



**ANALISIS SENTIMEN APLIKASI GOJEK DENGAN MENGGUNAKAN
*NATURAL LANGUAGE PROCESSING***

LAPORAN KERJA PRAKTIK



Oleh:

FEBRIAN DWI SUSANTO

21410200017

FAKULTAS TEKNOLOGI DAN INFORMATIKA

UNIVERSITAS DINAMIKA

2024

**ANALISIS SENTIMEN APLIKASI GOJEK DENGAN MENGGUNAKAN
*NATURAL LANGUAGE PROCESSING***

Diajukan sebagai salah satu syarat untuk menyelesaikan
Program Sarjana



Disusun Oleh:

Nama : FEBRIAN DWI SUSANTO

NIM : 21410200017

Program : S1 (Strata Satu)

Jurusan : Teknik Komputer

FAKULTAS TEKNOLOGI DAN INFORMATIKA

UNIVERSITAS DINAMIKA

2024



"Salah satu cara melakukan pekerjaan yang hebat adalah dengan mencintai apa yang kamu lakukan"

- **B. J. Habibie**

UNIVERSITAS
Dinamika

LEMBAR PENGESAHAN

**ANALISIS SENTIMEN APLIKASI GOJEK DENGAN MENGGUNAKAN
NATURAL LANGUAGE PROCESSING**

Laporan Kerja Praktik oleh

Febrian Dwi Susanto

NIM: 21410200017

Telah diperiksa, diuji, dan disetujui



Surabaya, 24 Juli 2024

Disetujui:

Pembimbing

Penyelia

cn=Harianto Harianto,
o=Universitas Dinamika,
ou=Prodi S1 Teknik Komputer,
email=hari@dinamika.ac.id,
c=ID
2024.07.31 21:26:26 +07'00'

Harianto, S.Kom., M.Eng.
NIDN. 0722087701

Lutfi Dwimulva

Mengetahui,

Ketua Program Studi S1 Teknik Komputer

cn=Pauladie Susanto, o=Universitas
Dinamika, ou=PS S1 Teknik
Komputer,
email=pauladie@dinamika.ac.id,
c=ID
2024.08.02 08:42:40 +07'00'

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

**PERNYATAAN
PERSETUJUAN PUBLIKASI DAN KEASLIAN KARYA ILMIAH**

Sebagai mahasiswa **Universitas Dinamika**, Saya :

Nama : **Febrian Dwi Susanto**
NIM : **21410200017**
Program Studi : **S1 Teknik Komputer**
Fakultas : **Fakultas Teknologi dan Informatika**
Jenis Karya : **Laporan Kerja Praktik**
Judul Karya : **ANALISIS SENTIMEN APLIKASI GOJEK DENGAN
MENGUNAKAN *NATURAL LANGUAGE
PROCESSING***

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, 26 Juli 2024



Febrian Dwi Susanto
NIM : 21410200017

ABSTRAK

Aplikasi layanan *on-demand* seperti Gojek telah menjadi bagian penting dari kehidupan masyarakat Indonesia. Namun, dengan pertumbuhan penggunaan, muncul tantangan dalam memahami dan merespons persepsi pengguna yang beragam. Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi Gojek dengan memanfaatkan teknologi *Natural Language Processing (NLP)*. Masalah yang dihadapi adalah tingginya volume ulasan dan opini yang sangat besar sehingga sangat tidak memungkinkan untuk melihat ulasan tersebut secara manual. Untuk mengatasi masalah ini, data yang didapatkan dari ulasan pengguna bersumber dari *website kaggle*. Proses analisis melibatkan beberapa tahap, termasuk pengumpulan data, *pre-processing* teks, dan penerapan algoritma klasifikasi sentimen. Hasil dari penelitian ini mencapai tingkat akurasi analisis sebesar 85% dan dapat memberikan kontribusi dalam meningkatkan pemahaman terhadap sentimen pengguna secara *real time*, serta menjadi dasar untuk pengembangan aplikasi yang lebih canggih dalam bidang analisis sentimen.

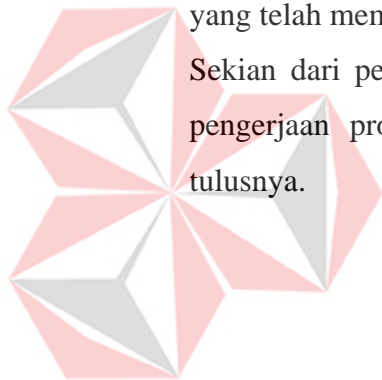
Kata Kunci: Analisis Sentimen, *Natural Language Processing*, Gojek.

KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa yang telah memberikan penulis kesehatan dan kemudahan sehingga menyelesaikan Kerja Praktik serta membuat laporan Kerja Praktik. Laporan ini dibuat berdasarkan Kerja Praktik serta hasil proyek yang dilakukan selama lima bulan di PT Hacktivate Teknologi Indonesia. Laporan Kerja Praktik ini membahas tentang Analisis Sentimen Aplikasi Gojek Dengan Menggunakan *Natural Language Processing*.

Penulis sangat berterima kasih kepada instruktur Kakak Sardi Irfansyah dan Kakak Louis Madaerdo sebagai mentor yang telah memberikan materi tentang *Artificial Intellegence*, serta penulis juga sangat berterima kasih kepada Bapak Harianto, S.Kom., M.Eng. sebagai dosen pembimbing laporan Kerja Praktik dan Bapak Pauladie Susanto, S.Kom., M.T. selaku ketua program studi S1 Teknik Komputer yang telah memberikan izin kepada penulis untuk melakukan Kerja Praktik.

Sekian dari penulis, apabila terdapat penulisan yang kurang dipahami maupun pengerjaan proyek kurang sempurna, penulis memohon maaf yang setulus-tulusnya.



UNIVERSITAS
Dinamika

Surabaya, 23 Juli 2024

Penulis

DAFTAR ISI

ABSTRAK	1
KATA PENGANTAR.....	2
DAFTAR ISI	3
DAFTAR TABEL.....	5
DAFTAR GAMBAR	6
DAFTAR LAMPIRAN	7
BAB I PENDAHULUAN	8
1.1 Latar Belakang.....	8
1.2 Rumusan Masalah.....	9
1.3 Batasan Masalah	9
1.4 Tujuan Penelitian	9
1.5 Manfaat Penelitian	9
BAB II GAMBARAN UMUM PERUSAHAAN.....	10
2.1 Latar Belakang Perusahaan.....	10
2.3 Profil Perusahaan	12
2.4 Visi dan Misi Perusahaan.....	12
2.5 Struktur Organisasi Perusahaan	13
BAB III LANDASAN TEORI.....	14
3.1 Natural Language Processing	14
3.2 Aplikasi Gojek	16
BAB IV DESKRIPSI PEKERJAAN	17
4.1 Kerja Praktik	17
4.2 Metode Pembelajaran Kerja Praktik	17
4.3 Lingkup Pekerjaan	18

4.4 Metode Penelitian	18
4.5 Deskripsi Capstone Project.....	19
4.6 Data Analisis Sentimen	19
4.7 Deskripsi Tugas Anggota	21
4.8 Tahap Pengerjaan	22
4.8.1 Memasukkan Library	22
4.8.2 Tahap Prepreprosesan Data	23
4.8.3 Tahap Text Preprocessing.....	24
4.8.4 Menampilkan Visualisasi	25
4.8.5 Tahap Undersampling	28
4.8.6 Tahap Long-Short Term Memory.....	28
4.9 Testing.....	31
BAB V PENUTUP.....	34
5.1 Kesimpulan	34
5.2 Saran	34
DAFTAR PUSTAKA	35
LAMPIRAN.....	36



DAFTAR TABEL

Tabel 4.1 Deskripsi Tugas Anggota	21
---	----



UNIVERSITAS
Dinamika

DAFTAR GAMBAR

Gambar 2. 1 Logo Kampus Merdeka.....	10
Gambar 2. 2 Logo PT Hacktivate Teknologi Indonesia.....	11
Gambar 2.3 Lokasi Kantor Hacktiv8	12
Gambar 2.4 Struktur Organisasi Hacktiv8	13
Gambar 3.1 Natural Language Processing.....	14
Gambar 4.1 Logo Gojek.....	19
Gambar 4.2 Sumber Data Analisis	20
Gambar 4.3 Timeline Pengerjaan Capstone Project.....	22
Gambar 4.4 Program untuk memunculkan visualisasi kata negatif.....	26
Gambar 4.5 Visualisasi kata-kata negatif.....	26
Gambar 4.6 Program untuk menampilkan visualisasi data positif.....	26
Gambar 4.7 Visualisasi kata-kata positif.....	27
Gambar 4.8 Program untuk menampilkan visualisasi data positif dan negatif.....	27
Gambar 4.9 Hasil dalam bentuk diagram batang visualisasi kata negatif dan.....	27
Gambar 4.10 Hasil diagram dengan menggunakan Undersampling.....	28
Gambar 4.11 Menampilkan kurva akurasi dan loss	29
Gambar 4.12 Program untuk menampilkan ROC	30
Gambar 4.13 Hasil analisis dengan ROC.....	30
Gambar 4.14 Program Training data untuk menampilkan LSTM	31
Gambar 4.15 Input Program Streamlit	32
Gambar 4.16 Hasil analisis ulasan positif.....	33
Gambar 4.17 Hasil sentimen analisis ulasan negatif.....	33

DAFTAR LAMPIRAN

Lampiran 1 Surat Balasan Dari Perusahaan.....	36
Lampiran 2 Logbook Bulanan Studi Independen	37
Lampiran 3 Sertifikat Kepesertaan & Mitra	43
Lampiran 4 Nilai Studi Independen	44
Lampiran 5 Form Bimbingan KP.....	45
Lampiran 6 Biodata Penulis	46



UNIVERSITAS
Dinamika

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam era digital yang berkembang pesat, aplikasi layanan *on-demand* seperti Gojek telah menjadi bagian penting dari kehidupan masyarakat Indonesia. Dengan berbagai layanan yang ditawarkan, mulai dari transportasi hingga pengiriman makanan, Gojek berhasil menarik jutaan pengguna aktif. Setiap harinya, pengguna memberikan ulasan dan feedback yang berharga di berbagai platform media sosial dan aplikasi. Ulasan ini mencerminkan pengalaman pengguna serta persepsi mereka terhadap kualitas layanan yang diberikan. Namun, volume data ulasan yang sangat besar dan beragamnya opini yang disampaikan menjadi tantangan tersendiri dalam memahami dan menganalisis sentimen pengguna. Ketidakpastian dalam menilai sentimen ini dapat mempersulit pengambilan keputusan yang strategis untuk meningkatkan layanan dan pengalaman pengguna. Di sinilah peran teknologi *Natural Language Processing (NLP)* menjadi sangat relevan. *Natural Language Processing* memungkinkan analisis otomatis dan efisien dari teks tidak terstruktur, sehingga membantu mengidentifikasi pola sentimen positif, negatif, atau netral.

Penelitian ini fokus pada analisis sentimen pengguna terhadap aplikasi Gojek menggunakan teknik *Natural Language Processing*. Dengan memahami sentimen pengguna secara lebih mendalam, penelitian ini bertujuan untuk memberikan wawasan berharga bagi pengembang aplikasi Gojek. Wawasan ini dapat digunakan untuk mengidentifikasi isu-isu utama yang dihadapi pengguna serta memberikan rekomendasi yang dapat diterapkan untuk meningkatkan kualitas layanan dan kepuasan pengguna.

1.2 Rumusan Masalah

Dalam analisis ulasan ini berdasarkan latar belakang, bahwa rumusan masalah pada project ini adalah sebagai berikut:

1. Bagaimana efektivitas teknik *Natural Language Processing* dalam mengkategorikan dan menilai sentimen dalam ulasan pengguna terhadap aplikasi Gojek?

1.3 Batasan Masalah

Dalam analisis ulasan ini berdasarkan latar belakang, bahwa Batasan masalah pada project ini adalah sebagai berikut:

1. Sample dataset Gojek sebanyak Sepuluh Ribu ulasan diambil dari *website Kaggle*.
2. Jumlah ulasan yang dianalisis mungkin dibatasi oleh ketersediaan data yang relevan dan representatif dari populasi pengguna.
3. Data ulasan yang dianalisis dibatasi pada periode tertentu untuk memberikan gambaran tren yang relevan dan terbaru.

1.4 Tujuan Penelitian

Tujuan penelitian dari proyek tersebut adalah menyelesaikan masalah analisis sentimen aplikasi Gojek untuk menentukan ulasan negatif, positif, dan netral dengan menggunakan *Natural Language Processing*.

1.5 Manfaat Penelitian

1. Mampu menerapkan *Machine Learning* untuk mengolah sebuah banyaknya data dari ulasan.
2. Proyek ini bisa membantu Gojek untuk mempercepat pengolahan review pengguna pada aplikasinya untuk meningkatkan kepuasan pelanggan.
3. Sentimen analisis, dari *wordcloud* bisa membantu Gojek cepat menemukan fitur aplikasi mana yang mendapatkan paling banyak tanggapan positif dan fitur yang membutuhkan improvisasi/perbaikan segera.
4. Mengetahui bahwa penggunaan *Exploratory Data Analyst* mampu menunjukkan bahwa data positif lebih dominan.

BAB II

GAMBARAN UMUM PERUSAHAAN

2.1 Latar Belakang Perusahaan

Program kegiatan Merdeka Belajar Kampus Merdeka merupakan sebuah program yang diluncurkan oleh Kemendikbudristek untuk mentransformasi system Pendidikan tinggi di Indonesia serta untuk menghasilkan lulusan yang lebih relevan. Tujuan diluncurkan Merdeka Belajar Kampus Merdeka adalah untuk mengasah kemampuan *hardskills* dan *softskills* pada sebuah Perusahaan juga membantu meningkatkan pengembangan potensi yang dimiliki dengan sesuai passion dan bakat melalui program *experiential learning* dengan jalur yang sangat fleksibel, dalam program ini juga sangat memungkinkan mahasiswa untuk mengambil program di luar kegiatan perkuliahan dan bisa mengubah beban studi atau SKS yang dapat diakui sesuai peraturan yang ada. Mahasiswa juga mampu mengembangkan inovasi untuk membantu dalam memecahkan masalah yang terdapat pada Perusahaan tersebut.



Gambar 2. 1 Logo Kampus Merdeka

(Sumber: <https://kampusmerdeka.kemdikbud.go.id/>)



Gambar 2. 2 Logo PT Hactivate
Teknologi Indonesia

(Sumber: <https://www.hactiv8.com/>)

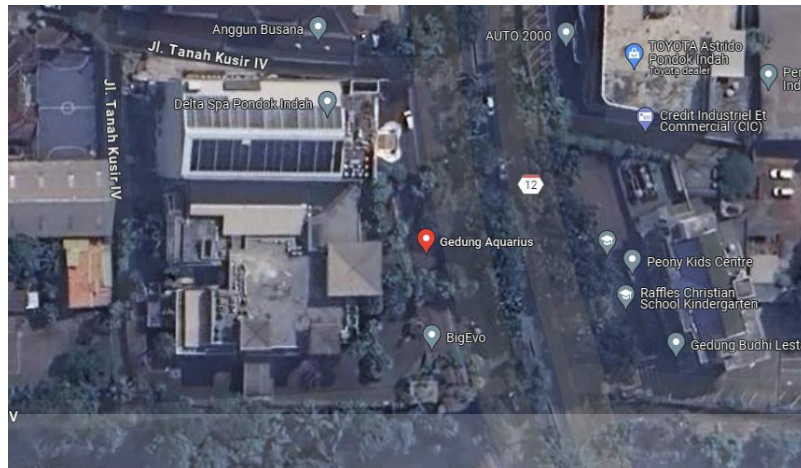
PT Hactivate Teknologi Indonesia merupakan program *reskilling* yang mendalam (*bootcamp*) yang mengubah pemula menjadi talenta digital siap kerja sejak tahun 2016, Hactiv8 telah berdedikasi mengatasi kesenjangan keterampilan teknologi melalui kurikulum komprehensif dan praktis yang mempersiapkan siswa untuk sukses berkarir serta berkontribusi pada pertumbuhan ekonomi di Indonesia. Program Hactiv8 menyediakan pelatihan intensif dalam bidang pengembangan *web full-stack*, ilmu data, dan berbagai keterampilan teknologi lainnya. PT Hactivate Teknologi Indonesia menjalankan program MSIB, dengan judul program “*IBM SkillsBuild for AI & Cybersecurity*”. Dengan rincian terkait program MSIB sebagai berikut:

1. Durasi aktivitas : 16 Februari – 30 Juni 2024
2. Jumlah kredit SKS : 20 SKS
3. Tipe aktivitas : Online (Daring)
4. Lokasi aktivitas : Online (Daring)

Program ini memberikan pengalaman di bidang *Artificial Intelligence* untuk membuat pemrograman *Natural Language Processing* dan *Computer Vision*.

2.3 Profil Perusahaan

Nama Instansi : PT Hacktivate Teknologi Indonesia
Alamat : Gedung Aquarius Lt. 1&2, Jl. Sultan Iskandar Muda No. 7,
RT. 005 RW. 009, Kebayoran Lama Selatan, Kebayoran
Lama, Jakarta Selatan 12240.



Gambar 2.3 Lokasi Kantor Hacktiv8



No Telp
Website
Email

: (021) 8067 5787

: <https://www.hacktiv8.com/>

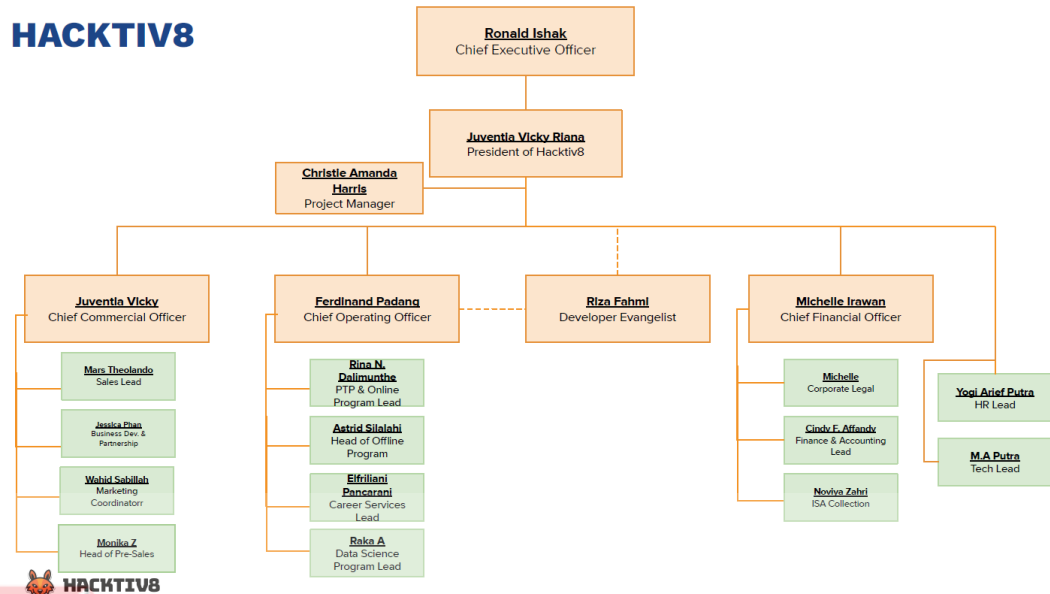
: halo@hacktiv8.com

2.4 Visi dan Misi Perusahaan

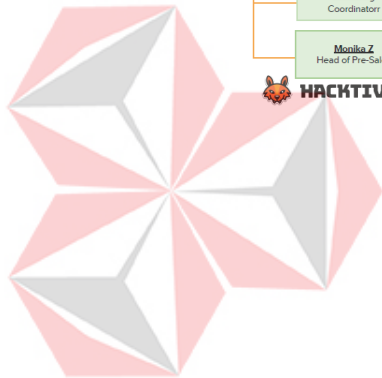
Visi PT Hacktivate Teknologi Indonesia adalah Pendidikan intensif atau disebut bootcamp yang mentransformasi pemula menjadi talenta digita siap kerja. Misi PT Hacktivate Teknologi Indonesia adalah Melalui solusi Hacktiv8, kami mewujudkannya agar orang dapat mempelajari keterampilan ini dengan cepat, ikuti terus dengan laju perubahan, dan berkembang di era digital.

2.5 Struktur Organisasi Perusahaan

Berikut adalah struktur organisasi Perusahaan PT Hacktivate Teknologi Indonesia:



Gambar 2.4 Struktur Organisasi Hacktiv8



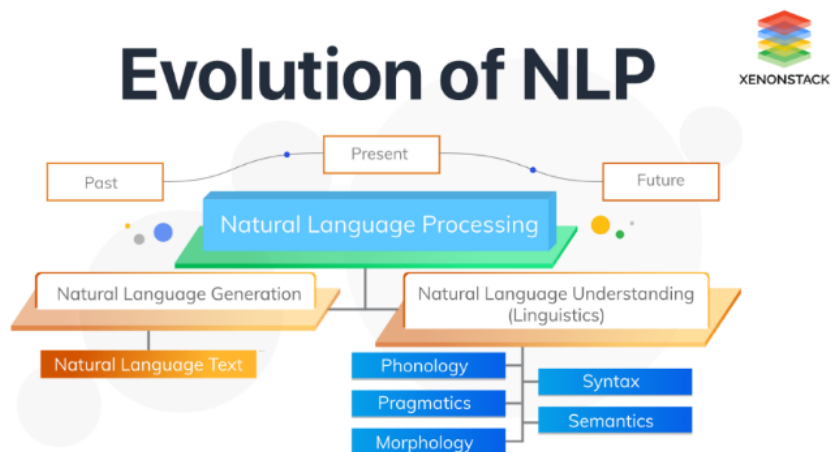
UNIVERSITAS
Dinamika

BAB III

LANDASAN TEORI

3.1 Natural Language Processing

Natural Language Processing merujuk pada cabang ilmu komputer dan lebih khususnya merupakan cabang dari *Artificial Intelligence* yang berkaitan dengan pemberian kemampuan kepada komputer untuk memahami teks dan kata-kata lisan dengan cara yang sama seperti manusia. *Natural Language Processing* sebagai disiplin ilmu yang melibatkan penggunaan komputer untuk memproses, menganalisis, dan menghasilkan bahasa manusia. Kemudian menekankan pada pemahaman dan representasi bahasa alami, mencakup aspek sintaksis, semantik, dan pragmatik. Perkembangan teknik pembelajaran mesin dalam *Natural Language Processing*, termasuk model berbasis statistik dan transformasi (Jurafsky, D. & Martin, J. H. 2022). *Natural Language Processing* menggabungkan linguistic komputasional pemodelan Bahasa manusia berbasis aturan dengan model statistik, pembelajaran mesin, dan pembelajaran mendalam atau *Deep Learning*. Teknologi ini memungkinkan komputer untuk melakukan proses bahasa manusia dalam bentuk teks atau data suara untuk memahami makna lengkap dengan maksud dan sentiment pembicara atau penulis.



Gambar 3.1 Natural Language Processing

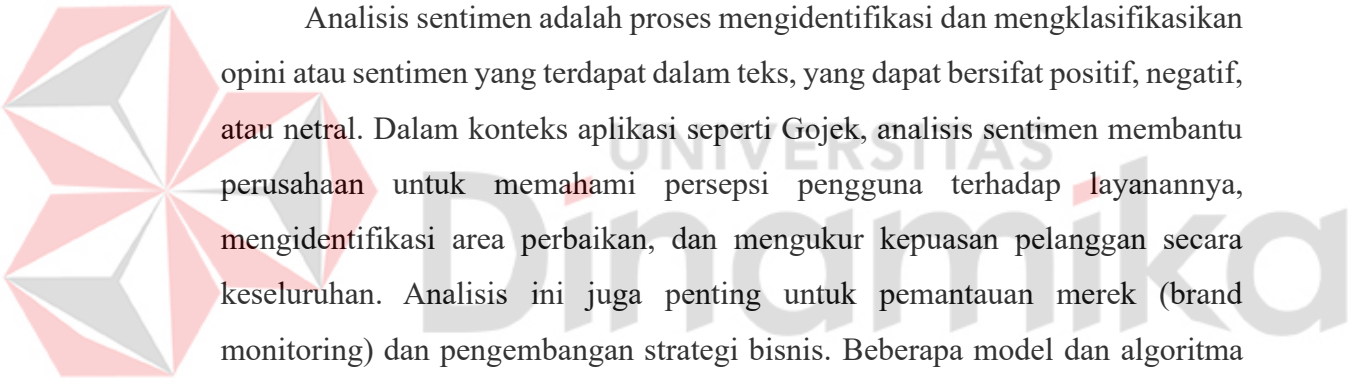
(Sumber: <https://www.xenonstack.com/blog/evolution-of-nlp>)

Natural Language Processing menggerakkan program komputer yang menerjemahkan teks dari satu bahasa ke bahasa lain, menanggapi perintah lisan, dan meringkas teks dalam jumlah besar dengan cepat bahkan secara *real time*. Ada kemungkinan besar telah berinteraksi dengan *Natural Language Processing* dalam bentuk sistem yang dioperasikan dengan suara, asisten digital, perangkat lunak yang diucapkan ke teks. *Natural Language Processing* bertujuan untuk memungkinkan komputer untuk memahami, memproses, dan menghasilkan bahasa manusia secara efektif. Berikut adalah beberapa bagian utama dari *Natural Language Processing*:

1. **Tokenisasi:** Tokenisasi melibatkan pemisahan teks menjadi unit-unit yang lebih kecil, seperti kata-kata atau frasa. Ini adalah langkah awal dalam pemrosesan teks.
2. **Text Cleaning:** Langkah ini mencakup menghapus karakter khusus, tanda baca, dan mengonversi teks ke format standar. Ini membantu dalam menghasilkan representasi yang lebih bersih dan seragam dari teks.
3. **Morfologi:** Memahami struktur internal dari kata-kata dalam bahasa. Ini mencakup aspek seperti pemenggalan kata dan analisis bentuk kata.
4. **Parsing:** Parsing adalah proses analisis gramatikal teks untuk menentukan struktur grammatikalnya, seperti hubungan antara kata-kata dalam kalimat.
5. **Semantik:** Memahami makna dari teks. Ini mencakup memahami hubungan antara kata-kata, frasa, atau kalimat dalam konteks tertentu.
6. **Klasifikasi dan pengenalan entitas:** Klasifikasi teks melibatkan mengkategorikan teks ke dalam kelas-kelas yang sudah ditentukan, sedangkan pengenalan entitas melibatkan mengidentifikasi dan mengekstraksi informasi tertentu seperti nama, tanggal, atau tempat dari teks.
7. **Evaluasi Sentimen:** Mengidentifikasi dan mengekstrak sentimen atau opini yang terkandung dalam teks, baik itu positif, negatif, atau netral.
8. **Pengelompokkan (*Clustering*):** Mengelompokkan dokumen atau teks berdasarkan kesamaan mereka dalam konteks tertentu.

3.2 Aplikasi Gojek

Gojek adalah aplikasi layanan *on-demand* yang populer di Indonesia dan beberapa negara Asia Tenggara lainnya. Gojek didirikan pada tahun 2010 oleh Nadiem Makarim, yang kemudian menjadi Menteri Pendidikan dan Kebudayaan Indonesia, kemudian Gojek memulai sebagai layanan ojek (taksi sepeda motor) di Jakarta, Indonesia. Seiring waktu, aplikasi ini berkembang untuk menawarkan berbagai layanan. Kepuasan pelanggan merupakan faktor krusial bagi kesuksesan dan keberlanjutan layanan berbasis aplikasi seperti Gojek. Pengguna sering kali memberikan ulasan atau review di platform seperti Google Play, App Store, dan media sosial untuk menyampaikan pengalaman mereka. Ulasan ini mencakup berbagai aspek layanan, termasuk kecepatan, kualitas layanan, harga, dan pengalaman pengguna.



Analisis sentimen adalah proses mengidentifikasi dan mengklasifikasikan opini atau sentimen yang terdapat dalam teks, yang dapat bersifat positif, negatif, atau netral. Dalam konteks aplikasi seperti Gojek, analisis sentimen membantu perusahaan untuk memahami persepsi pengguna terhadap layanannya, mengidentifikasi area perbaikan, dan mengukur kepuasan pelanggan secara keseluruhan. Analisis ini juga penting untuk pemantauan merek (brand monitoring) dan pengembangan strategi bisnis. Beberapa model dan algoritma yang sering digunakan dalam analisis sentimen seperti *Long Short-Term Memory (LSTM)*.

BAB IV

DESKRIPSI PEKERJAAN

4.1 Kerja Praktik

Kerja praktik merupakan sebuah program yang di khususkan untuk mahasiswa melakukan magang atau *internship* di sebuah Perusahaan milik swasta maupun Perusahaan milik negeri maupun organisasi yang sebagai bagian dari kurikulum pendidikannya. Program Kerja Praktik ini dirancang untuk memberikan mahasiswa pengalaman langsung dalam dunia kerja yang sesuai dengan bidang studinya. Tujuan utama Kerja Praktik adalah untuk mempersiapkan mahasiswa dengan keterampilan praktis yang dibutuhkan di dunia kerja setelah mahasiswa tersebut lulus. Selain itu, Kerja Praktik juga bisa menjadi kesempatan bagi mahasiswa untuk mengembangkan jaringan profesional, memahami bagaimana tentang dunia lingkungan kerja yang sesungguhnya, dan menerapkan teori yang dipelajari ke dalam situasi nyata

Selama dalam masa Kerja Praktik, diharuskan untuk memiliki tanggung jawab untuk melakukan pengembangan proyek agar dapat berjalan dengan sesuai jadwal yang sudah di tentukan. Komunikasi antar tim dalam Kerja Praktik sangat wajib untuk dilakukan, karena dengan adanya komunikasi yang baik dengan tim maka proyek akan berjalan sesuai jadwal dan agar tidak melakukan kesalahan dalam mengembangkan sebuah proyek, sehingga proyek tersebut selesai dengan sempurna. Selama dalam kesempatan kegiatan Kerja Praktik ini, harus memanfaatkan keterampilan komunikasi yang baik selama dalam pengerjaan proyek untuk memecahkan masalah yang ada di dalam proyek.

4.2 Metode Pembelajaran Kerja Praktik

Selama dalam kegiatan Kerja Praktik, terdapat metode pembelajaran untuk memahami materi yang akan dipelajari untuk mendapatkan pengetahuan. Metode-metode pembelajaran tersebut antara lain:

1. Akses Modul Pembelajaran.

Modul pembelajaran yang diberikan diakses melalui *website* kode.id beserta perintah penugasan serta panduan tentang *Capstone Project*.

2. Pembelajaran Kelas Daring.

Pelaksanaan untuk penjelasan materi dilakukan secara *online* atau daring dengan menggunakan platform *Google Meet*, kelas dibagi menjadi dua kelas, kelas pertama adalah *big class* penyampaian materi dijelaskan oleh instruktur, sedangkan kelas kedua adalah kelas mentoring penyampaian materi yang disampaikan oleh Mentor, kelas mentoring sangat berguna untuk berkomunikasi dengan mentor maupun dengan mahasiswa lain terkait pembahasan *Capstone Project*.

3. Google Classroom

Platform *Google Classroom* digunakan untuk mengumpulkan tugas yang diberikan oleh mentor dan penilaian tugas juga dilakukan menggunakan Platform *Google Classroom*.

4.3 Lingkup Pekerjaan

Program Magang dan Studi Independen Bersertifikat yang berada di perusahaan PT Hacktivate Teknologi Indonesia, nantinya setiap peserta akan mendaftar di platform kode.id untuk mendapatkan modul materi yang akan dipakai selama Studi Independen, setelah mendaftar akun kode.id maka peserta akan mendapatkan kode pembelajaran agar bisa menuju ke akses materi. Pembelajaran dilakukan secara online dengan menggunakan *Google Meet*, banyaknya peserta Studi Independen kelas daring dibagi menjadi dua kelas besar, yaitu Kelas Besar A dan Kelas Besar B instruktur materi setiap kelas berbeda dan juga *Mini Class* untuk sesi Bersama mentor. Pembelajaran semua dilakukan dengan daring dengan mengakses *IBM SkillsBuild* untuk sumber modul dan pengerjaan tugas dilakukan secara online dengan menggunakan *Google Classroom*.

4.4 Metode Penelitian

Pengerjaan proyek akhir dimulai dengan mengidentifikasi masalah dari beberapa ulasan aplikasi Gojek kemudian mengumpulkan data untuk

memecahkan masalah. Untuk memecahkan masalah analisis ini menggunakan *Natural Language Processing*

4.5 Deskripsi Capstone Project

Capstone Project merupakan bagian dari proyek akhir yang dilakukan mahasiswa dalam program Pendidikan tertentu, yang terdapat pada tingkat sarjana maupun magister. Proyek ini dirancang untuk mengintegrasikan dan mengaplikasikan pengetahuan yang didapat selama masa studi, serta menunjukkan kemampuan untuk memecahkan masalah dalam bidang studi yang relevan. *Capstone Project* ini.

Dalam proyek ini adalah menganalisis sentimen ulasan dari pengguna yang menggunakan aplikasi gojek, ulasan dibagi menjadi beberapa kategori yaitu ulasan positif, negatif, dan netral. Analisis ini berisikan tentang ulasan pengguna aplikasi Gojek terhadap *interface* yang terdapat pada aplikasi Gojek serta beberapa masalah lainnya yang muncul terkait aplikasi gojek. Aplikasi Gojek merupakan aplikasi penyedia jasa ojek berbasis *online* yang terdapat pada *smartphone*. Layanan gojek tidak hanya ojek saja namun terdapat beberapa jasa seperti *GoCar*, *GoFood*, serta *E-Wallet* yang diberi nama *Gopay* untuk bertransaksi di dalam aplikasi Gojek.

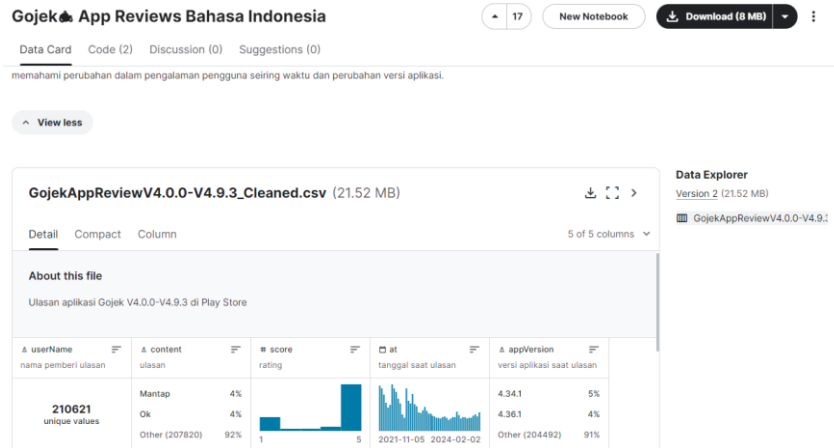


Gambar 4.1 Logo Gojek

(Sumber <https://www.gojek.com/blog/gojek/info-gojek>)

4.6 Data Analisis Sentimen

Data analisis sentimen ulasan bersumber dari *website Kaggle*. *Kaggle* merupakan sebuah platform *online* untuk menyediakan berbagai macam kompetensi dan proyek *data science*. Data yang dipakai dalam format CSV, format file CSV inilah yang berisikan tentang ulasan pengguna aplikasi gojek.



Gambar 4.2 Sumber Data Analisis

(Sumber: <https://www.kaggle.com/datasets/ucupsedaya/gojek-app-reviews-bahasa-indonesia/data>)

Pada dataset ini terdiri dari lima kolom utama yang mencakup informasi tentang pengguna aplikasi tersebut beserta ulasannya. Berikut adalah beberapa deskripsi singkat untuk setiap kolom:

1. Nama Pengguna: *Username* atau pemberi ulasan pada Gojek. Identifikasi unik atau nama samaran yang menunjukkan siapa yang memberikan ulasan.
2. Konten: Isi ulasan yang diberikan oleh pengguna. Kolom konten berisi teks dari ulasan yang mencakup pengalaman atau pendapat pengguna terkait aplikasi Gojek.
3. Skor: Skor yang diberikan oleh pengguna terhadap aplikasi Gojek. Nilai skor ini umumnya berkisar antara satu hingga lima, yang dimana satu merupakan menunjukkan ketidakpuasan dan lima menunjukkan kepuasan yang tinggi.
4. Tanggal: Tanggal ulasan diberikan oleh pengguna. Ini digunakan untuk mencatat kapan ulasan tersebut diunggah atau diberikan.
5. Versi Aplikasi: Versi aplikasi Gojek yang digunakan oleh pengguna saat memberikan ulasan. Informasi ini membantu untuk memahami konteks pembaruan aplikasi yang mungkin mempengaruhi pengalaman pengguna.

4.7 Deskripsi Tugas Anggota

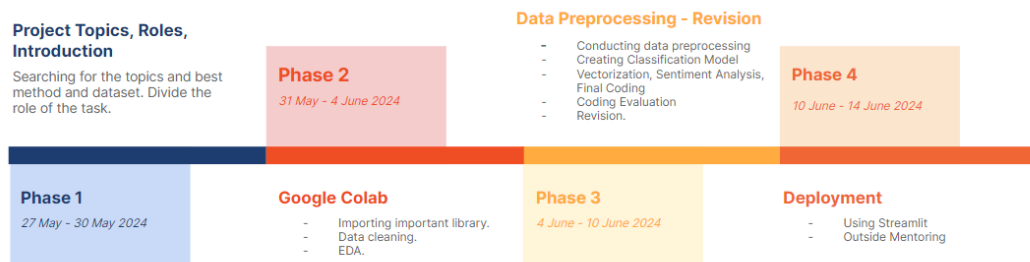
Pengerjaan *Capstone Project* dikerjakan secara berkelompok. Berikut adalah tabel terkait deskripsi pekerjaan setiap anggota kelompok:

Tabel 4.1 Deskripsi Tugas Anggota

Nama Kelompok	Deskripsi Pekerjaan
Fayza Nirwasita Ketua Kelompok	Mengawasi pekerjaan setiap anggota kelompok serta membantu dalam pembuatan pemrograman <i>Import Library</i> yang di butuhkan seperti <i>Pandas, Numpy, Matplotlib, Sklearn, Worldcloud</i> hingga <i>Tensor flow</i> .
Fakhirul Akmal Anggota Kelompok 1	Pembuatan bagian pemrograman <i>Data Preprocessing</i> yang meliputi <i>Check Nul, Handle Null, Check Duplicate</i> , dan <i>Handle Duplicate</i> . Pemrograman <i>Mapping</i> untuk membagi skor ulasan positif, negatif dan netral
Farrel Naufal Anggota Kelompok 2	Pembuatan pemrograman bagian <i>Text Preprocessing</i> yang meliputi <i>Stopword</i> untuk menghapus kata kata yang tidak diperlukan, <i>Tokenizing</i> untuk memecah kalimat menjadi kata-per-kata, <i>Stemming</i> agar kata menjadi lebih sederhana, dan melakukan standarisasi.
Febrian Dwi Susanto Anggota Kelompok 3	Pembuatan <i>Exploratory Data Analysis</i> untuk memvisualisasikan sentimen positif dan sentiment netral serta membuat distribusi sentiment sebelum dan sesudah menggunakan <i>undersampling</i>
Fadhillah Anggota Kelompok 4	Pembuatan program untuk <i>Long-Short Term Memory</i> untuk menampilkan model akurasi dan menampilkan grafik <i>Receiver Operating Characteristic</i>

4.8 Tahap Pengerjaan

Pengerjaan *Capstone Project* ini adalah dengan menggunakan platform *Google Colab* untuk membuat program analisis sentiment dengan menggunakan *Natural Language Processing*. Batas pengerjaan *Capstone Project* diberi waktu hingga tiga minggu.



Gambar 4.3 Timeline Pengerjaan *Capstone Project*

4.8.1 Memasukkan Library

Tahap pertama yang dilakukan adalah dengan *import library* dahulu untuk melakukan analisis sentimen, tujuan untuk diberi *library* ini adalah agar program bisa berjalan untuk melakukan analisis. Berikut adalah *library* yang digunakan untuk melakukan analisis :

1. *Pandas*: *library open-source* di Python yang sangat populer digunakan untuk manipulasi dan analisis data. *Library* ini menyediakan struktur data yang fleksibel dan efisien, terutama untuk bekerja dengan data tabular, seperti data dari *spreadsheet* atau *database*.
2. *Numpy*: *library open-source* di Python yang dirancang untuk komputasi numerik yang efisien. *Library* ini menyediakan dukungan untuk *array* multidimensi yang memungkinkan manipulasi data dengan cara yang cepat dan efisien. *Numpy* juga menyediakan sejumlah besar fungsi matematika dan aljabar linear yang berguna untuk analisis data, simulasi, dan pengembangan algoritma.
3. *Matplotlib*: *library visualisasi data* di Python yang sangat populer dan banyak digunakan. *Library* ini memungkinkan pengguna untuk membuat berbagai jenis grafik dan plot, seperti garis, batang, *scatter*, dan histogram.

Matplotlib sangat fleksibel dan dapat digunakan untuk membuat visualisasi yang sederhana hingga kompleks, termasuk visualisasi yang memerlukan penyesuaian detail, seperti label, warna, dan gaya.

4. *Sklearn: library open-source* di *Python* yang digunakan untuk *machine learning* dan analisis data. *Library* ini menyediakan alat yang mudah digunakan dan efisien untuk berbagai macam tugas *machine learning*, termasuk klasifikasi, regresi, pengelompokan, reduksi dimensi, dan evaluasi model. *scikit-learn* didasarkan pada *NumPy*, *SciPy*, dan *matplotlib*, sehingga sangat terintegrasi dengan ekosistem ilmiah *Python*.
5. *Wordcloud*: representasi visual dari data teks di mana kata-kata yang lebih sering muncul dalam teks ditampilkan lebih besar atau lebih tebal daripada kata-kata yang kurang sering. *Wordcloud* digunakan untuk memberikan gambaran visual yang cepat dan intuitif tentang frekuensi atau pentingnya kata-kata tertentu dalam sebuah teks.
6. *Natural Language Toolkit: Library open-source* di *Python* yang menyediakan alat untuk bekerja dengan data teks dalam bidang *Natural Language Processing (NLP)*. *Natural Language Toolkit* dirancang untuk memudahkan pengembangan aplikasi yang melibatkan pengolahan bahasa alami, seperti analisis teks, pemahaman bahasa, dan *Machine Learning* berbasis teks
7. *Tensorflow*: Merupakan *open-source library* untuk komputasi numerik dan *machine learning* yang awalnya dikembangkan oleh *Google Brain Team*. *TensorFlow* dirancang untuk memudahkan pengembangan dan *deployment model machine learning* dan *deep learning* pada berbagai platform, termasuk perangkat *mobile, desktop, server, dan cloud*.

4.8.2 Tahap Preprosesan Data

Tahap pertama yang dilakukan adalah dengan memasukkan data analisis ke dalam *Goggle Colab* untuk membuka format file CSV tujuan utama dilakukannya *Data Preprocessing* adalah agar data bisa lebih berpola dan muda di proses, dengan melakukan *checking null, duplicate* serta melakukan handlingnya. Kemudian melakukan *Mapping* untuk memertakan skor menjadi beberapa kategori seperti ulasan negatif, positif, dan netral, kategori tersebut

dibuat menjadi kolom terpisah. *Data preprocessing* adalah serangkaian langkah atau teknik yang dilakukan untuk membersihkan, mengorganisir, dan mempersiapkan data mentah sebelum diolah lebih lanjut oleh algoritma atau model *machine learning*. Tujuan utamanya adalah untuk meningkatkan kualitas data, memastikan data siap untuk analisis, dan meminimalkan kesalahan saat pemrosesan data. Berikut adalah beberapa langkah umum dalam data preprocessing:

1. Melakukan *Data Cleaning*: Mengatasi nilai-nilai yang hilang atau *missing values*, dan Mendeteksi dan memperbaiki data yang tidak valid atau tidak sesuai.
2. Transformasi data: Normalisasi atau standarisasi data untuk mengatasi skala yang berbeda, dan Transformasi variabel untuk memenuhi asumsi model (misalnya, transformasi logaritmik untuk data yang tidak terdistribusi normal).
3. Reduksi dimensi: Mengurangi jumlah fitur atau atribut untuk mengurangi kompleksitas model dan meningkatkan efisiensi.
4. Pemilihan fitur: Memilih fitur atau atribut yang paling relevan dan penting untuk membangun model yang efektif, dan Mengurangi overfitting dan meningkatkan akurasi model.
5. *Data Integration*: Menggabungkan data dari berbagai sumber untuk menciptakan dataset yang lebih lengkap dan bermakna.
6. Mengelompokkan data: Mengubah variabel kontinu menjadi variabel diskrit, misalnya untuk analisis kategorikal.

Setiap langkah ini penting untuk memastikan bahwa data yang digunakan untuk machine learning atau analisis data lainnya adalah valid, terstruktur dengan baik, dan siap untuk diproses oleh model atau algoritma yang akan digunakan. Dengan melakukan data preprocessing yang baik, hasil dari analisis atau model yang dibangun dapat lebih akurat dan bermakna.

4.8.3 Tahap Text Preprocessing

Tahapan selanjutnya adalah melakukan *Text Preprocessing*, *Text Preprocessing* adalah serangkaian langkah yang diterapkan pada data teks

mentah untuk menyiapkannya agar dapat digunakan dalam analisis lebih lanjut atau pemodelan, terutama dalam konteks *Natural Language Processing (NLP)*. Proses ini penting karena data teks mentah sering kali berantakan dan tidak terstruktur, yang dapat mengurangi efektivitas model *Natural Language Processing* jika tidak diproses dengan benar. Selanjutnya data analisis tersebut melewati beberapa tahapan berikut:

1. *Stopword*: Kata-kata umum yang sering muncul dalam teks tetapi biasanya tidak menambah banyak makna (seperti "dan", "adalah", "untuk"). Menghapus *stopwords* dapat membantu mengurangi kebisingan dalam data serta bisa mengurangi kata yang bermakna negatif dengan menggunakan *Library* Sastrawi.
2. Tokenisasi: Memecah teks menjadi unit-unit yang lebih kecil, seperti kata-kata atau kalimat. Ini adalah langkah dasar untuk analisis teks karena memungkinkan pemrosesan pada tingkat kata atau kalimat.
3. *Stemming*: Mengurangi kata ke bentuk dasar dengan menghilangkan akhiran. Kalimat yang sudah dipecah menjadi kata-per-kata dan dilakukan proses kembali, Langkah ini dilakukan agar penggabungan kata agar menjadi sebuah kalimat.
4. Standarisasi: Standarisasi dilakukan agar data analisis tersebut bisa untuk melakukan standarisasi dan normalisasi dari dataset.

4.8.4 Menampilkan Visualisasi

Exploratory Data Anlaysis digunakan untuk menampilkan kata-kata visualisasi, visualisasi kata dibagi menjadi dua berdasarkan ulasan negatif dan positif. Fungsi dari *Exploratory Data Anlaysis* untuk membangun pemahaman mendalam tentang distribusi, pola, dan hubungan dalam dataset. Ini mencakup mengidentifikasi sebuah anomali, memahami variabel yang paling penting, dan mengidentifikasi distribusi variabel. Berikut adalah merupakan pemrograman untuk pembuatan visualisasi kata beserta hasilnya.

```

▶ # Concatenating all negative clean text from the 'clean_content' column
all_negative_text = ''.join(word for word in negative['clean_content'])

# Creating a WordCloud object
wordcloud = WordCloud(width=1000, height=500, max_font_size=150, random_state=42).generate(all_negative_text)

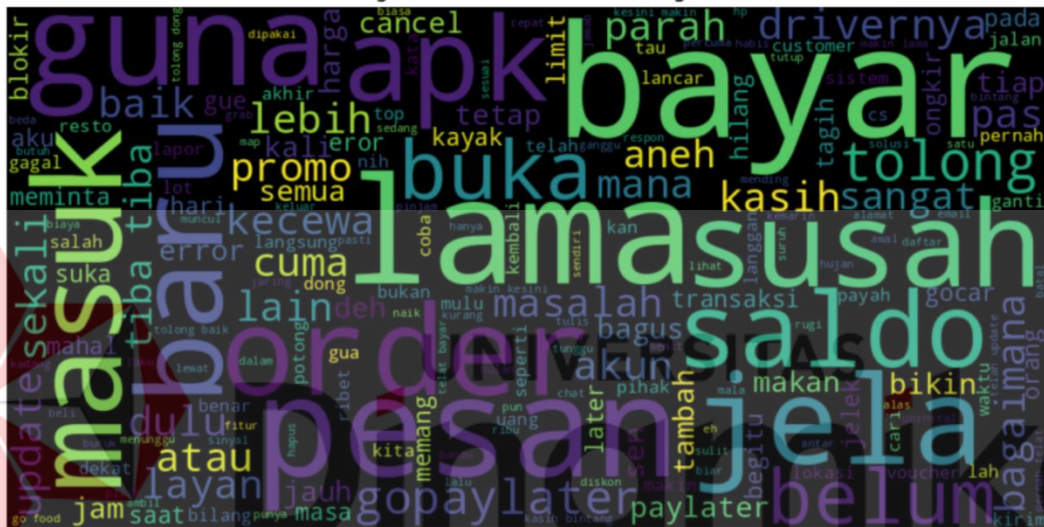
# Displaying the WordCloud
plt.figure(figsize=(10,6))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.title('Visualization of Negative Words')

plt.show()

```

Gambar 4.4 Program untuk memunculkan visualisasi kata negatif

Visualization of Negative Words (After Filtering Irrelevant Words)



Gambar 4.5 Visualisasi kata-kata negatif

```

▶ # Concatenating all positive clean text from the 'clean_content' column
all_positive_text = ''.join(word for word in positive['clean_content'])

# Creating a WordCloud object
wordcloud = WordCloud(width=1000, height=500, max_font_size=150, random_state=42).generate(all_positive_text)

# Displaying the WordCloud
plt.figure(figsize=(10,6))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.title('Visualization of Positive Words')

plt.show()

```

Gambar 4.6 Program untuk menampilkan visualisasi data positif

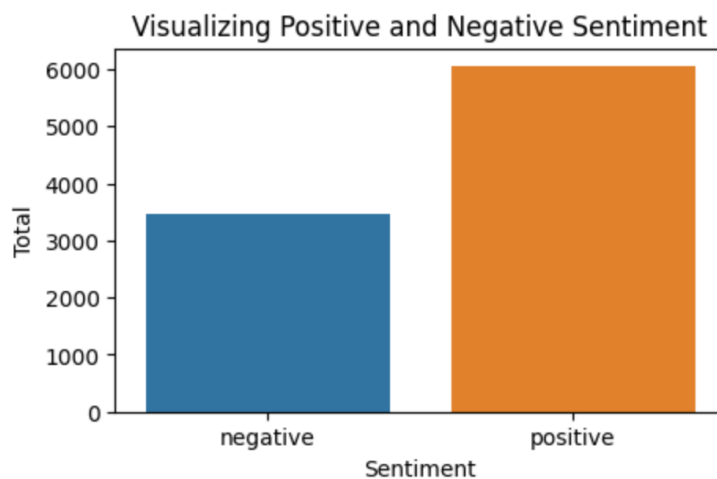
Visualization of Positive Words (After Filtering Irrelevant Words)



Gambar 4.7 Visualisasi kata-kata positif

```
[ ] # Visualizing Positive and Negative Sentiment  
  
plt.figure(figsize=(5,3))  
sns.countplot(data=data_clean, x='target', hue='target')  
plt.title('Visualizing Positive and Negative Sentiment')  
plt.xlabel('Sentiment')  
plt.ylabel('Total')  
  
plt.show()
```

Gambar 4.8 Program untuk menampilkan visualisasi data positif dan negatif



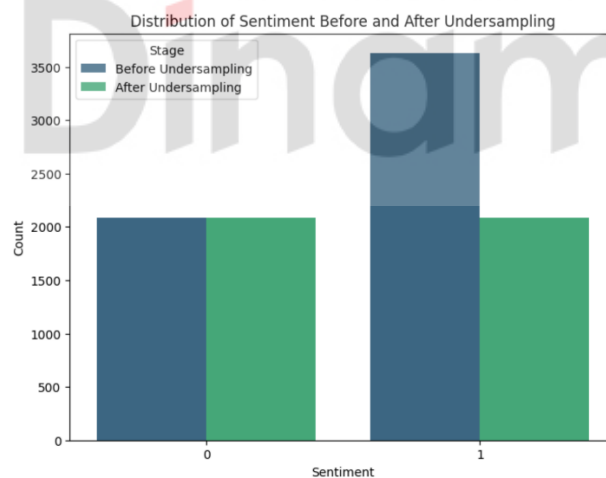
Gambar 4.9 Hasil dalam bentuk diagram batang visualisasi kata negatif dan

Positif

4.8.5 Tahap Undersampling

Undersampling adalah salah satu teknik yang digunakan untuk menangani masalah ketidakseimbangan data (imbalanced data) dalam set data, terutama dalam konteks klasifikasi. Ketidakseimbangan data terjadi ketika jumlah sampel dalam satu kelas jauh lebih banyak dibandingkan dengan kelas lainnya. Hal ini dapat menyebabkan model machine learning bias terhadap kelas yang dominan dan mengabaikan kelas minoritas.

Pada Konsep Undersampling melibatkan pengurangan jumlah sampel dari kelas yang dominan atau mayoritas untuk mencapai keseimbangan antara kelas-kelas yang ada. Dengan mengurangi sampel dari kelas mayoritas, proporsi antara kelas mayoritas dan minoritas menjadi lebih seimbang, sehingga model memiliki kesempatan lebih besar untuk mempelajari pola. Berikut adalah hasil diagram batang dengan menggunakan undersampling hijau merupakan data yang sudah melewati tahap undersampling sedangkan biru adalah sebelum melalui tahap undersampling



Gambar 4.10 Hasil diagram dengan menggunakan Undersampling

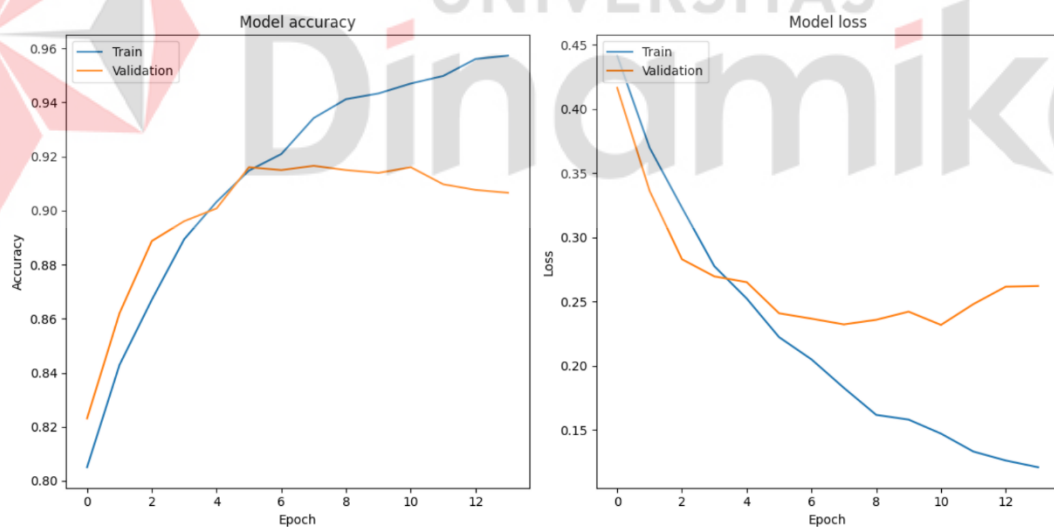
4.8.6 Tahap Long-Short Term Memory

Long Short-Term Memory (LSTM) adalah jenis arsitektur *Recurrent Neural Network (RNN)* yang dirancang untuk mengatasi masalah *vanishing gradient* dan *long-term dependencies* dalam pelatihan model *neural network*. LSTM sangat efektif dalam memproses dan membuat prediksi dari data sekuensial, seperti teks, audio, dan sinyal waktu lainnya. Kemampuan LSTM dalam

mempertahankan informasi penting selama urutan panjang dan mengatasi masalah gradient (Aggarwal, Charu C, 2021). Keunggulan dari mekanisme ini adalah kemampuan LSTM untuk menyimpan informasi yang relevan selama periode waktu yang panjang, dan menghapus informasi yang tidak relevan, yang sangat berguna dalam menangani masalah dengan dependensi jangka panjang. LSTM sangat efektif dalam berbagai aplikasi yang melibatkan data sekuensial, termasuk:

1. *Pemrosesan Bahasa Alami (NLP)*: Seperti prediksi teks, penerjemahan mesin, dan analisis sentimen.
2. *Pemrosesan Audio dan Suara*: Seperti pengenalan suara, pengenalan emosi, dan analisis musik.
3. *Analisis Sinyal Waktu*: Seperti prediksi harga saham, deteksi anomali, dan pengenalan pola.
4. *Generasi Data*: Seperti generasi teks, musik, dan gambar yang mengikuti pola tertentu.

Berikut adalah hasil dengan menggunakan *Long-Short Term Memory*:



Gambar 4.11 Menampilkan kurva akurasi dan *loss*

```

# Generate probabilities for the positive class
y_pred_prob = model.predict(X_test_pad)

# Compute ROC curve
fpr, tpr, _ = roc_curve(y_test_ohc[:, 1], y_pred_prob[:, 1])

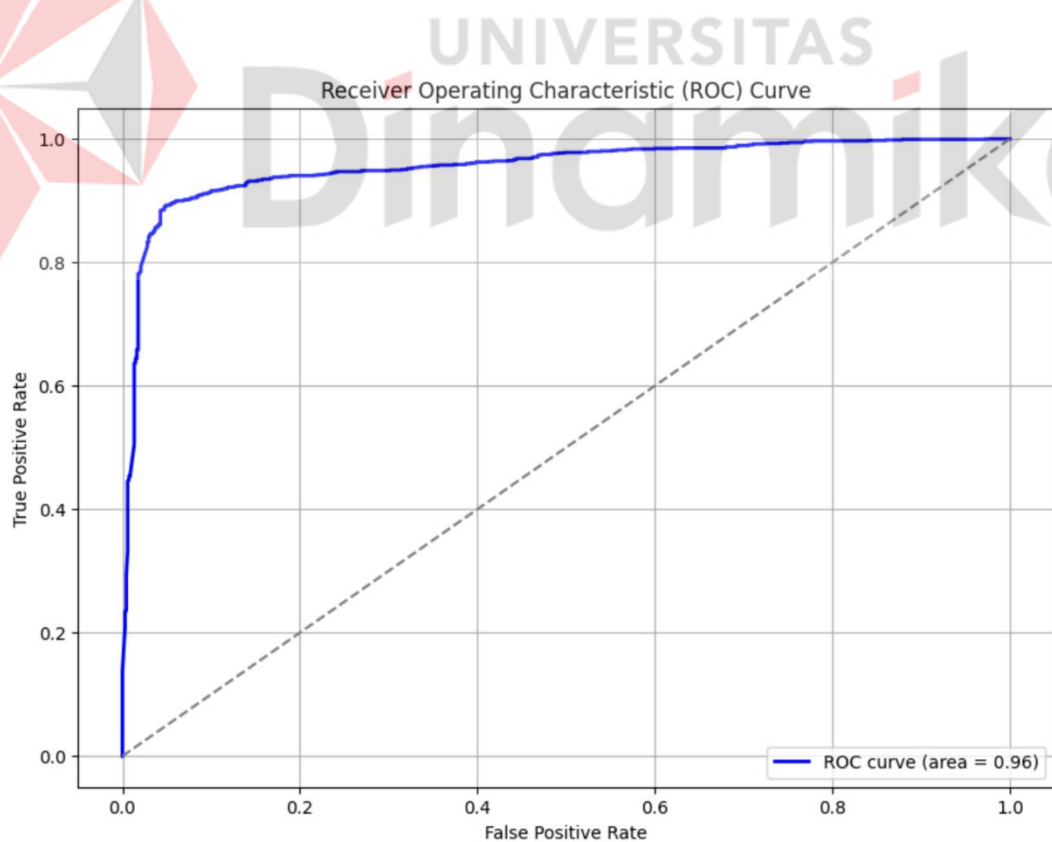
# Compute AUC score
roc_auc = auc(fpr, tpr)

# Plot ROC curve
plt.figure(figsize=(10, 7))
plt.plot(fpr, tpr, color='blue', lw=2, label=f'ROC curve (area = {roc_auc:.2f})')
plt.plot([0, 1], [0, 1], color='gray', linestyle='--')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('Receiver Operating Characteristic (ROC) Curve')
plt.legend(loc='lower right')
plt.grid(True)
plt.show()

# Print AUC score
print(f"AUC Score: {roc_auc:.2f}")

```

Gambar 4.12 Program untuk menampilkan ROC



Gambar 4.13 Hasil analisis dengan ROC

```
[ ] early_stopping = EarlyStopping(monitor='val_loss', patience=3, restore_best_weights=True)

[ ] # Define the LSTM model
model = Sequential()
model.add(Embedding(input_dim=vocab_size,
                    output_dim=embedding_dim,
                    weights=[embedding_matrix],
                    input_length=max_len,
                    trainable=True)) # Set trainable to True to update embeddings during training
model.add(SpatialDropout1D(0.2))
model.add(LSTM(100, dropout=0.2, recurrent_dropout=0.2, return_sequences=True))
model.add(LSTM(100, dropout=0.2, recurrent_dropout=0.2))
model.add(Dense(2, activation='softmax'))

[ ] # Compile the model
model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])

[ ] # Define the number of epochs
epochs = 20

# Fit the model
history = model.fit(
    X_train_pad,
    y_train_oh,
    epochs=epochs,
    batch_size=32,
    validation_data=(X_val_pad, y_val_oh),
    callbacks=[early_stopping]
)
```

Gambar 4.14 Program Training data untuk menampilkan LSTM

4.9 Testing

Testing analisis sentiment menggunakan *Streamlit*, *streamlit* merupakan *framework open-source* untuk membangun aplikasi web interaktif dengan mudah, terutama ditujukan untuk para data scientist dan *machine learning engineers*. *Streamlit* memungkinkan pengguna untuk membuat *User Interface* yang intuitif dan interaktif langsung dari script *Python* tanpa memerlukan pengetahuan mendalam tentang web development. Percobaan proyek ini menggunakan *Streamlit* dikarenakan:

1. Pengembangan cepat dan mudah: *Streamlit* dirancang untuk membuat pengembangan aplikasi web menjadi cepat dan intuitif. Dengan menggunakan sintaks *Python* yang sederhana, pengguna dapat dengan cepat membuat prototipe dan memvisualisasikan data atau model machine learning.
2. Memiliki integrasi *Python* yang kuat: Karena *Streamlit* berfokus pada *Python*, ini memungkinkan integrasi yang mulus dengan pustaka *Python* yang populer seperti *NumPy*, *Pandas*, *Matplotlib*, *Seaborn*, dan lain-lain. Hal ini memudahkan dalam memvisualisasikan data dan menyajikan hasil analisis.

3. Interaktivitas: Streamlit menyediakan komponen antarmuka pengguna yang interaktif seperti slider, tombol, dan input teks, yang memungkinkan pengguna aplikasi untuk berinteraksi dengan data atau model secara dinamis. Ini sangat berguna untuk eksplorasi data, analisis skenario, dan demonstrasi model.
4. Visualisasi data: Streamlit mendukung berbagai alat visualisasi, termasuk grafik statis dan interaktif, peta, dan grafik khusus. Ini memudahkan untuk menyajikan data dalam format yang mudah dipahami dan menarik secara visual.
5. Dukungan untuk *Machine Learning*: *Streamlit* sangat cocok untuk memvisualisasikan model machine learning, baik untuk pelatihan, validasi, maupun penyajian hasil. Ini memudahkan dalam membuat dashboard atau alat untuk menganalisis performa model, *tune hyperparameter*, atau mempresentasikan hasil kepada stakeholder.
6. Penyebaran mudah: Aplikasi Streamlit dapat dengan mudah dijalankan secara lokal atau di-deploy ke server untuk diakses oleh pengguna lain. Streamlit juga menyediakan platform untuk deploy aplikasi secara gratis yang disebut *Streamlit Sharing*.

Berikut adalah gambar hasil testing yang digunakan menggunakan streamlit, terdapat dua hasil yaitu analisis ulasan negatif dan ulasan positif.

```
[ ] !pip install streamlit

[ ] !wget -q -O - ipv4.icanhazip.com

▶ with open('app.py', 'w') as file:
  file.write('import streamlit as st\n'
            'from tensorflow.keras.models import load_model\n'
            'from tensorflow.keras.preprocessing.text import Tokenizer\n'
            'from tensorflow.keras.preprocessing.sequence import pad_sequences\n\n'
            '# Load the model\n'
            'model = load_model(\'model_lstm_undersampled.h5\')\n'
            'tokenizer = Tokenizer()\n\n'
            'st.title("Sentiment Analysis")\n\n'
            'st.write("Enter text for sentiment analysis:")\n\n'
            '# Input text for sentiment analysis\n'
            'text = st.text_input("Input Text:")\n\n'
            'if text:\n')
```

Gambar 4.15 Input Program Streamlit

Sentiment Analysis

Enter text for sentiment analysis:

Input Text:

bagus banget kinerjanya, sangat membantu

Sentiment Prediction:

Text: 'bagus banget kinerjanya, sangat membantu'

Sentiment: positive

Gambar 4.16 Hasil analisis ulasan positif

Sentiment Analysis

Enter text for sentiment analysis:

Input Text:

kinerjanya buruk, tidak bagus

Sentiment Prediction:

Text: 'kinerjanya buruk, tidak bagus'

Sentiment: negative

Gambar 4.17 Hasil sentimen analisis ulasan negatif

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan kegiatan Program Magang dan Studi Independen Bersertifikat di PT Hactivate Teknologi Indonesia, maka Kesimpulan yang di dapat terkait proyek tersebut adalah, teknik *Natural Language Processing* yang digunakan mampu mengkategorikan sentimen ulasan menjadi positif, negatif, atau netral dengan tingkat akurasi yang bisa mencapai 85%. Akurasi ini diukur melalui perbandingan dengan ulasan yang telah diklasifikasikan secara manual sebagai referensi. Hasil menunjukkan bahwa model *Natural Language Processing* dapat menangkap sentimen umum yang diekspresikan pengguna dengan cukup tepat.

5.2 Saran

Adapun saran yang diberikan untuk *Capstone Project* kali ini adalah sebagai berikut:

1. Melakukan analisis sentimen dengan menggunakan *Natural Language Processing* untuk membantu mencari ulasan negatif, positif dan netral
2. Hasil testing yang belum optimal karena menggunakan *Machine Learning* dengan metode *Clasification*

DAFTAR PUSTAKA

1. Diambil Dari Buku:

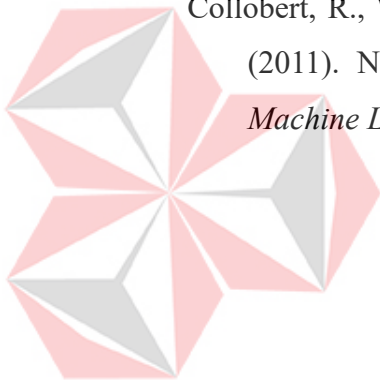
Jurafsky, D., & Martin, J. H. (2022). *Speech and language processing* (3rd ed., pp. 123-456). *Pearson*.

Patterson, David. *Introduction to Deep Learning and Neural Networks with Python*. *Apress*, 2021. pp. 65-68.

2. Diambil Dari Jurnal:

Ramadhanu, A., Zaky, M. R., Isra, M., Nengsih, N. S. W., Sularno, S., & A, M. R. (2023). Penerapan Machine Learning Untuk Menentukan Tingkat Kepuasan Tamu Hotel Dymens Menggunakan Metode Vader. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 5(3), 337-343.

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural Language Processing (Almost) from Scratch. *Journal of Machine Learning Research*, 12, 2493-2537.



UNIVERSITAS
Dinamika