memiliki kekurangan. Oleh karena itu, saya sangat terbuka terhadap kritik dan saran yang membangun untuk perbaikan di masa mendatang. Semoga laporan ini bermanfaat bagi pembaca dan memberikan kontribusi positif.

Surabaya, 21 Juli 2024

Penulis,



Salga Amar Rusyda



## DAFTAR ISI

ABSTRAK .....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI .....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR .....	xi
DAFTAR LAMPIRAN .....	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan.....	3
1.5 Manfaat.....	4
BAB II GAMBARAN UMUM.....	5
2.1 Latar Belakang Perusahaan .....	5
2.2 Identitas Perusahaan .....	6
2.3 Visi Perusahaan .....	6
2.4 Misi Perusahaan.....	6
2.5 Struktur Organisasi.....	7
2.6 Peta Lokasi .....	8
2.7 Produk dan Layanan.....	8
BAB III LANDASAN TEORI.....	11
3.1 Analisis Sentimen.....	11
3.2 Machine Learning.....	11
3.3 <i>Python (Data Science)</i> .....	12
3.4 <i>Long Short-Term Memory (LSTM)</i> .....	12
3.5 <i>Natural Language Processing (NLP)</i> .....	13
3.6 <i>IBM Watson Space</i> .....	14
BAB IV DESKRIPSI PEKERJAAN .....	15

4.1.	Latar Belakang Proyek .....	15
4.2	Tujuan Proyek.....	15
4.3	Struktur Tim Proyek .....	16
4.4	Metode Penelitian.....	16
	4.4.1 <i>Data Processing</i> .....	17
	4.4.2 <i>Exploratory Data Analysis (EDA)</i> .....	20
	4.4.3 <i>Training model</i> .....	27
	4.4.4 <i>Evaluation Model</i> .....	28
	4.4.5 <i>Deployment</i> .....	30
BAB V PENUTUP .....		31
5.1	Kesimpulan.....	31
5.2	Saran .....	31
	Daftar Pustaka .....	33
	LAMPIRAN.....	34



UNIVERSITAS  
**Dinamika**

## DAFTAR TABEL

Tabel 4. 1 Struktur Tim Proyek.....	16
-------------------------------------	----



UNIVERSITAS  
**Dinamika**

## DAFTAR GAMBAR

Gambar 2.1 Logo Hactiv8.....	5
Gambar 2.2 Struktur Organisasi PT. Hacktivate Teknologi Indonesia .....	7
Gambar 2.3 Lokasi pusat PT. Hacktivate Teknologi Indonesia .....	8
Gambar 4.1 Program dari <i>Data Loading</i> .....	17
Gambar 4.2 Program dari <i>Data Cleaning</i> .....	18
Gambar 4.3 Program dari <i>Text Normalization</i> .....	19
Gambar 4.4 Program dari <i>Stopword</i> .....	20
Gambar 4.5 Program dari <i>Word Frequency Analysis</i> .....	20
Gambar 4.6 Hasil dari program <i>Word Frequency Analysis</i> .....	21
Gambar 4.7 Program <i>Wordcloud</i> positif.....	22
Gambar 4.8 Hasil program <i>Wordcloud</i> positif .....	23
Gambar 4.9 Program <i>Wordcloud</i> negatif.....	23
Gambar 4.10 Hasil program <i>Wordcloud</i> negatif .....	24
Gambar 4.11 Program dari <i>Sentiment Distribution</i> .....	25
Gambar 4.12 Hasil dari program <i>Sentiment Distribution</i> .....	25
Gambar 4.13 Program dari <i>Emotion Distribution</i> .....	26
Gambar 4.14 Hasil dari program <i>Emotion Distribution</i> .....	26
Gambar 4.15 Training model menggunakan <i>early stopping</i> .....	27
Gambar 4.16 Training model tanpa menggunakan <i>early stopping</i> .....	28
Gambar 4.17 Program dari <i>evaluation model</i> .....	29
Gambar 4.18 Hasil dari program <i>evaluation model</i> .....	29

## DAFTAR LAMPIRAN

Lampiran 1. Surat Balasan dari Mitra .....	34
Lampiran 2. <i>Logbook</i> Bulanan .....	35
Lampiran 3. Laporan Aktivitas .....	41
Lampiran 4. Sertifikat Kepesertaan MSIB Batch 6 .....	42
Lampiran 5. Sertifikat Program MSIB Batch 6 .....	42
Lampiran 6. Transkrip Nilai .....	43
Lampiran 7. Kartu Bimbingan Kerja Praktik .....	44
Lampiran 8. Biodata Penulis .....	45



UNIVERSITAS  
**Dinamika**

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan pesat teknologi informasi dan komunikasi telah mengubah cara orang berinteraksi, termasuk dalam hal berbelanja. *E-commerce*, sebagai salah satu bentuk perdagangan elektronik, telah menjadi pilihan utama bagi banyak konsumen untuk memenuhi kebutuhan mereka. *Platform e-commerce* seperti Tokopedia menyediakan kemudahan bagi pelanggan untuk membeli berbagai produk secara online. Namun, dengan semakin banyaknya ulasan yang ditinggalkan oleh pelanggan, perusahaan harus menghadapi tantangan dalam menganalisis dan memahami informasi yang disampaikan dalam ulasan tersebut.

Ulasan pelanggan di *platform e-commerce* merupakan sumber data yang sangat berharga untuk mengevaluasi produk dan layanan. Informasi yang terdapat dalam ulasan dapat memberikan wawasan mengenai kepuasan pelanggan, kualitas produk, dan area yang perlu diperbaiki. Namun, volume ulasan yang besar membuat analisis manual menjadi tidak praktis. Oleh karena itu, diperlukan metode otomatis untuk menganalisis dan menginterpretasikan data tersebut secara efisien.

*Sentiment analysis* atau analisis sentimen adalah teknik yang digunakan untuk menentukan emosi atau sentimen di balik teks. Dalam konteks *e-commerce*, teknik ini dapat digunakan untuk mengategorikan ulasan pelanggan menjadi sentimen positif dan negatif. Dengan menggunakan *Natural Language Processing* (NLP) dan model pembelajaran mesin seperti *Long Short-Term Memory* (LSTM), *analysis sentiment*

dapat dilakukan secara otomatis untuk memberikan gambaran yang lebih jelas mengenai persepsi pelanggan terhadap produk yang ditawarkan.

Proyek ini bertujuan untuk menerapkan *analysis sentiment* pada ulasan pelanggan di Tokopedia untuk mengidentifikasi pola-pola sentimen dan memberikan wawasan yang dapat membantu perusahaan dalam meningkatkan layanan serta strategi pemasaran mereka. Dengan pendekatan ini, diharapkan perusahaan dapat lebih efektif dalam merespons umpan balik pelanggan dan mengoptimalkan pengalaman berbelanja di *platform* mereka.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam Kerja Praktik ini adalah bagaimana menerapkan *analysis sentiment* menggunakan teknik *Natural Language Processing* (NLP) untuk mengklasifikasikan ulasan pelanggan pada *platform e-commerce* Tokopedia ke dalam kategori sentimen positif dan negatif. Permasalahan utama pada Kerja Praktik adalah bagaimana performa penerapan *model Long Short-Term Memory* (LSTM) dalam mengidentifikasi sentimen dan ulasan pelanggan Tokopedia melalui parameter uji akurasi, presisi, dan *recall*. Selain itu, penelitian ini juga akan mengevaluasi efektivitas teknik *NLP* seperti *tokenisasi*, *stemming*, *lemmatization*, dan penghapusan *stop words* dalam memproses data teks untuk meningkatkan hasil klasifikasi.

## 1.3 Batasan Masalah

Berdasarkan uraian di atas, maka dalam pelaksanaan Kerja Praktik terdapat beberapa batasan masalah, antara lain:

1. Proyek ini hanya menggunakan dataset ulasan pelanggan yang diambil dari *Kaggle*, yang mencakup ulasan dalam bahasa Indonesia.
2. Fokus utama dari proyek ini adalah penerapan model *Long Short-Term Memory* (LSTM) untuk analisis sentimen, dan teknik *Natural Language Processing* (NLP) seperti *tokenisasi*, *stemming*, *lemmatization*, dan penghapusan *stop words*.
3. *Deployment* sistem akan dilakukan di *IBM Watson Studio* dengan API terpisah, dan tidak akan mencakup *platform deployment* atau metode lain.
4. Proyek ini tidak akan mempertimbangkan aspek lain dari ulasan pelanggan seperti analisis topik atau fitur-fitur tambahan.

#### 1.4 Tujuan

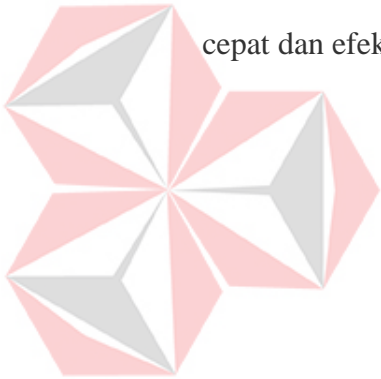
Berdasarkan uraian dari latar belakang dan rumusan masalah, maka tujuan dari Kerja Praktik ini yaitu untuk menerapkan *sentiment analysis* pada ulasan pelanggan di *platform e-commerce* Tokopedia dengan menggunakan teknik *Natural Language Processing* (NLP) dan model *Long Short-Term Memory* (LSTM) secara otomatis untuk mengidentifikasi pola-pola sentimen dalam ulasan memberikan wawasan mengenai kepuasan dan persepsi pelanggan terhadap produk, serta membantu perusahaan dalam meningkatkan layanan dan strategi pemasaran mereka agar lebih responsif dan efektif.



## 1.5 Manfaat

Adapun manfaat dari pelaksanaan Kerja Praktik ini antara lain sebagai berikut:

1. Membantu perusahaan untuk lebih memahami bagaimana pelanggan merasakan produk dan layanan mereka dengan mengidentifikasi pola sentimen dalam ulasan.
2. Mempercepat dan mempermudah analisis ulasan menggunakan teknik otomatis seperti *Natural Language Processing* (NLP) dan model *Long Short-Term Memory* (LSTM), dibandingkan dengan metode manual.
3. Memungkinkan perusahaan untuk merespons umpan balik pelanggan secara lebih cepat dan efektif, yang dapat memperbaiki hubungan dan citra merek.



UNIVERSITAS  
**Dinamika**

## **BAB II**

### **GAMBARAN UMUM**

#### **2.1 Latar Belakang Perusahaan**

Perusahaan yang memberikan tugas untuk mengerjakan proyek penelitian ini selama program MSIB Batch 6 adalah Hactiv8, yang merupakan sebuah coding bootcamp yang berbasis di Jakarta, Indonesia. Mereka berfokus pada pelatihan intensif untuk mempersiapkan individu dalam karir di bidang teknologi, terutama dalam pengembangan web, data science, dan Marketing. Hactiv8 menawarkan berbagai program yang dirancang untuk membantu siswa dari berbagai latar belakang untuk menjadi pengembang perangkat lunak yang kompeten dalam waktu singkat. Hactiv8 memiliki Kantor yang berlokasi di Jl. Sultan Iskandar Muda No.7, RT.5/RW.9, Kby. Lama Sel., Kec. Kby. Lama, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta 12240. Logo dari Hactiv8 sendiri dapat dilihat pada Gambar 2.1



Gambar 2.1 Logo Hactiv8

## 2.2 Identitas Perusahaan

Nama Instansi : PT. Hacktivate Teknologi Indonesia. (Hacktiv8)  
Alamat : Jl. Sultan Iskandar Muda No.7, RT.5/RW.9, Kby. Lama Sel.,  
Kec. Kby. Lama, Kota Jakarta Selatan, Daerah Khusus  
Ibukota Jakarta 12240  
No. Telepon : 02180675787  
Website : <https://www.hacktiv8.com/>  
Email : -

## 2.3 Visi Perusahaan

Hactiv8 memiliki visi untuk menjadi perusahaan terkemuka di Asia Tenggara yang menghasilkan talenta teknologi berkualitas tinggi dan siap kerja yang mampu bersaing di pasar global.

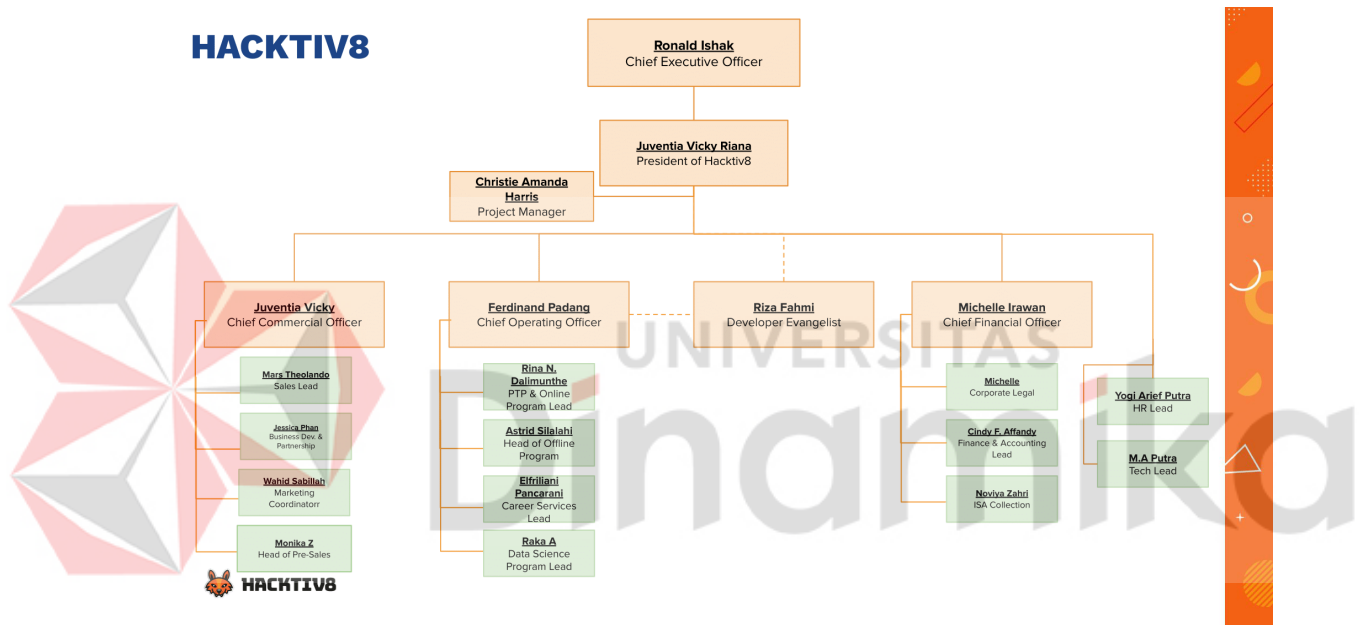
## 2.4 Misi Perusahaan

Misi dari Hacktiv8 adalah sebagai berikut :

1. Menyediakan program pelatihan yang berfokus pada keterampilan praktis dan relevan dengan perusahaan.
2. Membantu siswa mendapatkan pekerjaan di perusahaan teknologi terkemuka melalui jaringan dan kemitraan yang luas.
3. Mengembangkan metode pengajaran yang inovatif dan efektif untuk memastikan keberhasilan siswa dalam karir teknologi mereka.

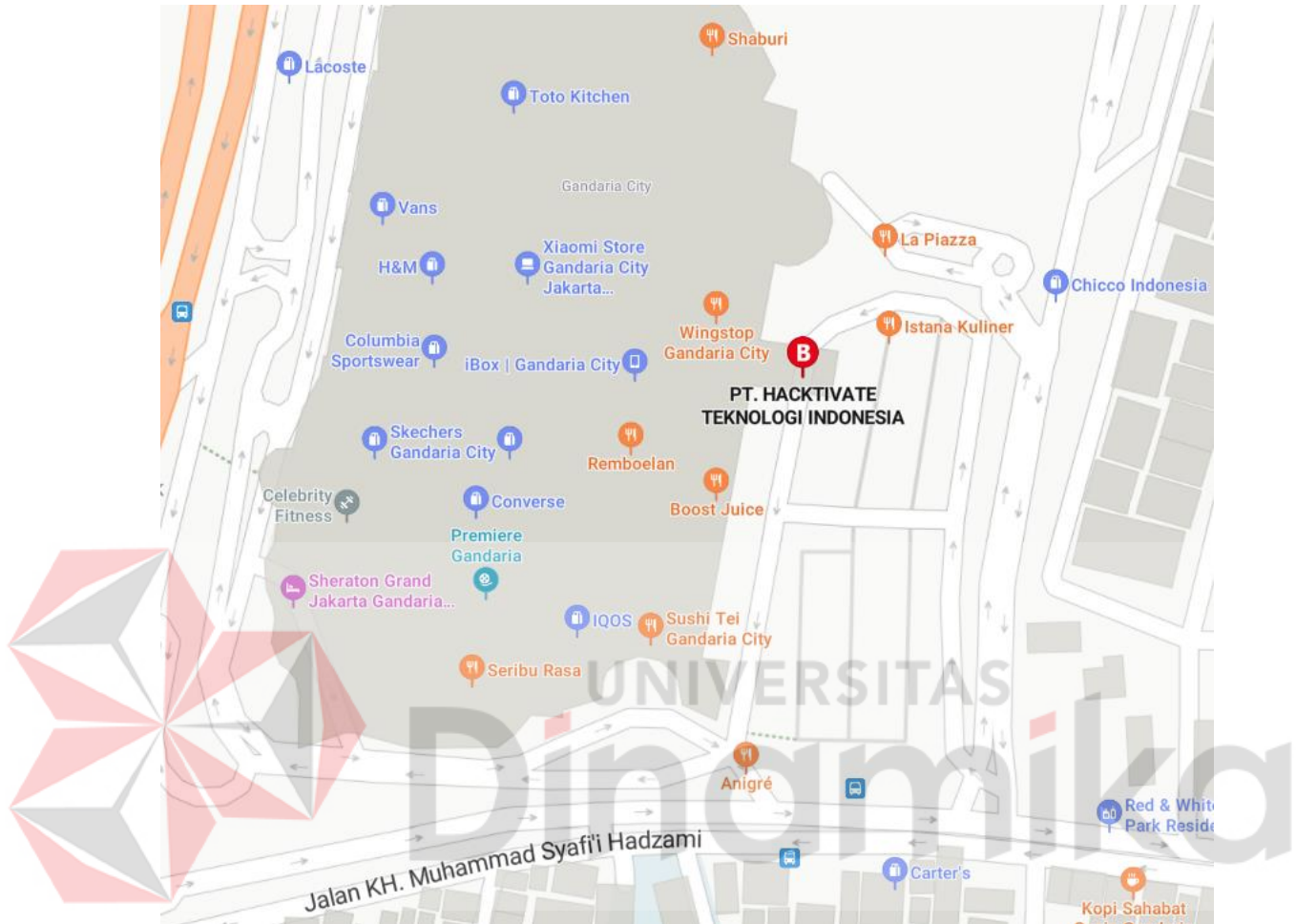
4. Membangun dan mendukung komunitas pengembang yang berkelanjutan dan berkontribusi pada ekosistem teknologi di Indonesia.
5. Memastikan bahwa setiap lulusan memiliki keterampilan dan pengetahuan yang diperlukan untuk sukses dalam karir mereka.

## 2.5 Struktur Organisasi



Gambar 2.2 Struktur Organisasi PT. Hacktivate Teknologi Indonesia

## 2.6 Peta Lokasi



Gambar 2.3 Lokasi pusat PT. Hacktivate Teknologi Indonesia

## 2.7 Produk dan Layanan

PT Hacktivate Teknologi Indonesia menyediakan berbagai produk dan layanan yang bertujuan untuk mendukung *reskilling* dan *upskilling* di industri teknologi. Berikut adalah beberapa produk dan layanan unggulan yang ditawarkan:

## 1. Bootcamp Coding

Hacktiv8 menawarkan program bootcamp intensif yang dirancang untuk mengubah pemula menjadi pengembang perangkat lunak siap kerja dalam waktu singkat. Program ini mencakup berbagai bahasa pemrograman dan teknologi terbaru seperti JavaScript, React, dan Node.js. Dengan kurikulum berbasis proyek yang praktis, siswa mendapatkan pengalaman nyata yang diperlukan untuk meraih sukses dalam karir mereka.

## 2. Program Spesialisasi

Selain *bootcamp coding*, Hacktiv8 juga menyediakan program spesialisasi untuk keterampilan teknologi tertentu. Program ini mencakup bidang seperti *data science*, *machine learning*, dan pengembangan aplikasi *mobile*. Setiap program dirancang untuk memberikan pengetahuan mendalam dan keterampilan praktis yang dapat langsung diterapkan di industri.

## 3. Kelas Online

Untuk menjangkau lebih banyak individu yang ingin meningkatkan keterampilan mereka, Hacktiv8 menyediakan kelas online yang fleksibel. Kelas ini memungkinkan peserta belajar dari mana saja dan kapan saja, dengan materi yang mencakup berbagai topik teknologi dan pengembangan keterampilan digital.

## 4. Pelatihan Perusahaan

Hacktiv8 bekerja sama dengan berbagai perusahaan untuk menyediakan pelatihan karyawan yang disesuaikan dengan kebutuhan spesifik perusahaan tersebut. Pelatihan ini membantu perusahaan dalam meningkatkan keterampilan digital

karyawan mereka dan mempersiapkan mereka untuk tantangan teknologi yang terus berkembang.

#### 5. **Mentorship dan Dukungan Karir**

Selain pelatihan teknis, Hactiv8 menawarkan dukungan karir yang komprehensif, termasuk *mentorship*, *workshop* pengembangan karir, dan bantuan penempatan kerja. Tim *mentorship* Hactiv8 terdiri dari profesional berpengalaman yang siap membantu siswa mencapai tujuan karir mereka.



UNIVERSITAS  
**Dinamika**

## BAB III

### LANDASAN TEORI

#### 3.1 Analisis Sentimen

Analisis Sentimen adalah teknik dalam *Natural Language Processing* (NLP) yang digunakan untuk mengidentifikasi dan mengklasifikasikan perasaan atau opini yang terkandung dalam teks. Teknik ini sangat penting dalam konteks analisis ulasan pelanggan pada *platform e-commerce* seperti Tokopedia, di mana pemahaman terhadap sentimen pelanggan dapat memberikan wawasan berharga mengenai kepuasan dan persepsi mereka terhadap produk atau layanan.

#### 3.2 Machine Learning

Machine Learning adalah cabang dari kecerdasan buatan (*artificial intelligence* atau AI) yang berfokus pada pengembangan algoritma dan teknik yang memungkinkan komputer untuk belajar dari data dan membuat prediksi atau keputusan berdasarkan data tersebut. Dalam konteks analisis sentimen, *machine learning* digunakan untuk mengembangkan model yang dapat mengklasifikasikan teks ulasan pelanggan ke dalam kategori sentimen positif dan negatif. Teknik ini memungkinkan sistem untuk meningkatkan kinerjanya dalam menyelesaikan tugas spesifik tanpa diprogram secara eksplisit. Secara umum, *machine learning* dapat dibagi menjadi beberapa kategori utama, yaitu *supervised learning*, *unsupervised learning*, dan *reinforcement learning*. Dalam *supervised learning*, model dilatih menggunakan data yang telah diberi label. *Unsupervised learning* melibatkan model yang dilatih



menggunakan data tanpa label untuk menemukan pola tersembunyi. Sementara itu, *reinforcement learning* melibatkan model yang belajar melalui trial and error dengan menerima umpan balik dalam bentuk hadiah atau hukuman.

### **3.3 Python (Data Science)**

*Python* adalah bahasa pemrograman tingkat tinggi yang sangat populer dalam pengembangan aplikasi AI, termasuk *machine learning* dan pemrosesan bahasa alami (*natural language processing* atau NLP). *Python* dikenal karena sintaksnya yang sederhana dan kemampuannya untuk mendukung berbagai perpustakaan dan kerangka kerja seperti Scikit-Learn, TensorFlow, dan Keras yang sangat penting dalam pengembangan model *machine learning*. *Python* menawarkan banyak keunggulan, termasuk kemudahan penggunaan, fleksibilitas, dan komunitas pengguna yang besar. Dokumentasi yang luas dan dukungan dari komunitas pengguna membuat *Python* menjadi pilihan utama bagi banyak pengembang dan peneliti dalam bidang *machine learning* dan NLP. Selain itu, *Python* mendukung integrasi dengan berbagai alat dan platform, memungkinkan pengembang untuk membangun, melatih, dan menerapkan model AI dengan lebih efisien.

### **3.4 Long Short-Term Memory (LSTM)**

*Long Short-Term Memory* (LSTM) adalah jenis jaringan saraf tiruan yang termasuk dalam kategori *recurrent neural networks* (RNN). LSTM dirancang untuk menangani masalah gradien yang menghilang dan mampu mengingat informasi penting dalam urutan data untuk jangka waktu yang panjang. Arsitektur LSTM terdiri dari memori sel yang mampu menyimpan informasi untuk waktu yang lama serta

gerbang input, output, dan lupa yang mengontrol aliran informasi masuk, keluar, dan dilupakan oleh sel memori. LSTM sangat efektif untuk tugas-tugas NLP seperti analisis sentimen karena kemampuannya untuk menangani data sekuensial. Dalam analisis sentimen, LSTM dapat digunakan untuk mempelajari pola sentimen dalam teks ulasan pelanggan dan membuat prediksi yang akurat mengenai sentimen tersebut. Keunggulan LSTM dalam mengelola dependensi jangka panjang membuatnya sangat cocok untuk analisis data teks yang memiliki urutan dan konteks yang kompleks.

### **3.5 Natural Language Processing (NLP)**

*Natural Language Processing (NLP)* adalah cabang dari kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia. NLP memungkinkan komputer untuk membaca, memahami, dan menghasilkan bahasa manusia dengan cara yang bermanfaat. Teknik-teknik utama dalam NLP meliputi *tokenization*, yaitu proses memecah teks menjadi unit-unit kecil seperti kata atau kalimat; *stemming* dan *lemmatization*, yang mengurangi kata ke bentuk dasarnya; serta *stop words removal*, yang menghapus kata-kata umum yang tidak memberikan informasi penting.

Dalam analisis sentimen, NLP digunakan untuk mengategorikan ulasan pelanggan ke dalam sentimen positif dan negatif dengan memproses dan menganalisis teks ulasan. NLP juga melibatkan teknik lain seperti pengenalan entitas bernama (*named entity recognition*), analisis topik (*topic modeling*), dan penguraian sintaksis (*syntactic parsing*), yang semuanya membantu dalam pemahaman dan interpretasi teks bahasa alami. Dengan menggunakan teknik NLP, sistem dapat mengidentifikasi emosi

dan opini yang diungkapkan dalam ulasan pelanggan, memberikan wawasan yang berharga bagi perusahaan untuk meningkatkan kualitas produk dan layanan mereka.

### **3.6 IBM Watson Space**

*IBM Watson Space* adalah platform kecerdasan buatan yang menyediakan berbagai alat dan layanan untuk pengembangan aplikasi AI. *Watson Space* menyediakan lingkungan untuk mengembangkan, melatih, dan menerapkan model AI dengan fitur utama seperti API kecerdasan buatan yang memungkinkan integrasi mudah dengan aplikasi untuk pemrosesan bahasa alami, analisis sentimen, dan pembelajaran mesin. Lingkungan pengembangan yang didukung oleh *Watson Space* memungkinkan perusahaan untuk menerapkan model yang telah dikembangkan dan mendapatkan wawasan dari ulasan pelanggan secara otomatis, mendukung proses *deployment* model analisis sentimen dalam proyek ini.

*IBM Watson Space* menawarkan berbagai layanan seperti pemrosesan bahasa alami, analitik prediktif, dan analisis data, yang semuanya dapat diakses melalui API yang mudah digunakan. Dengan menggunakan *Watson*, pengembang dapat dengan cepat mengintegrasikan kemampuan AI ke dalam aplikasi mereka, mengurangi waktu dan usaha yang diperlukan untuk mengembangkan solusi AI dari awal. Selain itu, *Watson Space* mendukung kolaborasi tim, memungkinkan beberapa pengembang untuk bekerja bersama pada proyek yang sama, meningkatkan efisiensi dan produktivitas.

## **BAB IV**

### **DESKRIPSI PEKERJAAN**

#### **4.1. Latar Belakang Proyek**

Dengan kemajuan teknologi internet, aktivitas berbelanja telah beralih ke platform digital secara signifikan. Platform belanja online menjadi pilihan utama bagi banyak orang karena kemudahan akses, harga yang lebih terjangkau, dan beragam promo yang ditawarkan. Namun, meski populer, kegiatan belanja online juga menghadapi beberapa masalah, seperti ketidaksesuaian produk dengan deskripsi, pengiriman lambat, pelayanan buruk, dan risiko penipuan. Meningkatkan pengalaman pengguna dalam belanja online menjadi penting untuk mempertahankan minat mereka.

Model yang dibangun bertujuan untuk mengklasifikasikan emosi atau sentimen dalam ulasan produk. Masalah yang dipecahkan adalah memahami dan menganalisis ulasan produk untuk mengetahui emosi atau sentimen yang terkandung di dalamnya. Hal ini dapat memberikan wawasan berharga kepada pemilik bisnis atau perusahaan *e-commerce* tentang bagaimana pelanggan merespons produk mereka, dan dapat membantu dalam meningkatkan strategi pemasaran, pengembangan produk, dan layanan pelanggan.

#### **4.2 Tujuan Proyek**

Tujuan dari analisis sentimen ulasan pelanggan adalah untuk memahami dan menganalisis berbagai ulasan produk yang diberikan oleh pelanggan. Dengan melakukan analisis ini, kita dapat mengetahui dan mengidentifikasi emosi atau

sentimen yang terkandung di dalam setiap ulasan, baik itu positif maupun negatif. Selain itu, analisis ini juga membantu dalam mengungkapkan pandangan dan perasaan pelanggan terhadap produk tersebut, yang dapat digunakan untuk meningkatkan kualitas produk dan pelayanan di masa depan.

### 4.3 Struktur Tim Proyek

Proyek dengan tema *setiment analysis* ini dikerjakan oleh tim yang tertera pada Tabel 4.1

Judul Proyek : *Sentimeny Analysis Customers Review* pada *Platform E-commerce* Tokopedia

Kelas : SI6-12

Mentor : M Riziq Maulana

Tabel 4. 1 Struktur Tim Proyek

No	Nama	Perguruan Tinggi	Program Studi
1	Siti Fatimah	Universitas Pamulang	Teknik Informatika
2	Sadat Ardiansyah	Universitas Esa Unggul	Teknik Informatika
3	Sangga Dewa Ramadhan	Politeknik Negeri Jember	Teknik Informatika
4	Sonny Fahmi Wicaksono	Universitas Muhammadiyah Jakarta	Teknik Informatika
5	Ronald Wira Utama	Universitas Gunadarma	Teknik Informatika
6	Salga Amar Rusyda	Universitas Dinamika	Teknik Komputer

### 4.4 Metode Penelitian

Dataset ini adalah kumpulan data ulasan produk dari Tokopedia yang diberi label emosi dan sentimen. Dataset ini mencakup ulasan dalam bahasa Indonesia dari 29 kategori produk. Setiap ulasan diberi satu dari lima label emosi: cinta, kebahagiaan, kemarahan, ketakutan, atau sedih. Proses anotasi dilakukan oleh kelompok annotator berdasarkan kriteria dari ahli psikologi klinis. Selain itu, dataset ini juga mencakup atribut lain seperti lokasi, harga, rating keseluruhan, jumlah terjual, total ulasan, dan

rating pelanggan untuk mendukung penelitian lebih lanjut. Proses Pembuatan proyek ini terdiri dari beberapa tahapan, yaitu :

#### 4.4.1 Data Processing

*Data Processing* adalah tahap awal yang penting untuk menyiapkan data teks agar bisa digunakan oleh model pembelajaran mesin. Tujuan utama dari *data processing* adalah membersihkan, memformat, dan mengubah data teks menjadi format yang bisa dipahami dan digunakan oleh algoritma *Natural Language Processing* (NLP). Adapun beberapa langkah dalam *Data Processing* yaitu:

##### 1. Data Loading

Langkah program pada Gambar 4.1 adalah langkah dimana kita mengambil data dari file *Kaggle* dan memasukkannya ke dalam program agar bisa diolah dan dianalisis lebih lanjut. Data yang digunakan di sini adalah dataset ulasan produk dalam bahasa Indonesia.



```
[ ] df = pd.read_csv('Product Reviews Dataset for Emotions Classification Tasks - Indonesian (PRDECT-ID) Dataset/PRDECT-ID Dataset.csv')
df.head()
```

	Category	Product Name	Location	Price	Overall Rating	Number Sold	Total Review	Customer Rating	Customer Review	Sentiment	Emotion
0	Computers and Laptops	Wireless Keyboard i8 Mini TouchPad Mouse 2.4G ...	Jakarta Utara	53500	4.9	5449	2369	5	Alhamdulillah berfungsi dengan baik. Packaging...	Positive	Happy
1	Computers and Laptops	PAKET LISENSI WINDOWS 10 PRO DAN OFFICE 2019 O...	Kota Tangerang Selatan	72000	4.9	2359	1044	5	barang bagus dan respon cepat, harga bersaing ...	Positive	Happy
2	Computers and Laptops	SSD Midasforce 128 Gb - Tanpa Caddy	Jakarta Barat	213000	5.0	12300	3573	5	barang bagus, berfungsi dengan baik, seller ram...	Positive	Happy
3	Computers and Laptops	ADAPTOR CHARGER MONITOR LCD LED TV LG merek LG...	Jakarta Timur	55000	4.7	2030	672	5	bagus sesuai harapan penjual nya juga ramah. t...	Positive	Happy
4	Computers and Laptops	ADAPTOR CHARGER MONITOR LCD LED TV LG merek LG...	Jakarta Timur	55000	4.7	2030	672	5	Barang Bagus, pengemasan Aman, dapat Berfungsi...	Positive	Happy

Gambar 4.1 Program dari *Data Loading*

## 2. Data Cleaning

Langkah program pada Gambar 4.2 adalah digunakan untuk memeriksa dan menampilkan informasi dasar tentang dataset, menghapus baris yang memiliki nilai yang hilang, menghapus duplikat, dan menghapus

```
# Checking the status of the dataset in rows, columns, duplicate entries & missing values
print(f'\033[94mJumlah catatan (baris) dalam dataset adalah: {df.shape[0]}')
print(f'\033[94mJumlah fitur (kolom) dalam dataset adalah: {df.shape[1]}')
print(f'\033[94mJumlah entri duplikat dalam dataset adalah: {df.duplicated().sum()}')
print(f'\033[94mJumlah nilai yang hilang dalam dataset adalah: {sum(df.isna().sum())}')

Jumlah catatan (baris) dalam dataset adalah: 5400
Jumlah fitur (kolom) dalam dataset adalah: 11
Jumlah entri duplikat dalam dataset adalah: 7
Jumlah nilai yang hilang dalam dataset adalah: 0

[ ] df.dropna(subset=['Customer Review'], inplace=True)

[ ] # Removing duplicates
df.drop_duplicates(subset=['Customer Review'], keep='first', inplace=True)

[ ] # Removing irrelevant columns
df.drop(['Location', 'Price', 'Number Sold'], axis=1, inplace=True)
```

Gambar 4.2 Program dari *Data Cleaning*

## 3. Text Normalization

*Text normalization* adalah proses mengubah teks menjadi format yang konsisten dan lebih mudah diolah oleh komputer. Lalu program pada Gambar 4.3 bertujuan untuk membersihkan teks ulasan pelanggan dengan menghapus karakter-karakter yang tidak relevan seperti URL, angka, tanda baca, dan mengubah semua teks menjadi huruf kecil.

```
[ ] # import library natural Language toolkit
import nltk
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
nltk.download('punkt')
nltk.download('stopwords')

# membuat stopwords list
# mengembalikan daftar kata-kata umum dalam bahasa Indonesia yang biasanya diabaikan dalam analisis teks
stopwords_list = stopwords.words('indonesian')

[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Unzipping tokenizers/punkt.zip.
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Unzipping corpora/stopwords.zip.

[ ] def remove_special_characters(text):
    text = text.lower() # Mengubah teks menjadi lower case
    text = re.sub(r'https?://\S+|www\.\S+', '', text) # Menghapus URL
    text = re.sub(r'[-+]?[0-9]+', '', text) # Menghapus angka
    text = re.sub(r'^\w\s|$', '', text) # Menghapus karakter tanda baca
    text = re.sub(r' ', ' ', text) # Menghapus spasi ganda
    text = text.strip()
    return text

df['Customer Review'] = df['Customer Review'].apply(remove_special_characters)
```

Gambar 4.3 Program dari *Text Normalization*

#### 4. Stopword

*Stopword* adalah kata-kata umum yang sering muncul dalam teks tetapi biasanya tidak memiliki makna penting dalam analisis teks. Pada program Gambar 4.4 digunakan untuk menghapus kata – kata yang tidak penting sehingga teks menjadi lebih fokus pada kata-kata yang membawa makna penting dan menghapus kata-kata umum sehingga teks menjadi lebih bersih dan lebih mudah untuk dianalisis lebih lanjut.



```

# Fungsi untuk menghapus stopwords dari teks
def remove_stopwords(text):
    tokens = text.split() # Pisahkan teks menjadi token
    filtered_tokens = [word for word in tokens if word.lower() not in stopwords_list] # Hapus stopwords
    cleaned_text = ' '.join(filtered_tokens) # Gabungkan token yang sudah difilter kembali menjadi teks
    return cleaned_text.strip() # Hapus spasi di depan dan di belakang teks

# Gunakan fungsi remove_stopwords pada kolom 'Customer Review'
df['Customer Review'] = df['Customer Review'].apply(remove_stopwords)

```

Gambar 4.4 Program dari *Stopword*

#### 4.4.2 Exploratory Data Analysis (EDA)

EDA adalah proses awal dalam analisis data yang membantu kita memahami data dengan cara yang sederhana dan ditampilkan dengan visual gambar. Adapun beberapa langkah dalam proses *Exploratory Data Analysis* (EDA) yaitu:

##### 1. *Word Frequency Analysis*

*Word Frequency Analysis* adalah bagian dari *Exploratory Data Analysis* (EDA) yang berfokus pada menganalisis kata-kata yang paling sering muncul dalam data teks.

```

[ ] # Make a copy of cleaned dataframe for visualization purpose
df_vis = df.copy()

from collections import Counter
df_vis['temp_list'] = df_vis['Customer Review'].apply(lambda x:str(x).split())
top = Counter([item for sublist in df_vis['temp_list'] for item in sublist])
temp = pd.DataFrame(top.most_common(20))
temp.columns = ['Common_words', 'count']
temp.style.background_gradient(cmap='Blues')

```

Gambar 4.5 Program dari *Word Frequency Analysis*



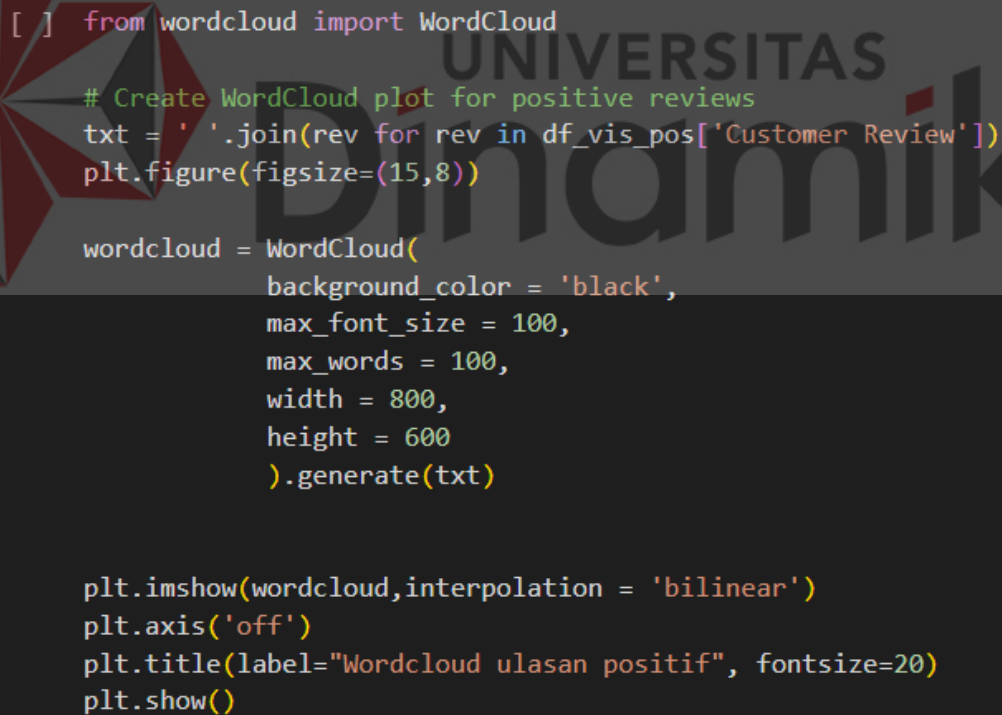
	Common_words	count
0	barang	1473
1	sesuai	1084
2	bagus	873
3	cepat	774
4	pengiriman	658
5	beli	589
6	seller	540
7	produk	517
8	kecewa	414
9	dikirim	402
10	aman	384
11	harga	369
12	respon	364
13	pesanan	362
14	packing	361
15	kualitas	343
16	rapi	308
17	terima	305
18	barangnya	294
19	toko	283

Gambar 4.6 Hasil dari program *Word Frequency Analysis*

## 2. Wordcloud

*Wordcloud* adalah cara visual untuk menampilkan kata-kata yang paling sering muncul dalam teks. Pada program Gambar 4.7 adalah program untuk *wordcloud* kata – kata positif yang paling sering muncul dan pada program Gambar 4.9 adalah program untuk *wordcloud* kata – kata negatif yang paling sering muncul.

Untuk hasil dari program Gambar 4.7 dan Gambar 4.9 dapat dilihat pada Gambar 4.8 untuk *wordcloud* positif dan Gambar 4.10 untuk *wordcloud* negatif.



```
[ ] from wordcloud import WordCloud
# Create WordCloud plot for positive reviews
txt = ''.join(rev for rev in df_vis_pos['Customer Review'])
plt.figure(figsize=(15,8))

wordcloud = WordCloud(
    background_color = 'black',
    max_font_size = 100,
    max_words = 100,
    width = 800,
    height = 600
).generate(txt)

plt.imshow(wordcloud,interpolation = 'bilinear')
plt.axis('off')
plt.title(label="Wordcloud ulasan positif", fontsize=20)
plt.show()
```

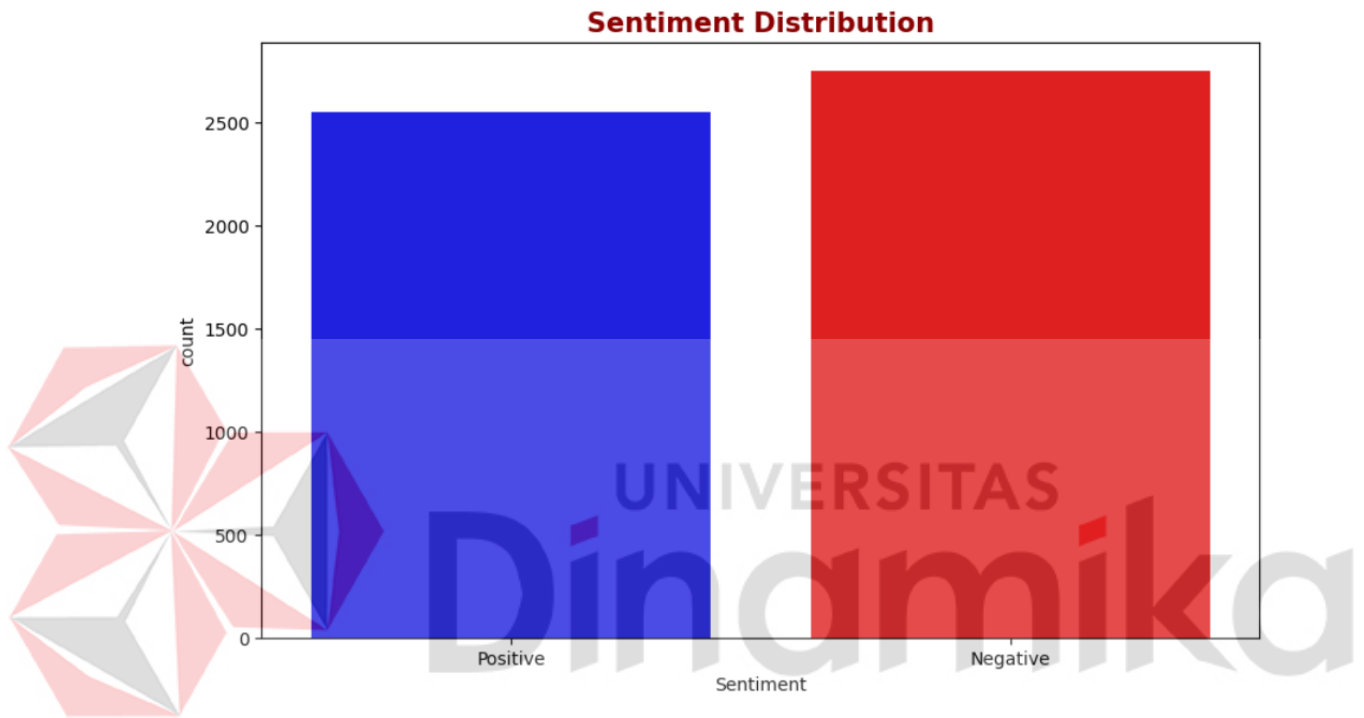
Gambar 4.7 Program *Wordcloud* positif





```
[ ] # B. Sentiment Distribution
plt.figure(figsize=(10, 6))
sns.countplot(x='Sentiment', data=df, palette=['blue', 'red'])
plt.title('Sentiment Distribution', fontsize = 15, fontweight = 'bold', color = 'darkred')
plt.savefig('Sentiment Distribution.png')
plt.show()
```

Gambar 4.11 Program dari *Setiment Distribution*



Gambar 4.12 Hasil dari program *Setiment Distribution*

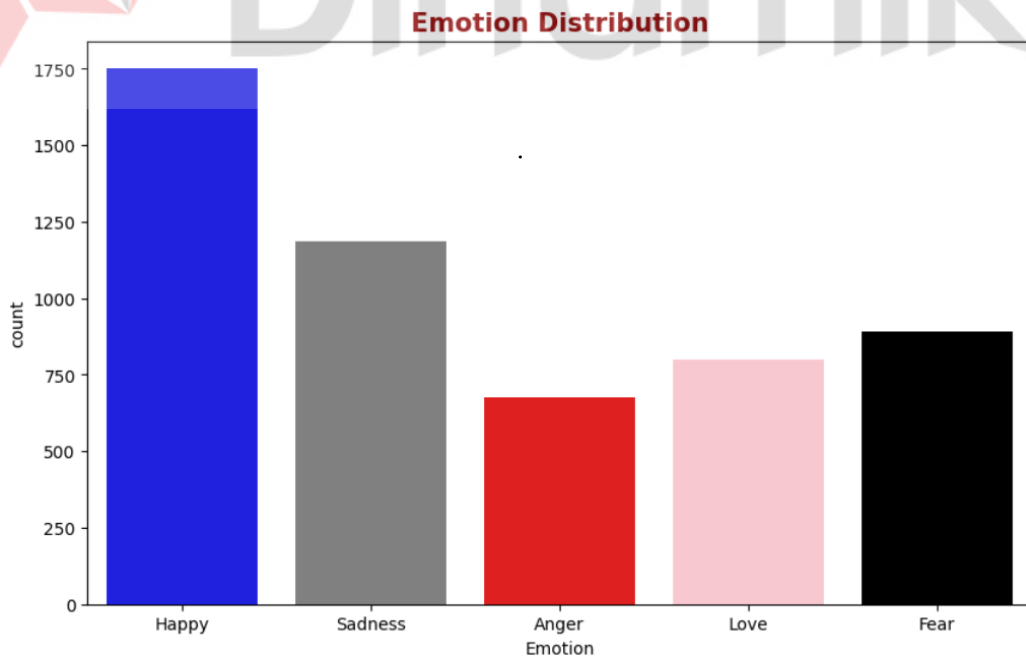
#### 4. Emotion Distribution

*Emotion Distribution* adalah digunakan untuk menunjukkan seberapa sering setiap kategori emosi muncul dalam data.

Pada program Gambar 4.13 digunakan untuk menggambarkan bagaimana distribusi emosi dalam dataset. Ini membantu dalam memahami seberapa banyak data yang termasuk dalam setiap kategori emosi. Hasil dari program dapat dilihat pada Gambar 4.14 bahwa perasaan negatif lebih banyak seperti *sadness*, *anger* dan *fear* sedangkan perasaan positif hanya *happy* dan *love*.

```
# B. Emotion Distribution
plt.figure(figsize=(10, 6))
sns.countplot(x='Emotion', data=df, palette=['blue','grey','red','pink','black'])
plt.title('Emotion Distribution', fontsize = 15, fontweight = 'bold', color = 'darkred')
plt.savefig('Emotion Distribution.png')
plt.show()
```

Gambar 4.13 Program dari *Emotion Distribution*

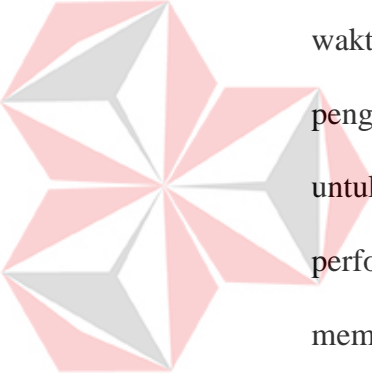


Gambar 4.14 Hasil dari program *Emotion Distribution*

### 4.4.3 Training model

Program pada Gambar 4.15 digunakan untuk melatih model dengan menerapkan *early stopping*. *Early stopping* akan menghentikan pelatihan jika model tidak menunjukkan peningkatan dalam indeks kinerja. Manfaat dari *early stopping* adalah mencegah overfitting dan menghemat waktu pelatihan dengan menghentikan proses lebih awal ketika model telah mencapai performa optimal.

Jika tanpa menggunakan metode *early stopping* dan dengan melatih model selama 30 *epoch* maka akan ada risiko overfitting, peningkatan waktu pelatihan, potensi penurunan kinerja setelah titik tertentu, dan penggunaan sumber daya yang lebih besar. *Early stopping* membantu untuk menghentikan pelatihan lebih awal jika model sudah mencapai performa optimal, sehingga mengurangi risiko-risiko tersebut dan mempercepat proses pelatihan. Untuk training model tanpa *early stopping* dapat dilihat pada Gambar 4.16.



```
# Latih model dengan early stopping
history = model.fit(X_train_pad, y_train_encoded,
                    epochs=20,
                    validation_data=(X_test_pad, y_test_encoded),
                    batch_size=32,
                    callbacks=[early_stopping]) # Menggunakan early stopping

Epoch 1/20
117/117 [=====] - 97s 419ms/step - loss: 0.5976 - accuracy: 0.7167 - val_loss: 0.3129 - val_accuracy: 0.8832
Epoch 2/20
117/117 [=====] - 57s 489ms/step - loss: 0.2115 - accuracy: 0.9292 - val_loss: 0.2573 - val_accuracy: 0.9058
Epoch 3/20
117/117 [=====] - 46s 392ms/step - loss: 0.2214 - accuracy: 0.9434 - val_loss: 0.2745 - val_accuracy: 0.8982
Epoch 4/20
117/117 [=====] - 45s 383ms/step - loss: 0.1028 - accuracy: 0.9736 - val_loss: 0.2986 - val_accuracy: 0.8876
Epoch 5/20
117/117 [=====] - 45s 389ms/step - loss: 0.0795 - accuracy: 0.9782 - val_loss: 0.2959 - val_accuracy: 0.9045
```

Gambar 4.15 Training model menggunakan *early stopping*



```
history = model.fit(X_train_padded, y_train_encoded,
                    epochs=30,
                    validation_data=(X_test_padded, y_test_encoded),
                    batch_size=32)
Epoch 27/30
117/117 [=====] - 33s 283ms/step - loss: 0.0331 - accuracy: 0.9890 - val_loss: 0.4702 - val_accuracy: 0.8882
Epoch 28/30
117/117 [=====] - 35s 303ms/step - loss: 0.0252 - accuracy: 0.9903 - val_loss: 0.4418 - val_accuracy: 0.8932
Epoch 29/30
117/117 [=====] - 35s 303ms/step - loss: 0.0229 - accuracy: 0.9919 - val_loss: 0.4509 - val_accuracy: 0.8920
Epoch 30/30
117/117 [=====] - 35s 299ms/step - loss: 0.0221 - accuracy: 0.9935 - val_loss: 0.4635 - val_accuracy: 0.8907
```

Gambar 4.16 *Training model* tanpa menggunakan *early stopping*

#### 4.4.4 *Evaluation Model*

Pada Gambar 4.17 merupakan program evaluation model yang digunakan untuk menilai seberapa baik model bekerja setelah pelatihan. Dengan mengevaluasi model pada data uji, kita dapat melihat seberapa akurat model dalam membuat prediksi pada data yang belum pernah dilihat sebelumnya dan seberapa baik model tersebut dalam beradaptasi dengan data baru. Hasil evaluasi ini membantu kita memahami kinerja model dan menentukan apakah perlu penyesuaian lebih lanjut.

Pada program Gambar 4.18 digunakan untuk visualisasi hasil pelatihan model dengan matplotlib yaitu untuk membuat grafik dan gambar dari data kita, membantu kita memahami data dengan lebih baik dan sklearn yaitu untuk membuat dan mengevaluasi model pembelajaran mesin, memungkinkan kita untuk membuat prediksi atau keputusan berdasarkan data.

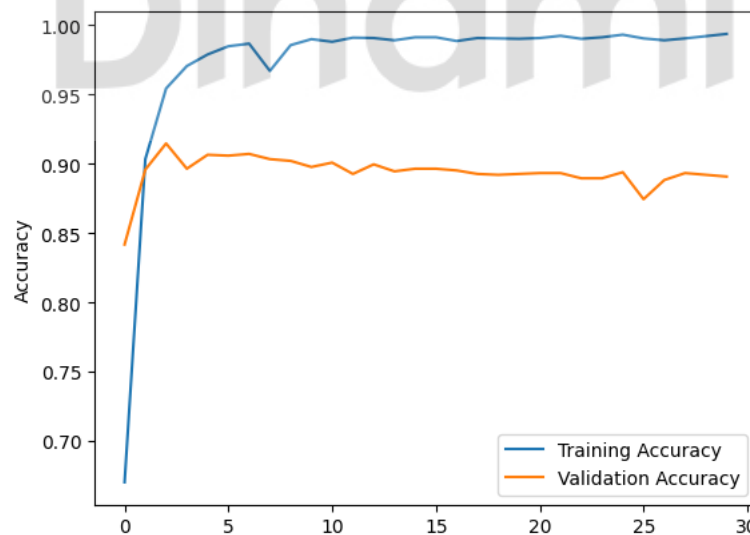
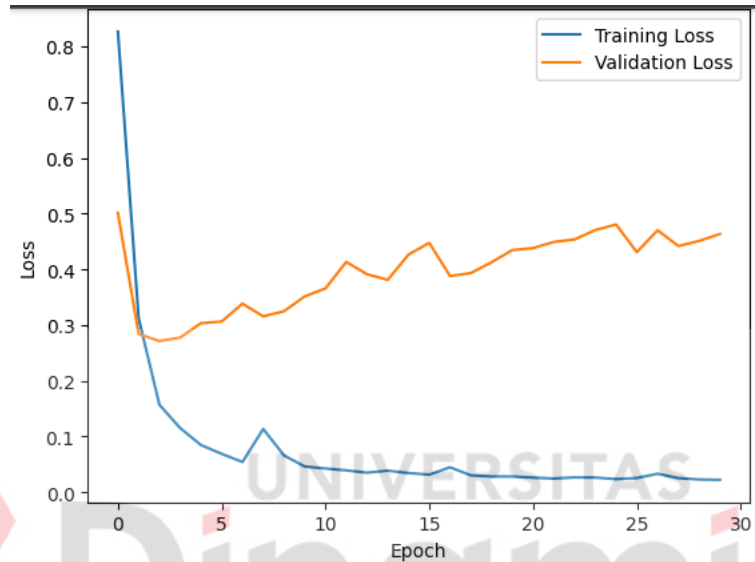
Bedasarkan hasil grafik *evaluation model* pada gambar 4.18 dapat diketahui bahwa hasil *accuracy* sebesar 89.07% dan tingkat *loss* sebesar 0.4

```

] # Evaluasi model
  loss, accuracy = model.evaluate(X_test_padded, y_test_encoded)
  print("Test Accuracy:", accuracy)

```

Gambar 4.17 Program dari *evaluation model*



```

50/50 [=====] - 4s 73ms/step - loss: 0.4635 - accuracy: 0.8907
Test Accuracy: 0.8907034993171692

```

Gambar 4.18 Hasil dari program *evaluation model*

#### 4.4.5 *Deployment*

Tujuan *deployment* ke *IBM Watson Space* adalah untuk membuat model yang telah dilatih bisa digunakan secara online oleh aplikasi atau pengguna lain. Dengan *deployment*, model bisa menerima data baru dan memberikan prediksi secara real-time melalui internet. Ini memudahkan integrasi model ke dalam sistem atau aplikasi yang membutuhkan prediksi secara otomatis dan cepat.



UNIVERSITAS  
**Dinamika**

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Proyek ini berhasil menerapkan analisis sentimen terhadap ulasan pelanggan di *platform e-commerce* Tokopedia menggunakan metode *Natural Language Processing* (NLP) dan model *Long Short-Term Memory* (LSTM). Hasilnya menunjukkan bahwa teknik NLP yang diterapkan, seperti tokenisasi, *stemming*, *lemmatization*, dan penghapusan *stop words*, mampu memproses teks dengan efektif. Model LSTM yang digunakan berhasil mengklasifikasikan sentimen dengan tingkat akurasi yang cukup tinggi, yaitu: 89%, menunjukkan kinerja yang lebih baik dibandingkan model tradisional. Wawasan yang diperoleh dari analisis ini memberikan gambaran mendalam tentang persepsi pelanggan terhadap produk dan layanan di Tokopedia, yang dapat digunakan untuk meningkatkan kualitas layanan dan strategi pemasaran perusahaan.

#### **5.2 Saran**

1. Untuk meningkatkan akurasi dan kinerja model, disarankan untuk mengeksplorasi dan mengimplementasikan teknik machine learning lainnya, seperti Transformer atau BERT, yang memiliki kemampuan lebih baik dalam memahami konteks dan dependensi jangka panjang dalam teks.
2. Menambahkan lebih banyak data ulasan pelanggan dari berbagai sumber dan periode waktu yang berbeda dapat membantu meningkatkan generalisasi model. Ini

juga dapat mencakup ulasan dalam berbagai bahasa untuk memperluas cakupan analisis.

3. Selain analisis sentimen, melakukan analisis topik pada ulasan pelanggan dapat memberikan wawasan lebih lanjut tentang isu-isu spesifik yang sering muncul dalam ulasan. Hal ini dapat membantu perusahaan mengidentifikasi area yang perlu diperbaiki lebih detail.
4. Untuk memastikan implementasi yang lebih efektif dan efisien, disarankan untuk menerapkan sistem ini di lingkungan produksi yang lebih terintegrasi, dengan menggunakan platform seperti *IBM Watson Studio* atau alternatif lain yang sesuai dengan kebutuhan perusahaan.



UNIVERSITAS  
**Dinamika**

## Daftar Pustaka

- Guetterman, T. C., Chang, T., DeJonckheere, M., Basu, T., Scruggs, E., & Vydiswaran, V. (2018). Augmenting Qualitative Text Analysis with Natural Language Processing: Methodological Study. *JOURNAL OF MEDICAL INTERNET RESEARCH*, 40, 1-13. doi:10.2196/jmir.9702
- J., A. S., & K., K. P. (2020, August). A Literature Review on Application of Sentiment Analysis Using Machine Learning Techniques. *International Journal of Applied Engineering and Management Letters (IJAEML)*, 4, 41-77. doi:http://doi.org/10.5281/zenodo.3977576
- Li, Y., & Cao, H. (2018). Prediction for Tourism Flow based on LSTM Neural Network. *Procedia Computer Science*, 129, 277-283. doi:https://doi.org/10.1016/j.procs.2018.03.076



UNIVERSITAS  
Dinamika