

LAPORAN KERJA PRAKTIK

**ANALISIS SENTIMEN ULASAN APLIKASI TASPEN
OTENTIKASI MENGGUNAKAN METODE KLASIFIKASI
ALGORITMA NAÏVE BAYES**



Disusun oleh:

MARIO DESENDI MANURUNG
120450082

PROGRAM STUDI SAINS DATA

JURUSAN SAINS
INSTITUT TEKNOLOGI SUMATERA

2023

LAMAR PENGESAHAN

Proposal Kerja Praktik

Analisis Sentimen Ulasan Aplikasi Taspen Otentifikasi Menggunakan Metode Klasifikasi Algoritma Naïve Bayes

Oleh

Mario Desendi Manurung

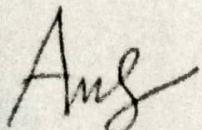
120450082

Lampung Selatan, 13 November 2023

Mengetahui,

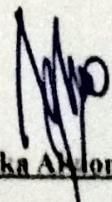
Pembimbing

Kordinator Kerja Praktik



Ardika Satrin, M.Si

NRK. 1997111020231425

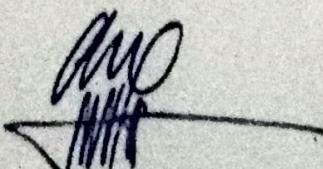


Mika Alphonita S, M.Si

NRK. 1993050920212258

Menyetujui,

Koordinator Program Studi Sains Data ITERA



Tirta Setiawan, S.Pd., M.Si

NIP. 199008222022031003

ABSTRAK

Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi "Taspen Otentikasi" dengan menggunakan metode Naïve Bayes. Alasan dipilihnya Multinomial Naïve Bayes karena algoritma ini merupakan salah satu algoritma machine learning yang paling populer digunakan untuk memproses klasifikasi teks, prosesnya sederhana, efisien, dan berkinerja baik. Data ulasan pengguna diambil dari Google Play Store menggunakan web scraping. Hasil penelitian menunjukkan bahwa model Multinomial Naive Bayes memiliki akurasi sebesar 91.49%. Hasil analisis menunjukkan mayoritas sentimen negatif dari pengguna, dengan fokus pada kemudahan penggunaan. Temuan ini memberikan wawasan penting bagi PT. Taspen untuk meningkatkan aplikasi mereka guna memenuhi harapan pengguna.

Kata Kunci: Analisis Sentimen, Taspen Otentikasi, Naïve Bayes, Web Scrapping, Pengalaman Pengguna.

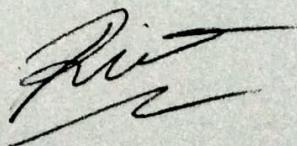
KATA PENGANTAR

Puji dan syukur Penulis ucapkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan laporan kerja praktik ini sebagai syarat lulus mata kuliah Kerja Praktik Program Studi Sains Data Institut Teknologi Sumatera. Dalam proses pembuatan laporan yang telah penulis kerjakan, tidak terlepas dari bantuan, dukungan, kritik dan saran membangun yang telah diberikan oleh berbagai pihak sehingga laporan ini dapat dikerjakan dengan baik, maka dari itu penulis ingin menyampaikan rasa terima kasih kepada :

1. PT. TASPEN (PERSERO) KANTOR CABANG BEKASI yang telah memberikan kesempatan kepada penulis untuk belajar melalui kerja praktik di instansi tersebut.
2. Bapak Ikhhat Permana, selaku Pembimbing lapangan yang telah memberikan arahan dan bimbingan dalam melaksanakan kerja praktik.
3. Bapak Ardika Satria M,Si. selaku dosen pembimbing yang telah memberikan banyak pengetahuan dan telah meluangkan waktu untuk berdiskusi dengan Penulis dalam rangka penyusunan laporan akhir kerja praktik.
4. Keluarga yang telah memberikan doa dan dukungan kepada Penulis selama proses kerja praktik berlangsung.
5. Seluruh teman-teman Kontrakan Jahanam yang telah memberikan dukungan, semangat dan membantu secara langsung maupun tidak langsung dalam penyusunan Laporan Kerja Praktik.

Penulis menyadari bahwa Laporan kerja praktik ini masih jauh dari kata sempurna, sehingga Penulis mengharapkan evaluasi, kritik dan saran perbaikan dari berbagai pihak. Akhir kata Penulis mengucapkan terima kasih.

Bandar Lampung, 13 November 2023



Mario Desendi Manurung

DAFTAR ISI

LEMBAR PENGESAHAN	i
ABSTRAK	ii
KATA PENGANTAR.....	iii
DAFTAR ISI.....	iv
DAFTAR GAMBAR	vi
DAFTAR LAMPIRAN	vii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
BAB II TINJAUAN PUSTAKA	4
2.1 PT. Taspen (Persero) Kantor Cabang Bekasi.....	4
2.2 Analisis Sentimen.....	5
2.3 Aplikasi Taspen Otentifikasi.....	6
2.4 Klasifikasi Naïve Bayes	7
2.5 Pembobotan TF-IDF.....	9
BAB III METODOLOGI	10
3.1 Waktu dan Tempat Pelaksanaan Kerja Praktik	10
3.2 Diagram Alir Penelitian.....	10
3.3 Metode Pengolahan Data.....	11
3.3.1 Pengambilan Data	11
3.3.2 Pemberian Label.....	12
3.3.3 Text Preprocessing	12
3.3.4 Splitting Data	14
3.3.5 Pembobotan TF-IDF	14
3.3.6 Pelatihan Model	15
3.3.7 Evaluasi Model.....	15
BAB IV HASIL DAN PEMBAHASAN.....	16
4.1 Data Ulasan	16
4.2 Hasil Text Preprocessing	17
4.3 Hasil Evaluasi Model	18

BAB V PENUTUP.....	22
5.1 Kesimpulan.....	22
DAFTAR PUSTAKA.....	23
LAMPIRAN.....	26

DAFTAR GAMBAR

Gambar 1. Diagram Alir.....	10
Gambar 2. Data Awal	16
Gambar 3. Diagram Batang Jumlah Ulasan berdasarkan Rating.....	16
Gambar 4. Data setelah dilakukan filtering kolom	17
Gambar 5. Pemberian Label.....	17
Gambar 6. Hasil Text Preprocessing	18
Gambar 7. Hasil Evaluasi Model	19
Gambar 8. Confusion Matrix	20
Gambar 9. Pie Chart Persentase Sentimen.....	21

DAFTAR LAMPIRAN

Lampiran 1. Source Code.....	26
Lampiran 2. Kegiatan Saat Kerja Praktik	26
Lampiran 3. Sertifikat Magang	26

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pemerintah sebagai salah satu elemen penting bagi sebuah negara disamping berperan sebagai pengembang suatu amanat pemerintah juga dituntut agar dapat memberikan pelayanan terbaik untuk memenuhi kebutuhan Masyarakat diberbagai segala aspek seperti yang telah di cantumkan dalam udang-undang sebuah negara. pelayanan masyarakat merupakan peran yang paling mendasar dari keberadaan pemerintah indonesia. Pelayanan yang diberikan pemerintah bertujuan untuk bisa memberikan keadilan dan kesejahteraan di tengah-tengah warga negara melalui penyediaan berbagai pelayanan kehidupan mayarakat dalam bidang sosial, ekonomi dan kebudayaan yang di dukung dan disediakan oleh negara [1].

Pemerintah membangun Instansi Badan Usaha Milik Negara (BUMN) yang dipercaya oleh pemerintah sebagai suatu penyelengara asuransi sosial dalam sistem perekonomian yang berfokus terhadap kesejahteraan Masyarakat [2]. Menurut Undang- Undang Nomor 19 Tahun 2003 Badan Usaha Milik Negara (BUMN) merupakan suatu bidang usaha dimana modalnya dimiliki oleh pemerintah yang berasal dari kekayaan Negara. Pemerintah memberikan salah satu diantaranya mengelola program jaminan sosial bagi Aparatur Sipil Negara (ASN) dan pejabat negara dengan membentuk ketetapan pada Aparatur Sipil Negara “Dana Pensiun” yang terdapat dalam Undang- Undang Nomor 11 Tahun 1996 yang dikelola oleh PT. Taspen sebagai bentuk Perlindungan pemeliharaan serta peningkatan kesejahteraan terhadap Aparatur Sipil Negara Pegawai Negera Sipil (PNS) [3].

PT. Taspen (Persero) merupakan pelayanan pembayaran dana pensiun Pegawai Negeri sipil (PNS). PT. Taspen juga termasuk kedalam salah satu perusahaan yang di kelola oleh Badan Usaha Milik Negara untuk melayani dan memudahkan masyarakat dan aparatur sipil negara seperti Pegawai Negeri Sipil (PNS) dalam pengambilan dana pensiun. PT. Taspen mulai sejak Januari 2019 di Seluruh Indonesia termasuk kantor Cabang Kota Bekasi menerapkan kebijakan sistem otentifikasi untuk para pensiunan dengan melakukan biometric berupa wajah, suara

dan sidik jari dengan meluncurkan program Aplikasi Taspen dengan motto lima 5T yaitu: tepat orang, tepat waktu, tepat jumlah, tepat tempat, dan tepat administrasi [4].

Aplikasi Taspen menjembatani para penerima pensiun agar menjadi lebih mudah dalam melakukan pengambilan dana pensiun dengan verifikasi wajah pensiun. Melalui data penerima pensiun yang sudah ada, para pensiun dapat melakukan otentikasi data dimanapun dan kapanpun melalui aplikasi secara otentikasi menggunakan smartphone tanpa harus mengunjungi bank, pos maupun mitra bayar. Data dari enrollment tersebut juga akan disematkan dalam bentuk smartcard taspen yang nantinya dapat digunakan sebagai ATM untuk pengambilan dana pensiun [5].

Namun setelah dikeluarkan Aplikasi Taspen otentikasi ini ternyata bukan semakin memudahkan namun malah membuat peserta pensiun merasa kesulitan yang diakibatkan saat verifikasi enrollment atau data biometric yang sering kali terjadi kegagalan atau sistem aplikasi yang eror sehingga para peserta pensiun tetap harus datang langsung secara manual ke mitra bayar untuk verifikasi data [6].

Oleh karena itu PT. Taspen harus meningkatkan kualitas aplikasi pelayanannya berdasarkan ulasan dari aplikasi “Taspen Otentikasi”. Untuk memudahkan pengelolaan data ulasan, penelitian ini menggunakan analisis sentimen. Analisis sentimen adalah proses untuk mengevaluasi sentimen atau pendapat yang dinyatakan oleh pengguna dalam ulasan yang mereka tulis tentang aplikasi tertentu. Analisis sentimen aplikasi sering digunakan oleh pengembang dan perusahaan untuk memperbaiki produk mereka, memahami kebutuhan pelanggan, dan meningkatkan kepuasan pelanggan. Analisis sentimen dilakukan dengan mengelompokkan respon menjadi kelas positif dan negatif, metode klasifikasi yang digunakan adalah Multinomial Naïve Bayes (MNB) [7].

Metode Naive Bayes adalah salah satu metode klasifikasi teks yang banyak digunakan dalam analisis sentimen. Metode ini berdasarkan teori probabilitas Bayes dan diasumsikan bahwa setiap fitur dalam teks tidak berkorelasi satu sama lain. Metode Naive Bayes telah terbukti efektif dalam banyak kasus analisis

sentimen pada teks, termasuk ulasan aplikasi. Sentimen analisis adalah sebuah pemrosesan bahasa dengan menggunakan sebuah pendekatan untuk mendefinisikan bahasa tersebut mengarah ke arah positif atau negatif. Sentimen analisis merupakan aspek yang sangat popular dan memberikan keuntungan yaitu prediksi penjualan dan pengambilan keputusan kepada para investor [8]. Penerapan analisis sentimen pada ulasan aplikasi TASPEN Otentikasi menggunakan metode klasifikasi algoritma Naïve Bayes adalah langkah yang penting dalam upaya untuk memahami dan meningkatkan penggunaan aplikasi ini.

1.2 Rumusan Masalah

Adapun rumusan masalah yang menjadi fokus penelitian adalah :

1. Bagaimana metode klasifikasi algoritma Naïve Bayes dapat diterapkan untuk melakukan analisis sentimen pada ulasan aplikasi TASPEN Otentikasi?
2. Bagaimana tingkat kepuasan pengguna aplikasi TASPEN Otentikasi dapat diukur berdasarkan analisis sentimen ulasan pengguna?

1.3 Tujuan

Tujuan utama dari penelitian ini adalah :

1. Mengembangkan sistem analisis sentimen yang menggunakan metode klasifikasi algoritma naïve bayes untuk mengklasifikasikan ulasan aplikasi Taspen Otentikasi menjadi kategori sentiment positif atau negatif.
2. Mengukur kepuasan pengguna aplikasi Taspen Otentikasi berdasarkan analisis sentimen ulasan pengguna.

BAB II

TINJAUAN PUSTAKA

2.1 PT. Taspen (Persero) Kantor Cabang Bekasi

PT. Taspen (Persero) Kantor Cabang Bekasi adalah badan usaha milik negara Indonesia yang bergerak dibidang asuransi tabungan dari hari tua dan dana pensiun bagi aparatur sipil negara dan pejabat negara. Berawal dari Konferensi Kesejahteraan Pegawai Negeri yang diselenggarakan pada tanggal 25-26 Juli 1960 di Jakaa yang menghasilkan Keputusan Menteri Pertama RI Nomor 388/MP/1960 tanggal 25 Agustus 1960. Dalam Keputusan tersebut, Pemerintah menetapkan pentingnya pembentukan jaminan sosial sebagai bekal bagi Pegawai Negeri dan keluarganya di masa purna bakti. Kemudian pada tanggal 17 April 1963, Pemerintah mendirikan Perusahaan Negara Dana Tabungan dan Asuransi Pegawai Negeri (PN TASPEN) melalui Peraturan Pemerintah Nomor 15 Tahun 1963. Pembentukan Program Tabungan Hari Tua Pegawai Negeri ditetapkan dalam Peraturan Pemerintah No 9 tahun 1963 tentang Pembelanjaan Pegawai Negeri dan Peraturan Pemerintah Nomor 10 tahun 1963 tentang Tabungan Asuransi dan kepegawaian Indonesia.

Seiring dengan adanya peningkatan jumlah Pegawai Negeri dan semakin luasnya cakupan layanan, pada tanggal 18 November 1970 melalui Keputusan Menteri Keuangan RI Nomor Kep.749/MK/IV/11/1970 PN TASPEN bertransformasi menjadi Perusahaan Umum. Peningkatan status dari Perusahaan Umum menjadi Perseroan dilakukan berdasarkan Peraturan Pemerintah RI Nomor 26 Tahun 1981 dan disahkan dengan Akta Notaris Imas Fatimah Nomor: 4 tanggal 4 Januari 1982 dengan nama PT TASPEN (PERSERO) yang menyelenggarakan Program Tabungan Hari Tua dan Program Pensiun.

Pada tahun 2014, dalam rangka memfokuskan diri sebagai perusahaan yang melayani jaminan sosial bagi Aparatur Sipil Negara berdasarkan Pasal 92 ayat (4) dan Pasal 107 Undang- undang nomor 5 Tahun 2014 tentang Aparatur Sipil Negara, Pemerintah telah menerbitkan Peraturan Pemerintah Nomor 70 Tahun 2015 tanggal 16 September 2015 tentang Jaminan Kecelakaan Kerja dan Jaminan Kematian Bagi Pegawai Aparatur Sipil Negara yang kemudian telah diubah dengan Peraturan Pemerintah Nomor 66 Tahun 2017 tanggal 29 Desember 2017

tentang Perubahan Atas Peraturan Pemerintah Nomor 70 Tahun 2015 tentang Jaminan Kecelakaan Kerja dan Jaminan Kematian bagi Pegawai Aparatur Sipil Negara. Oleh karena itu, PT TASPEN (PERSERO) dipercaya untuk mengelola Program Asuransi Sosial yaitu Program Jaminan Kecelakaan Kerja (JKK) dan Jaminan Kematian (JKM) terhitung mulai 1 Juli 2015.

Dengan pengalaman yang sudah terbukti dalam memberikan pelayanan program asuransi sosial bagi ASN dan Pejabat Negara dan lebih memberikan layanan yang terbaik bagi Peserta, TASPEN melakukan pembenahan dan mendekatkan diri kepada Peserta melalui 57 Kantor Cabang yang tersebar di Seluruh Indonesia terdiri dari 6 Kantor Cabang Utama, 7 Kantor Cabang tipe A, 14 Kantor Cabang tipe B, 19 Kantor Cabang tipe C dan 11 Kantor Cabang tipe D. Sebagai salah satu instansi pelayanan publik, TASPEN berkomitmen untuk senantiasa meningkatkan layanan bagi Aparatur Sipil Negara dan Pejabat Negara melalui berbagai macam inovasi, mulai dari digital-based service, layanan klaim otomatis, layanan kunjungan nasabah hingga layanan klaim satu jam yang telah memperoleh seiksi ISO 9001: 2015. Hal ini merupakan komitmen Perseroan untuk terus meningkatkan kualitas layanan guna tercapainya kesejahteraan Aparatur Sipil Negara yang berkelanjutan [9].

PT Tabungan dan Asuransi Pegawai Negeri (Persero) kantor cabang Bekasi merupakan salah satu kantor cabang PT.TASPEN yang terdapat di Indonesia dengan jumlah karyawan 17 dengan 6 penjaga keamanan beralamat di Jalan Kemakmuran No.39, RT. 004/005, Marga Jaya, Kecamatan Bekasi Selatan, Kota Bekasi, Provinsi Jawa Barat 1714. PT Dana Tabungan dan Asuransi Pegawai Negeri (Persero) Kantor Cabang Bekasi memiliki 5 bagian yaitu: Kepesertaan, Layanan dan Manfaat, Kas dan Verifikasi Surat Pertanggungjawaban, Administrasi Keuangan, serta Umum dan Sumber Daya Manusia.

2.2 Analisis Sentimen

Analisis sentimen adalah bidang studi text mining yang menganalisa opini, sentimen, evaluasi, penilaian, sikap dan emosi orang terhadap entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atributnya

[10]. Analisis sentimen atau opinion mining merupakan proses memahami, mengestrak, dan mengolah data textual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini [11]. Analisis sentimen bertujuan untuk mengetahui kecenderungan opini atau pandangan seseorang terhadap sebuah masalah atau objek. Analisis sentimen merupakan kombinasi antara text mining dan natural language processing.

Analisis sentiment terbagi ke dalam dua kelompok jika didasarkan pada sumber datanya [12] yaitu, coarse-grained sentiment analysis dan fined grained sentiment analysis. Coarse-grained sentiment analysis adalah analisis sentimen di level dokumen, dimana analisis ini mengklasifikasikan keseluruhan isi dokumen sebagai sentimen positif dan sentimen negatif. Sedangkan fined grained sentiment analysis adalah analisis sentimen di level kalimat. Kelas sentimen dapat dibagi menjadi kelas sangat positif, positif, netral, negatif, dan sangat negatif [13].

2.3 Aplikasi Taspen Otentikasi

Otentikasi adalah sebuah usaha pengecekan identitas seseorang pengguna sistem komunikasi pada proses login ke dalam sebuah system. Proses ini diawali dengan pengiriman kode seperti, password, username, pin, sidik jari oleh pengguna kepada server. Setelah itu sistem akan menerima kode tersebut, dan kemudian akan dibandingkan dengan kode data yang telah disimpan dalam database server yang sudah disimpan sebelumnya. Jika hasil tersebut sama dan dapat dikonfirmasi, maka server akan menerima hak akses dan pengguna dapat menggunakan dan dapat mengakses secara penuh. Sedangkan jika hasilnya gagal atau tidak sama, maka server akan mengirimkan pesan kegagalan dan menolak hak akses pengguna.

Otentikasi Taspen merupakan proses verifikasi yang dilakukan untuk memastikan bahwa dana pensiun bulanan diterima kepada pihak yang berhak mendapatkan tunjangan pensiunan dari Negara melalui PT. Taspen. Dari perkembangan teknologi digital yang sangat pesat saat ini PT. Taspen memanfaatkan teknologi untuk mengamankan dan mempermudah dalam melayani para nasabahnya untuk pengambilan dan pendistribusian gaji pensiunan. Untuk itu, PT. Taspen menggunakan aplikasi Taspen Otentikasi. Aplikasi Taspen

Otentikasi merupakan layanan digital yang dihadirkan oleh PT. Taspen untuk mempermudah nasabahnya melakukan otentikasi atau absen secara berkala. Langkah tersebut dinilai sebagai wujud pengembangan layanan agar tidak ada keterlambatan bagi pensiunan menerima gaji pensiunnya.

Aplikasi Otentikasi memudahkan pensiunan untuk melakukan pembuktian fisik atas pembayaran dan pencairan hak pensiunan [14]. Pasalnya mereka hanya perlu mengarahkan kamera ponsel ke wajah. Autentikasi ini wajib dilakukan agar dana pensiun diterima pensiunan itu sendiri dan tak salah sasaran. Aplikasi Taspen Otentikasi dapat diakses dengan cara mengunduh aplikasi pada telepon pintar masing-masing nasabah, setelah itu nasabah akan diminta untuk memasukkan nomor identitas yang sudah disediakan oleh Taspen. Sebelum melakukan otentikasi menggunakan aplikasi ini nasabah diharapkan sudah melakukan perekaman data diri yang sudah dilakukan oleh mitra bayar Taspen.

2.4 Klasifikasi Naïve Bayes

Teorema Bayes merupakan dasar dari teknik pengklasifikasian yang diterapkan pada klasifikasi Naïve Bayes. Teorema Bayes sendiri ditemukan oleh ilmuwan Inggris bernama Thomas Bayes. Dalam melakukan analisis sentimen banyak yang menerapkan algoritma Naïve Bayes sebagai metodenya. Klasifikasi Naïve Bayes dapat memprediksi probabilitas keanggotaan suatu kelas yang bekerja berdasarkan prediksi asumsi independen. Dengan demikian, peluang masa depan dapat diidentifikasi berdasarkan pengalaman sebelumnya [15]. Algoritma Naïve Bayes memiliki keunggulan, yaitu efisien karena mampu membuat proses analisis sentimen menjadi singkat [16]. Selain itu, algoritma Naïve Bayes memiliki akurasi yang cenderung tinggi walaupun memiliki data latih yang sedikit [17].

Naïve Bayes juga merupakan salah satu metode machine learning yang memanfaatkan perhitungan probabilitas dan statistika [18]. Pada penelitian yang dilakukan menggunakan algoritma multinomial naïve bayes (MNB). Algoritma ini mengimplementasikan algoritma naïve bayes untuk data yang didistribusikan secara multinasional dan merupakan salah satu dari dua varian naïve bayes klasik yang digunakan dalam klasifikasi teks [19]. MNB cocok untuk mengklasifikasikan teks yang memiliki banyak fitur (misalnya kata-kata dalam

dokumen) dengan menggunakan metode probabilitas. Algoritma ini bekerja dengan menghitung probabilitas kemunculan setiap kata di setiap kelas dan kemudian menggunakan probabilitas ini untuk memprediksi kelas baru dari suatu teks. Implementasi Multinomial Naïve Bayes biasanya digunakan pada klasifikasi email sebagai spam atau bukan spam, klasifikasi dokumen sebagai topic tertentu, atau klasifikasi sentimen pada teks (positif, negatif, atau netral). Kode yang digunakan saat proses eksperimen dengan mengimport kelas MultinomialNB dari modul naïve bayes pada library scikit-learn. Dengan menggunakan MultinomialNB dapat melakukan pelatihan dan prediksi dengan mudah pada data teks. Rumus Multinomial Naïve Bayes digunakan untuk menghitung probabilitas dokumen terhadap setiap kategori yang ada. Probabilitas ini digunakan untuk memprediksi kategori mana yang paling mungkin untuk dokumen tersebut. Rumus dasar Multinomial Naïve Bayes adalah sebagai berikut:

$$P\left(\frac{c}{d}\right) = P(c) * P\left(\frac{d}{c}\right)/P(d) \quad (2.1)$$

Keterangan :

$P\left(\frac{c}{d}\right)$ = probabilitas dokumen d masuk ke dalam kategori c.

$P(c)$ = probabilitas prior untuk kategori c.

$P\left(\frac{d}{c}\right)$ = probabilitas kemunculan fitur (kata atau istilah) dalam dokumen d, jika dokumen d masuk dalam kategori c.

$P(d)$ = probabilitas dari dokumen d.

Untuk menghitung probabilitas $P(d|c)$, digunakan rumus sebagai berikut :

$$P((d|c)) = \prod (P(t|c)^{nt}) \quad (2.2)$$

Keterangan :

$P\left(\frac{t}{c}\right)$ = Probabilitas kemunculan fitur t (kata atau istilah) dalam kategori c.

nt = Jumlah kemunculan fitur t dalam dokumen d.

Untuk menghitung probabilitas $P(c)$ dan $P(d)$, digunakan rumus sebagai berikut :

$$P(c) = \frac{N_c}{N} \quad (2.3)$$

Keterangan :

N_c = Jumlah dokumen dalam kategori c.

N = Jumlah total dokumen.

$P(d)$ dapat dihitung dengan cara yang sama seperti menghitung $P(d|c)$, dengan asumsi bahwa dokumen tersebut termasuk dalam semua kategori yang ada. Setelah semua probabilitas diperoleh, dokumen dapat diklasifikasikan ke dalam kategori dengan probabilitas tertinggi. Kategori dengan probabilitas tertinggi dianggap sebagai kategori dokumen yang paling mungkin.

2.5 Pembobotan TF-IDF

Pembobotan kata menggunakan Term Frequency-Inverse Document Frequency (TF-IDF), yaitu pembobotan kata berfungsi untuk mengubah data berupa teks menjadi numerik. Bidang pembelajaran mesin atau *Machine learning* dan *deep learning* akan bekerja dengan baik dan maksimal jika data yang diolah berupa data numerik. Maka setiap penelitian pada bidang ini khususnya analisis sentimen harus mengubah data berupa kumpulan kata-kata atau kalimat menjadi numerik. Proses tersebut biasanya disebut dengan proses pembobotan kata.

Metode TF-IDF memiliki representasi lebih baik dibandingkan dengan metode fitur ekstraksi lainnya untuk beberapa kasus analisis sentimen [20]. Adapun konsep dan teknisnya merujuk ke penelitian sebelumnya [20], [21]. Teknisnya dari TF-IDF ialah dokumen dipecah menjadi satu kata atau unigram. Proses menghitung Term Frequency (TF) dan Inverse Document Frequency (IDF) adalah bagian dari proses pemberian bobot pada kata dengan TF-IDF. Tujuan dari TF-IDF adalah untuk mengidentifikasi kata-kata yang paling penting dalam suatu dokumen atau kumpulan dokumen.

BAB III

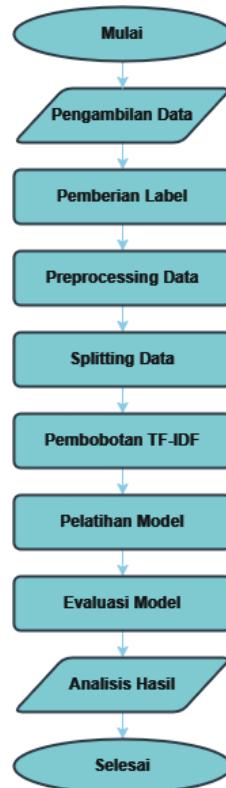
METODOLOGI

3.1 Waktu dan Tempat Pelaksanaan Kerja Praktik

Dalam pelaksanaan Kerja Praktik, saya berkesempatan melakukan kerja praktik di PT. Taspen (Persero) Kantor Cabang Kota Bekasi.

Nama Instansi	: PT. Taspen (Persero) Kantor Cabang Kota Bekasi
Alamat	: Jl. Kemakmuram No.39, RT.004/RW.005, Marga Jaya, Kec. Bekasi Sel, Kota Bekasi 17141
No. Telpon	: (021) 28519766 / 081280127643
Email	: taspenkcbekasi@gmail.com
Web	: www.taspen.co.id
Waktu Pelaksanaan	: 6 Juli 2023 – selesai

3.2 Diagram Alir Penelitian



Gambar 1. Diagram Alir

3.3 Metode Pengolahan Data

3.3.1 Pengambilan Data

Google Play merupakan aplikasi multiplatform yang dapat diakses secara mobile maupun website. Pada penelitian ini, data ulasan aplikasi Taspen Otentikasi diambil melalui website Google Play. Teknik yang digunakan dalam tahap pengambilan data ulasan aplikasi Taspen Otentikasi adalah teknik web scraping. Data yang diambil pada proses web scraping merupakan data yang berasal dari halaman website dengan bahasa markup. Prosesnya sendiri dilakukan secara semi-terstruktur [22]. Cara kerja dari web scraping yaitu dengan mengakses halaman website, memilih informasi tertentu yang ingin didapat, kemudian dilakukan proses ekstraksi data, dan terakhir menyimpan data tersebut menjadi dataset yang terstruktur [23].

Proses web scraping dilakukan dengan menggunakan library google-play-scraping di Google Colab. Pertama, Anda harus menginstal library ini dengan perintah pip. Setelah itu, Anda dapat mengimpor fungsi app dan reviews_all untuk mengambil informasi tentang aplikasi yang ingin Anda analisis dan ulasan pengguna terkait aplikasi tersebut. Dengan mengidentifikasi ID paket aplikasi, Anda dapat mengakses informasi seperti judul, deskripsi, peringkat, jumlah ulasan, dan lainnya. Selanjutnya, Anda dapat mengambil semua ulasan pengguna dengan berbagai parameter opsional seperti bahasa dan negara. Data ulasan tersebut dapat diolah sesuai kebutuhan, disimpan dalam format CSV dan digunakan untuk analisis sentimen.

Dari scrapping data ulasan aplikasi Taspen Otentikasi didapat banyak sekali kolom, yaitu review ID, userName, userImage, content, score, thumbsUpCount, reviewCreatedVersion, at, replyContent, repliedAt, appVersion, dan sebagainya. Kemudian, karena kita hanya membutuhkan kolom content dan score, maka kitalakukan filter kolom hingga menyisakan kolom content dan score.

3.3.2 Pemberian Label

Melakukan pemberian label sentimen pada ulasan yang tersimpan dalam DataFrame `my_df`. Untuk melakukan pelabelan ini, sebuah fungsi bernama `pelabelan(score)` telah dibuat. Fungsi tersebut menerima parameter `score` yang merupakan nilai atau skor yang menggambarkan sentimen dari suatu ulasan. Dalam fungsi pelabelan, terdapat aturan-aturan yang diterapkan untuk menentukan sentimen ulasan:

- Jika nilai `score` kurang dari 3, maka ulasan akan diberi label "Negatif".
- Sedangkan jika nilai `score` adalah 4 atau 5, maka ulasan akan diberi label "Positif".

Kemudian, kode menerapkan fungsi `pelabelan` tersebut pada setiap baris data dalam kolom 'score' dari DataFrame `my_df`. Hasil dari penerapan ini disimpan dalam kolom baru yang dinamakan 'Label'. Dengan demikian, setiap ulasan dalam DataFrame `my_df` akan mendapatkan label sentimen "Positif" jika skornya adalah 4 atau 5, dan label "Negatif" jika skornya kurang dari 3.

Hasil akhir dari proses ini adalah DataFrame `my_df` yang telah diperkaya dengan kolom 'Label', yang menggambarkan sentimen dari setiap ulasan berdasarkan skor yang diberikan. Data ini dapat digunakan untuk analisis lebih lanjut atau pemahaman lebih mendalam tentang sentimen ulasan yang terkandung dalam dataset.

3.3.3 Text Preprocessing

- **Data Cleaning**

Pada tahap ini, data ulasan aplikasi TASPEN Otentifikasi akan melewati proses pembersihan. Data cleaning adalah proses penting dalam penelitian ini untuk memastikan integritas dan kualitas data yang digunakan. Dalam konteks ini, salah satu

tindakan yang dilakukan adalah mengidentifikasi dan menangani nilai null atau kosong dalam dataset.

- **Case Folding**

Tahap Case Folding adalah proses mengubah dataset menjadi huruf kecil [24]. Untuk menjaga konsistensi dalam analisis teks, semua huruf dalam ulasan akan diubah menjadi huruf kecil (lowercase). Hal ini memastikan bahwa tidak ada perbedaan antara kata yang ditulis dalam huruf besar dan kecil, sehingga menghindari kesalahan dalam pemrosesan kata.

- **Stopword Removal**

Kata-kata umum yang tidak memberikan informasi penting seperti "dan," "atau," "yang," dan sebagainya, yang sering disebut sebagai stop words, akan dihapus dari ulasan. Stopword dapat diartikan juga untuk menghilangkan kata yang kurang efektif [25]. Hal ini bertujuan untuk mengurangi dimensi data dan memfokuskan analisis pada kata-kata yang memiliki makna sentimen.

- **Tokenizing**

Tokenizing adalah proses pemecahan teks pada kalimat menjadi potongan kata [26]. Pada langkah ini, ulasan akan dipecah menjadi token atau kata-kata individu. Ini merupakan langkah penting dalam pemrosesan teks karena memungkinkan analisis lebih lanjut pada tingkat kata. Setiap token akan dianggap sebagai entitas terpisah yang dapat dianalisis.

- **Stemming**

Stemming adalah proses mengubah kata berimbuhan menjadi kata dasar [24]. Stemming menyaring kata yang terdapat kata sambung, kata ganti, kata depan menjadi kata dasar, yaitu dengan menghilangkan awalan dan akhiran kata tersebut [25]. Misalnya, kata-kata seperti "makanan," "makananan," dan "makan" akan diubah menjadi "makan." Untuk melakukan stemming bahasa

Indonesia kita dapat menggunakan library Python Sastrawi. Hal ini membantu mengurangi variasi kata yang memiliki akar yang sama, sehingga kata-kata yang sebenarnya memiliki makna yang sama dapat dikenali dengan benar dalam analisis sentimen.

3.3.4 Splitting Data

Data ulasan akan dibagi menjadi dua subset utama: data pelatihan (training data) dan data pengujian (testing data) dengan `test_size = 0.20` atau 20% dari keseluruhan data akan digunakan sebagai data pengujian dan `random_state` nya 0. Data pengujian digunakan untuk menguji seberapa baik model tersebut dapat mengklasifikasikan sentimen pada ulasan yang belum pernah dilihat sebelumnya. Pembagian data ini penting untuk menghindari overfitting dan memvalidasi kinerja model.

3.3.5 Pembobotan TF-IDF

Proses TF-IDF digunakan untuk mengubah teks menjadi vektor numerik dengan bobot TF-IDF, yang kemudian digunakan dalam pelatihan model klasifikasi Naïve Bayes. Proses TF-IDF adalah cara yang umum digunakan dalam pemrosesan teks untuk mengatasi masalah dalam analisis teks. Dengan mengubah teks menjadi representasi vektor numerik ini, kita dapat menggunakan algoritma pembelajaran mesin, seperti Naïve Bayes, untuk mengklasifikasikan ulasan ke dalam kategori sentimen positif atau negatif berdasarkan vektor TF-IDF yang dihasilkan. Jadi, penggunaan TF-IDF dalam konteks ini adalah untuk mewakili setiap dokumen sebagai vektor numerik yang memperhitungkan pentingnya setiap kata dalam dokumen tersebut dalam konteks korpus secara keseluruhan. Hal ini membantu model klasifikasi Naive Bayes untuk memahami dan membedakan dokumen berdasarkan kata-kata yang muncul di dalamnya.

3.3.6 Pelatihan Model

Algoritma Multinomial Naïve Bayes digunakan dalam tahap klasifikasi sentimen. Model ini dilatih menggunakan data pelatihan yang telah dibobotkan dengan TF-IDF. Algoritma Naïve Bayes adalah algoritma klasifikasi yang berdasarkan teorema Bayes dengan asumsi bahwa semua atribut (kata-kata dalam konteks ini) adalah independen satu sama lain. Model ini digunakan untuk mengklasifikasikan ulasan pengguna pada data pengujian ke dalam kategori sentimen yang berbeda, seperti "Positif," atau "Negatif."

3.3.7 Evaluasi Model

Evaluasi hasil klasifikasi adalah langkah penting dalam penelitian ini. Hasil klasifikasi sentimen oleh model Algoritma Naïve Bayes akan dievaluasi menggunakan berbagai metrik seperti akurasi (tingkat keberhasilan), presisi (kemampuan model dalam mengidentifikasi sentimen yang benar), recall (kemampuan model dalam mengidentifikasi semua sentimen yang seharusnya diidentifikasi), F1-score (ukuran gabungan dari presisi dan recall), dan confusion matrix yang menggambarkan seberapa baik model dapat mengklasifikasikan ulasan. Evaluasi ini memberikan pemahaman yang lebih mendalam tentang kinerja model dalam mengenali sentimen ulasan dan berpotensi memberikan wawasan penting bagi TASPEN dalam meningkatkan layanannya.

BAB IV

HASIL DAN PEMBAHASAN

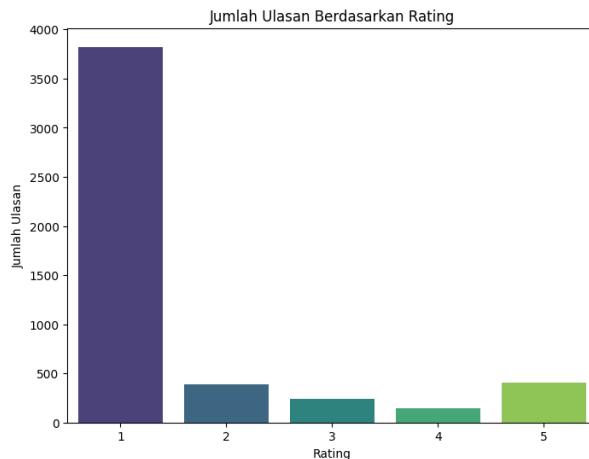
4.1 Data Ulasan

Total data ulasan yang diperoleh melalui proses scraping, terdapat 5000 data ulasan terbaru. Berikut merupakan tampilan 5 data teratasnya :

	reviewId	userName	userImage	content	score	thumbsUpCount	reviewCreatedVersion	at	replyContent	repliedAt	appVersion
0	1645bed6-e5f1-43c9-a7b5-9e342818d3f8	M T	lh.googleusercontent.com/a-/ALV-U...	Otentifikasi berulang-ulang juga selalu gagal....	1	7	1.6.9	2023-10-28 02:19:32	None	NaT	1.6.9
1	5fdc21a-c1bd-488f-85db-0541cd9d41e0	Relly Wicahyadi	lh.googleusercontent.com/a/ACg8oc...	Terima kasih TASPEN, setelah update yang v 169...	5	154	1.6.9	2023-10-12 07:05:09	None	NaT	1.6.9
2	765fbcd9-7cc8-49cc-91f0-dcf2aaad0815	Budi Sugono	lh.googleusercontent.com/a/ACg8oc...	Waahhh... harus berapakali mengulang dan mengu...	5	50	1.6.9	2023-10-10 10:33:48	None	NaT	1.6.9
3	9f2c9382-68f4-4b75-a9f5-17979bcd9e82	Linda Lisyani	lh.googleusercontent.com/a/ACg8oc...	Maaf sebelum nya kenapa sudah di upgrade aplik...	1	12	1.6.9	2023-10-25 00:38:04	None	NaT	1.6.9
4	09a811ca-40cf-4e25-a032-dff594179c54	REFNA WIRDASE	lh.googleusercontent.com/a-/ALV-U...	Aplikasi sering eror, jangan kan orang tua beru...	2	60	1.6.9	2023-10-07 07:26:16	None	NaT	1.6.9

Gambar 2. Data Awal

Setelah itu dilakukan Visualisasi Jumlah Ulasan Berdasarkan Rating atau kolom Score.



Gambar 3. Diagram Batang Jumlah Ulasan berdasarkan Rating

Dari diagram batang diatas terlihat bahwa Jumlah Ulasan terbanyak ada pada Rating 1 sedangkan Rating Lainnya tidak lebih dari 500 ulasan.

Dari Scrapping data yang telah dilakukan didapatkan banyak sekali kolom, kemudian kolom – kolom tersebut di filter karena hanya membutuhkan kolom content dan score.

		content	score
906		Simpel dan mudah	5
3337	Sebaiknya aplikasinya disesuaikan dengan kemam...		5
2234		Baik tanpa kendala	5
0	Otentikasi berulang-ulang juga selalu gagal.. ...		1
884		Bagus dan memudahkan	5

Gambar 4. Data setelah dilakukan filtering kolom

Setelah itu dilakukan Pemberian label dengan menambahkan kolom baru yang dinamakan 'Label' berdasarkan parameter score yang merupakan nilai atau skor yang menggambarkan sentimen dari suatu ulasan. Dengan ketentuan setiap ulasan dalam DataFrame my_df akan mendapatkan label sentimen "Positif" jika skornya adalah 4 atau 5, dan label "Negatif" jika skornya kurang dari 3.

		content	score	Label
906		Simpel dan mudah	5	Positif
3337	Sebaiknya aplikasinya disesuaikan dengan kemam...		5	Positif
2234		Baik tanpa kendala	5	Positif
0	Otentikasi berulang-ulang juga selalu gagal.. ...		1	Negatif
884		Bagus dan memudahkan	5	Positif

Gambar 5. Pemberian Label

4.2 Hasil Text Preprocessing

Setelah itu dilakukan Text Preprocessing melalui beberapa tahapan, yaitu : Data Cleaning, Case Folding, Stopword Removal, Tokenizing, dan Stemming. Berikut merupakan tampilan data setelah dilakukan proses tersebut.

	content	score	Label	text_clean	text_stopword	text_tokens	text_stemindo
0	Simpel dan mudah	5	Positif	simpel dan mudah	simpel mudah	[simpel, mudah]	simpel mudah
1	Sebaiknya aplikasinya disesuaikan dengan kemam...	5	Positif	sebaiknya aplikasinya disesuaikan dengan kemam...	aplikasinya disesuaikan kemampuan orang orang ...	[aplikasinya, disesuaikan, kemampuan, orang, o...]	aplikasi sesuai mampu orang orang tua pensiun ...
2	Baik tanpa kendala	5	Positif	baik tanpa kendala	kendala	[kendala]	kendala
3	Otentifikasi berulang-ulang juga selalu gagal... ...	1	Negatif	otentifikasi berulangulang juga selalu gagal ken...	otentifikasi berulangulang gagal diperbaiki cari...	[otentifikasi, berulangulang, gagal, diperbaiki,...	otentifikasi berulangulang gagal baik cari solus...
4	Bagus dan memudahkan	5	Positif	bagus dan memudahkan	bagus memudahkan	[bagus, memudahkan]	bagus mudah
5	MOHON PERBAIKI APLIKASI INI AGAR MEMPERMUDAH P...	1	Negatif	mohon perbaiki aplikasi ini agar mempermudah p...	mohon perbaiki aplikasi mempermudah proses ote...	[mohon, perbaiki, aplikasi, mempermudah, proses...	mohon baik aplikasi mudah proses otentifikasi
6	Sangat susah untuk otentifikasi dan harus diul...	1	Negatif	sangat susah untuk otentifikasi dan harus diul...	susah otentifikasi diluangulang capek	[susah, otentifikasi, diluangulang, capek]	susah otentifikasi diluangulang capek
7	Lebih baik untuk otentifikasi	4	Positif	lebih baik untuk otentifikasi	otentifikasi	[otentifikasi]	otentifikasi
8	Biasanya juga lancar ko sekarang server sibuk ...	4	Positif	biasanya juga lancar ko sekarang server sibuk ...	lancar ko server sibuk updet	[lancar, ko, server, sibuk, updet]	lancar ko server sibuk updet
9	kenapa dalam mata susah ke prediksi	4	Positif	kenapa dalam mata susah ke prediksi	mata susah prediksi	[mata, susah, prediksi]	mata susah prediksi

Gambar 1. Hasil Text Preprocessing

Setelah melalui tahapan ini, teks ulasan telah dibersihkan, diubah menjadi huruf kecil, dan diproses sehingga siap untuk digunakan dalam analisis sentimen. Selanjutnya data dibagi menjadi data latih dan data uji dengan perbandingan 80% data latih dan 20% data uji.

Kemudian dilakukan Pembobotan teks menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency) untuk mengubah teks ulasan menjadi vektor numerik yang dapat digunakan dalam model klasifikasi. Objek TfidfVectorizer dari sklearn.feature_extraction.text digunakan untuk mengubah teks ulasan menjadi vektor TF-IDF. Data latih (X_train) diubah menjadi vektor TF-IDF dengan fit_transform, sedangkan data pengujian (X_test) diubah menjadi vektor TF-IDF dengan transform. Hasil pembobotan TF-IDF data latih dan data uji disimpan dalam tfidf_train dan tfidf_test.

Tahap Selanjutnya Pelatihan Model Naive Bayes Menggunakan algoritma klasifikasi Naive Bayes (MultinomialNB) dari sklearn.naive_bayes. Model Naive Bayes dilatih dengan menggunakan vektor TF-IDF data latih (tfidf_train) dan label data latih (y_train) dengan perintah nb.fit(tfidf_train, y_train).

4.3 Hasil Evaluasi Model

Hasil dari analisis sentimen ulasan aplikasi TASPEN Otentifikasi menggunakan metode klasifikasi Algoritma Naïve Bayes dapat dilihat dari evaluasi kinerja model. Model Multinomial Naïve Bayes menunjukkan hasil sebagai berikut:

```

MultinomialNB Accuracy: 0.9149159663865546
MultinomialNB Precision: 0.9161220043572985
MultinomialNB Recall: 0.9952662721893492
MultinomialNB f1_score: 0.9540555870674986
confusion_matrix:
[[841  4]
 [ 77 30]]
=====
      precision    recall   f1-score   support
Negatif        0.92     1.00     0.95      845
Positif        0.88     0.28     0.43      107
accuracy          -         -         -      952
macro avg       0.90     0.64     0.69      952
weighted avg    0.91     0.91     0.89      952

```

Gambar 7. Hasil Evaluasi Model

Dalam hasil evaluasi model pada penelitian ini, digunakan metode klasifikasi Multinomial Naive Bayes untuk menganalisis sentimen ulasan aplikasi “Taspen Otentikasi.” Hasil evaluasi menunjukkan kinerja model yang baik dengan beberapa catatan yang perlu diperhatikan.

Akurasi model mencapai sekitar 91.49%, yang menunjukkan kemampuan model dalam memprediksi sentimen ulasan dengan tingkat akurasi yang tinggi. Presisi model sekitar 91.61%, menunjukkan bahwa sebagian besar prediksi sentimen positif adalah benar. Namun, recall model sangat tinggi, sekitar 99.53%, yang menunjukkan kemampuan model dalam mengidentifikasi ulasan yang seharusnya positif. Skor F1 model sekitar 95.41%, mencerminkan kinerja keseluruhan yang baik dalam memprediksi sentimen ulasan.

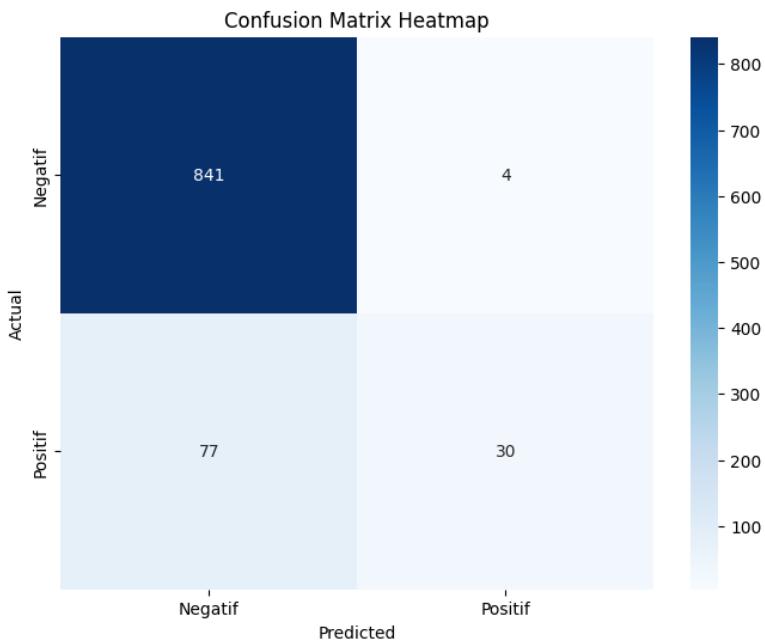
Confusion Matrix menunjukkan hasil prediksi model terhadap sentimen ulasan. Model memiliki True Positives (TP) sebanyak 30, True Negatives (TN) sebanyak 841, False Positives (FP) sebanyak 4, dan False Negatives (FN) sebanyak 77. Model sangat baik dalam mengidentifikasi sentimen negatif, tetapi memiliki kesalahan dalam mengidentifikasi sentimen positif.

Classification Report memberikan perincian lebih lanjut tentang kinerja model untuk setiap kelas sentimen. Model memiliki presisi yang tinggi untuk sentimen ‘Negatif’ (0.92) dan recall yang sempurna (1.00). Namun, kinerja model

lebih rendah untuk sentimen “Positif” dengan presisi 0.88 dan recall 0.28. Hal ini mungkin disebabkan oleh ketidakseimbangan kelas di mana jumlah ulasan dengan sentimen positif jauh lebih sedikit dibandingkan dengan sentimen negatif. Hasil rata-rata tertimbang menunjukkan akurasi sekitar 91%, yang menggambarkan kinerja keseluruhan model. Classification report memberikan insight yang lebih detail tentang kinerja model pada setiap kelas.

Dalam keseluruhan, model Multinomial Naive Bayes telah menunjukkan kinerja yang baik dalam mengidentifikasi sentimen negatif, tetapi masih ada beberapa perbaikan yang diperlukan dalam mengidentifikasi sentimen positif. Dengan akurasi sekitar 91%, model ini dapat digunakan untuk menganalisis ulasan aplikasi “Taspen Otentikasi” dengan baik, namun perlu diperhatikan dalam pengidentifikasiannya sentimen positif.

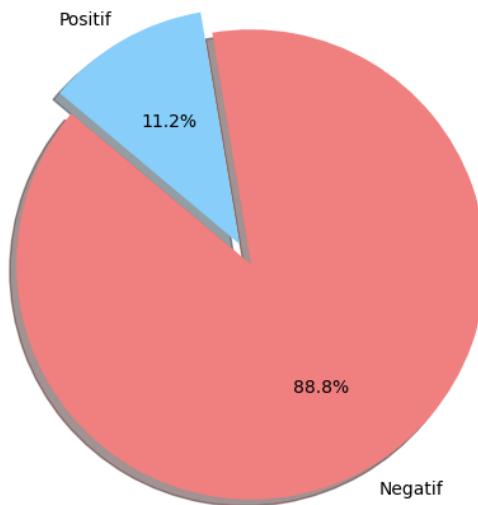
Confusion Matrix ditampilkan dalam bentuk heatmap menggunakan seaborn untuk membantu dalam visualisasi performa model.



Gambar 8. Confusion Matrix

Persentase kelas sentimen ("Negatif" dan "Positif") pada data uji ditampilkan dalam bentuk diagram lingkaran (pie chart) untuk memberikan gambaran distribusi sentimen.

Persentase Sentimen Positif dan Negatif



Gambar 9. Pie Chart Persentase Sentimen

Dari total ulasan yang dianalisis, terdapat persentase sentimen positif sebesar 11.2%, sementara sentimen negatif mendominasi dengan persentase yang signifikan, mencapai 88.8%. Hasil ini memberikan gambaran yang jelas tentang persepsi pengguna terhadap kualitas dan kinerja aplikasi TAS PEN Otentikasi.

Dominator sentimen negatif yang tinggi dapat menunjukkan adanya kekurangan atau ketidakpuasan yang mungkin dialami oleh sebagian pengguna dalam penggunaan aplikasi ini. Faktor-faktor seperti masalah fungsionalitas, antarmuka pengguna, atau performa aplikasi dapat menjadi penyebab potensial dari respon negatif ini. Oleh karena itu, pemahaman mendalam terhadap ulasan-ulasan negatif ini dapat memberikan panduan yang berharga untuk melakukan perbaikan dan peningkatan pada aplikasi TAS PEN Otentikasi.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat menarik kesimpulan sebagai berikut Penelitian ini telah berhasil mengimplementasikan metode klasifikasi algoritma Naïve Bayes untuk menganalisis sentimen ulasan pengguna pada aplikasi TASPEN Otentikasi. Hasil penelitian menunjukkan bahwa metode Naïve Bayes, khususnya model Multinomial Naïve Bayes, memiliki akurasi yang baik sekitar 91.49%.

Hasil penelitian menunjukkan bahwa sentimen negatif mendominasi, sementara sentimen positif memiliki persentase yang lebih rendah. Meskipun demikian, model klasifikasi yang digunakan dalam penelitian ini menunjukkan kinerja yang baik dalam hal akurasi dan presisi. Meskipun masih ada ruang untuk perbaikan dalam mengidentifikasi sentimen positif secara lebih optimal. Dengan demikian, analisis sentimen menjadi alat yang berharga bagi pengembang aplikasi untuk memahami persepsi dan respon pengguna terhadap aplikasi TASPEN Otentikasi. Hasil penelitian ini diharapkan dapat memberikan panduan yang berharga dalam upaya meningkatkan kualitas dan kepuasan pengguna terhadap aplikasi tersebut.

DAFTAR PUSTAKA

- [1] Nuriah Jahan, "Analisi Kepuasan Nasabah Terhadap Pelayanan Pembayaran Program Pensiun Pegawai Negeri Sipil (PNS) Di PT Taspen (Persero) Cabang Yogyakarta", (Yogyakarta: Fakultas Ekonomi,2017). Hal 1
- [2] Dita Destia Anggraini "Evektifitas Penggunaan Aplikasi Taspen Otentikasi Di PT Taspen (Persero) Kator Cabang Palembang", (Fakultas Ilmu Sosial dan Ilmu Politik, 2020), hal 6.
- [3] Undang-Undang Nomor 19 Tahun 2003 Tentang Badan Usaha Milik Negara
- [4] Priskilla Grace Langkai, "Kualitas Pelayanan Pembayaran Pensiun Pegawai Negeri Sipil (PNS) Pada PT Taspen (Persero) Cabang Manado" (Jurnal Administrasi Bisnis:2018). Hal 74. Vol.7, hal 26.
- [5] Adek Nurhasana, "Evektifitas Penggunaan Sistem Aplikasi Otentikasi Penerima Pensiun PT.Taspen Secara Digital Dengan Menggunakan Smartphone Dikota Padang" (Jural Ilmu Administrasi: 2020).hal 17, Vol 2, hal 43.
- [6] Lola, " Analisi Peran Perseroan Terbatas Taspen Dalam Pelayanan Sistem Aplikasi Otentikasi Penerima Dana Pensiunan Di Kota Banda Aceh" (Fakultas Ilmu Sosial: 2022), hal 5.
- [7] Bahary A, Sibaroni Y, Mubarok M. (1, 2019) : Sentiment analysis of student responses related to information system services using Multinomial Naive Bayes (Case study: Telkom University). Journal of Physics: Conference Series, Vol. 1192.
- [8] Singh, Vivek and Piryani, Rajesh and Uddin, Ashraf and Waila, Pranav. (2013) : Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification. 2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), pp. 712-717.
- [9] PT TASPEN (PERSERO), "sejarah perusahaan," 2018. [Online]. Available: <https://taspen.co.id/tentang-kami/sejarah>.

- [10] B. Liu, “Sentiment Analysis and Opinion Mining”, Synthesis Lecturers on Human Languages Technologies, vol. 5, no. 1, pp. 1-167, 2012.
- [11] I. F. Rozi, S. H. Pramono, dan E. A. Dahlan, “Implementasi Opinion Mining (Analisis Sentimen) untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi”, Jurnal EECCIS, vol. 6, no. 1, 2012.
- [12] Falahah dan D. D. A. Nur, “Pengembangan Aplikasi Sentiment Analysis Menggunakan Naïve Bayes (Studi Kasus Sentiment Analysis dari Media Twitter)”, Seminar Nasional Sistem Informasi Indonesia, November 2015.
- [13] W G. A. Buntoro, T. B. Adji, dan A. E. Purnamasari, “Sentiment Analysis Candidates of Indonesian Presiden 2014 with Five Class Attribute”, International Journal of Computer Applications, vol. 123, no. 2, pp. 23-29, Februari 2016.
- [14] TASPERN. (2020). TASPERN Otentifikasi: Aplikasi Mobile. Diakses dari <https://www.taspen.co.id/>.
- [15] R. Watrianthos, S. Suryadi, D. Irmayani, M. Nasution, and E. F. S. Simanjorang, “Sentiment Analysis of Traveloka App Using Naïve Bayes Classifier Method,” Int. J. Sci. Technol. Res., Vol. 8, No. 7, pp. 786–788, 2019, doi: <https://doi.org/10.31227/osf.io/2dbe4>.
- [16] N. Herlinawati, Y. Yuliani, S. Faizah, and W. Gata, “Analisis Sentimen ZOOM Cloud Meetings di Play Store Menggunakan Naïve Bayes dan Support Vector Machine,” Vol. 5, No. 2, pp. 293–298, 2020, doi: <https://doi.org/10.24114/cess.v5i2.18186>.
- [17] R. N. Devita et al., “Perbandingan Kinerja Metode Naïve Bayes dan K-Nearest Neighbor Untuk Klasifikasi Artikel Berbahasa Indonesia,” Vol. 5, No. 4, pp. 427–434, 2018, doi: <https://doi.org/10.25126/jtiik.201854773>.
- [18] D. Darwis, N. Siskawati, and Z. Abidin, “Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional,” vol. 15, no. 1, 2021.
- [19] “1.9.NaïveBayes.” https://scikitlearn.org/stable/modules/naive_bayes.html (accessed Mar. 13, 2023).
- [20] A. Fauzi, E. B. Setiawan, and Z. K. A. Baizal, “Hoax News Detection on Twitter using Term Frequency Inverse Document Frequency and Support

Vector Machine Method,” in Journal of Physics: Conference Series, 2019.
Doi: 10.1088/1742-6596/1 192/1/01025.

- [21] I. M. Karo Karo, R. Romia, S. Dewi, and P. M. Fadilah, “Hoax Detection on Indonesian Tweets using Naïve Bayes Classifier with TF-IDF,” Journal of Information System Research (JOSH), vol. 4, no. 3, pp. 914-919, Apr. 2023, doi: 10.47065/josh.v4i3.3317.
- [22] Nurdin, M. Hutomi, M. Qamal, and B. Bustami, “Sistem Pengecekan Toko Online Asli Atau Dropship pada Shopee Menggunakan Algoritma Breadth First Search,” J. RESTI (Rekayasa Sist. dan Teknol. Informasi), Vol. 4, No. 6, pp. 1117–1123, Dec. 2020, doi: <https://doi.org/10.29207/resti.v4i6.2514>.
- [23] Boeing, G., Wadell, P. (2016). New Insight into Rental Housing Markets across the United States: Web Scrapping and Analyzing Craigslist Rental Listings. Journal of Planning Education and Research, 37(4), 457-476
- [24] M. A. Rosid, A. S. Fitriani, I. R. I. Astutik, N. I. Mulloh, and H. A. Gozali, “Improving Text Preprocessing for Student Complaint Document Classification Using Sastrawi,” in IOP Conference Series: Materials Science and Engineering, Jul. 2020, vol. 874, no. 1. doi: 10.1088/1757-899X/874/1/012017.
- [25] A. Rahman Isnain, A. Indra Sakti, D. Alita, and N. Satya Marga, “Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma SVM,” JDMSI, vol. 2, no. 1, pp. 31–37, 2021, [Online]. Available: <https://t.co/NfhnfMjtXw>.
- [26] M. I. Fikri, T. S. Sabrila, and Y. Azhar, “Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter,” SMATIKA, vol. 10, no. 2, pp. 71–76, 2020.

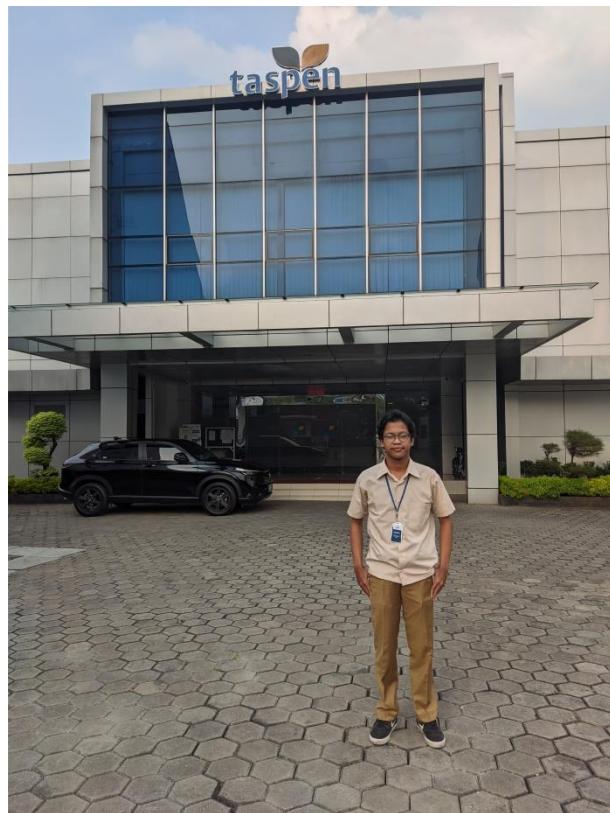
LAMPIRAN

Lampiran 1. Source Code

https://github.com/mariomanroe/Analisis-Sentimen-Aplikasi-Taspen-Otentifikasi-Menggunakan-Metode-Klasifikasi-Algoritma-Naive-Bayes/blob/main/KP_MARIO.ipynb

Lampiran 2. Kegiatan Saat Keja Praktik





Lampiran 3. Sertifikat Magang





KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI

INSTITUT TEKNOLOGI SUMATERA

Jalan Terusan Ryacudu, Way Hui, Jati Agung, Lampung Selatan 35365

Telepon (0721) 8030188, Fax. (0721) 8030189, Email: pusat@itera.ac.id

www.itera.ac.id

DATA PELAKSANAAN KERJA PRAKTIK

INSTANSI

Nama Perusahaan : PT. TASPEN (Persero)
Alamat : Jl. Kemakmur NO.39, Marga Jaya
No. Telp : 02128519766
Pembimbing Lapangan : Ikhhat Permana

TUGAS/PROYEK

Judul

: Analisis Sentimen Ulasan Aplikasi Taspen Otentikasi Menggunakan Metode Naive Bayes

Ringkasan Tugas

: Pengambilan data → Preprocessing Data → Splitting Data → Pembelahan TF-IDF → Pelatihan Model → Evaluasi Model → Analisis Hasil → Selanjutnya

Hasil yang diharapkan

: - Mengembangkan sistem analisis sentimen menggunakan metode klasifikasi algoritma naive bayes untuk memklifikasi ulasan aplikasi Taspen Otentikasi menjadi kategori sentimen positif atau negatif
- Mengukur kepuasan Pengguna aplikasi Taspen Otentikasi berdasarkan analisis sentimen ulasan Pengguna.

Waktu Pelaksanaan

: 6 Juni - 6 Agustus 2023

Mengetahui,
Pembimbing Lapangan

(Ikhhat Permana)



KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI

INSTITUT TEKNOLOGI SUMATERA

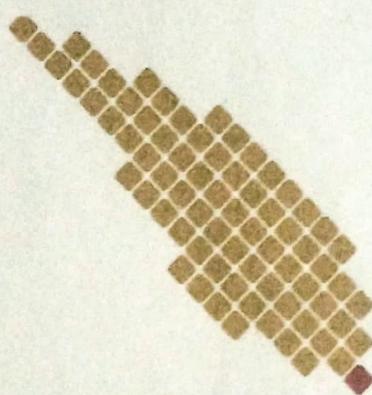
Jalan Terusan Ryacudu, Way Hui, Jati Agung, Lampung Selatan 35365

Telepon (0721) 8030188, Fax. (0721) 8030189, Email: pusat@itera.ac.id

www.itera.ac.id

Lampiran 3

LAPORAN HARIAN KERJA PRAKTIK



ITERA

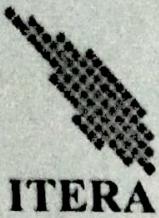
Nama	: Mario Desendi Manurung
NIM	: 120450082
Judul	: Analisis Sentimen Ulasan Aplikasi Taspen Dicintikari Menggunakan Metode Naive Bayes
Pembimbing Lapangan	: Ikhhat Permana
Dosen Pembimbing	: Ardika Satria, M.Si

PRODI SAINS DATA

JURUSAN SAINS

INSTITUT TEKNOLOGI SUMATERA

TAHUN 2023



KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI
INSTITUT TEKNOLOGI SUMATERA

Jalan Terusan Ryacudu, Way Hui, Jati Agung, Lampung Selatan 35365
Telepon (0721) 8030188, Fax. (0721) 8030189, Email: pusat@itera.ac.id
www.itera.ac.id

**Lembar Kegiatan Harian Kerja Praktik
Di Perusahaan/ Instansi**

Hari ke-	Hari/Tanggal	Uraian Kegiatan	Paraf Pembimbing Lapangan
1.	16 -06 -2023	Olahraga Pagi , Pengenalan , dan Penempatan Posisi kerja	/
2.	19 -06 -2023	Memilah Dokumen Peserta Pensiun , scan dokumen , upload data Pada aplikasi ELO,Mendisir Dokumen	/
3.	20 -06 -2023	Memilah Dokumen , scan dokumen , upload dokumen Pada aplikasi Elo , Mendisir Dokumen .	/
4.	21 -06 -2023	Memilah Dokumen , scan dokumen , upload dokumen Pada aplikasi ELO , Mendisir Dokumen	/
5.	22 -06 -2023	Memilah Dokumen , scan dokumen , upload dokumen Pada aplikasi ELO , Mendisir Dokumen .	/
6.	23 -06 -2023	Mencabut dokumen yang sudah Penuh , Melakukan Create Path dokumen yang sudah Penuh , Olahraga Pagi	/
7.	26 -06 -2023	Memilah dokumen , scan dokumen , upload dokumen Pada Aplikasi ELO , Mendisir Dokumen	/

**KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI
INSTITUT TEKNOLOGI SUMATERA**

Jalan Terusan Ryacudu, Way Hui, Jati Agung, Lampung Selatan 35365
Telepon (0721) 8030188, Fax. (0721) 8030189, Email: pusat@itera.ac.id
www.itera.ac.id

**Lembar Kegiatan Harian Kerja Praktik
Di Perusahaan/ Instansi**

Hari ke-	Hari/Tanggal	Uraian Kegiatan	Para/Pembimbing Lapangan
8.	27-06-2023	Memilah dokumen, Scan dokumen, upload dokumen Pada Aplikasi ELO, Mendasir Dokumen	/z
9.	3 -07-2023	Memilah dokumen, scan dokumen, upload dokumen Pada Aplikasi Elo, Mendasir Dokumen	/z
10.	4 -07-2023	Memilah dokumen, scan dokumen, upload dokumen Pada Aplikasi Elo, Mendasir Dokumen	/z
11.	5-07-2023	Memilah dokumen, scan dokumen, upload dokumen Pada Aplikasi ELO, Mendasir Dokumen	/z
12.	6-07-2023	Memilah dokumen, Scan dokumen, Upload dokumen Pada APLIKASI ELO, Mendasir Dokumen	/z
13.	7-07-2023	Olahraga Pagi, mencabut dokumen. Sang sudah punah, Melakukan Cuci lahan dokumen sang sudah punah.	/z
14.	10-07-2023	Verifikasi berkas SPTB, Scan E-klim	/z
15.	11-07-2023	Verifikasi berkas SPTB, Scan E-klim	/z

Lembar Kegiatan Harian Kerja Praktik
Di Perusahaan/ Instansi

Hari ke-	Hari/Tanggal	Uraian Kegiatan	Para Pembimbing Lapangan
16.	12 -07-2023	Verifikasi Berkas SPTB , Scan E-klim	✓
17.	13 -07-2023	Verifikasi Berkas SPTB , Scan E-klim	✓
18.	14 -07-2023	Verifikasi Berkas SPTB , Scan E-klim, Olahraga Pagi	✓
19.	17 -07-2023	Memilah dokumen , Scan dokumen , upload dokumen Pada Aplikasi Elo , Mendisit dokumen .	✓
20.	18 -07-2023	Memilah dokumen , Scan dokumen , upload dokumen Pada Aplikasi Elo , Mendisit dokumen .	✓
21.	20 -07-2023	Memilah dokumen , Scan dokumen , upload dokumen Pada Aplikasi Elo , Mendisit dokumen .	✓
22.	21 -07-2023	Verifikasi SPTB , Scan E-klim, Olahraga	✓
23.	24 -07-2023	Verifikasi SPTB , Scan E-klim	✓
24.	25 -07 - 2023	Memilah dokumen , upload dokumen Pada Aplikasi Elo , input data dari Pernah FC BEKASI	✓
25.	26 -07 - 2023	Input data dari Pernah FC BEKASI	✓



KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI
INSTITUT TEKNOLOGI SUMATERA

Jalan Terusan Ryacudu, Way Hui, Jati Agung, Lampung Selatan 35365
Telepon (0721) 8030188, Fax. (0721) 8030189, Email: pusat@itera.ac.id
www.itera.ac.id

Lembar Kegiatan Harian Kerja Praktik Di Perusahaan/ Instansi



KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI

INSTITUT TEKNOLOGI SUMATERA

Jalan Terusan Ryacudu, Way Hui, Jati Agung, Lampung Selatan 35365

Telepon (0721) 8030188, Fax. (0721) 8030189, Email: pusat@itera.ac.id

www.itera.ac.id

FORMULIR PENILAIAN PEMBIMBING LAPANGAN

Nama Mahasiswa : Mario Desendi Manurung
NIM : 120450082
Pembimbing Lapangan : Ikhhat Permana
Perusahaan : PT. TASPEN (Persero)

No.	Aspek Penilaian	Indikator	Nilai				
			SK	K	C	B	SB
1	Integritas(etika, moral, dan kesungguhan)	Peserta KP hadir tepat waktu sesuai jam kerja perusahaan	1	2	3	4	(5)
		Presentase kehadiran peserta KP tinggi	1	2	3	4	(5)
		Peserta KP mematuhi semua peraturan perusahaan	1	2	3	4	(5)
		Peserta KP berlaku disiplin dalam setiap tindakan	1	2	3	4	(5)
2	Ketepatan waktu dalam bekerja	Peserta KP mampu menyelesaikan pekerjaan sesuai dengan jadwal yang ditetapkan	1	2	3	4	(5)
3	Kerja sama dalam tim	Peserta KP mampu berinteraksi dengan tim kerja	1	2	3	(4)	5
4	Komunikasi	Peserta KP mampu melakukan komunikasi secara aktif	1	2	3	(4)	5
		Peserta KP mampu beradaptasi dengan lingkungan perusahaan	1	2	3	(4)	5
5	Penggunaan teknologi informasi	Peserta KP mampu memanfaatkan teknologi informasi secara efektif	1	2	3	4	(5)
6	Keahlian berdasarkan bidang ilmu	Peserta KP mampu menyelesaikan pekerjaan sesuai dengan pengetahuan bidang prodinya	1	2	3	4	(5)
7	Pengembangan diri	Peserta KP mampu memanfaatkan informasi formal untuk menyelesaikan pekerjaan	1	2	3	4	(5)



KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI

INSTITUT TEKNOLOGI SUMATERA

Jalan Terusan Ryacudu, Way Hui, Jati Agung, Lampung Selatan 35365
Telepon (0721) 8030188, Fax. (0721) 8030189, Email: pusat@itera.ac.id
www.itera.ac.id

	Peserta KP memiliki inisiatif untuk menyampaikan ide-ide atau pun metode-metode baru	1	2	3	4	5
	Peserta KP mampu memberikan kontribusi nyata dalam memberikan solusi terhadap permasalahan yang terjadi di perusahaan	1	2	3	4	5
	TOTAL				60	

Bekasi, 6 Agustus 2023

Pembimbing Lapangan

(Ikhat Permana)

INSTITUT TEKNOLOGI SUMATERA

Jalan Terusan Ryacudu, Way Hui, Jati Agung, Lampung Selatan 35365
Telepon (0721) 8030188, Fax. (0721) 8030189, Email: pusat@itera.ac.id
www.itera.ac.id

LAPORAN BIMBINGAN

NO.	Tanggal	Kegiatan	Paraf Dosen Pembimbing
1.	6 September 2023	Konsultasi tentang judul dan data yang didapat dari instansi.	Ang
2.	1 November 2023	Revisi Laporan (Perbaiki format, Pengetahuan kata, mengganti flowchart, Perbaiki Paper Pustaka)	Ang

INSTITUT TEKNOLOGI SUMATERA

Jalan Terusan Ryacudu, Way Hui, Jati Agung, Lampung Selatan 35365
Telepon (0721) 8030188, Fax. (0721) 8030189, Email: pusat@itera.ac.id
www.itera.ac.id

Formulir Pendaftaran Seminar Kerja Praktik Program Studi Sains Data

Nama : Mario Desendi Manurung
NIM : 120450082
JUDUL : Analisis Sentimen Ulasan Aplikasi Taspen Otentifikasi Menggunakan Metode Klasifikasi Algoritma Naïve Bayes

Dengan ini mengajukan Seminar Kerja Praktik pada
Hari/Tanggal : Senin, 13 November 2023
Tempat : F 206 (Pukul 15.00 - selesai)

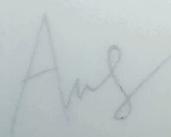
Lampung Selatan, 09 November 2023

Pemohon,



Mario Desendi Manurung
120450082

Menyetujui,
Dosen Pembimbing



Ardika Satria, M.Si
NRK. 1997111020231425

Mengetahui,
Koordinator Prodi Sains Data



Tirta Setiawan, S.Pd., M.Si
NIP. 199008222022031003