

**ANALISIS SENTIMEN BERBASIS ASPEK PADA APLIKASI LINKAJA
BERDASARKAN ULASAN PENGGUNA DENGAN MENGGUNAKAN
LATENT DIRECTIONAL ALLOCATION DAN
ATTENTION BASED-LSTM**

SKRIPSI

ANNISA PUTRI DAULAY

191402030



**PROGRAM STUDI S1 TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA
2024**

**ANALISIS SENTIMEN BERBASIS ASPEK PADA APLIKASI LINKAJA
BERDASARKAN ULASAN PENGGUNA DENGAN MENGGUNAKAN
LATENT DIRECTIONAL ALLOCATION DAN
ATTENTION BASED-LSTM**

SKRIPSI

Diajukan untuk melengkapi tugas dan memenuhi syarat memperoleh ijazah Sarjana
Teknologi Informasi

ANNISA PUTRI DAULAY

191402030



**PROGRAM STUDI S1 TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA**

2024

PERSETUJUAN

Judul : ANALISIS SENTIMEN BERBASIS ASPEK PADA
APLIKASI LINKAJA BERDASARKAN ULASAN
PENGGUNA DENGAN MENGGUNAKAN LATENT
DIRECHLET ALLOCATION DAN ATTENTION
BASED-LSTM

Kategori : SKRIPSI

Nama Mahasiswa : ANNISA PUTRI DAULAY

Nomor Induk Mahasiswa : 191402030

Program Studi : SARJANA (S1) TEKNOLOGI INFORMASI

Fakultas : ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA

Medan, 12 Juli 2024

Komisi Pembimbing:

Pembimbing 2,



Fanindia Purnamasari S.TI., M.IT
NIP.198908172019032023

Pembimbing 1,





Dedy Arisandi S.T., M.Kom.
NIP.197908312009121002

Diketahui/Disetujui oleh

Program Studi S1 Teknologi Informasi

Ketua,

Dedy Arisandi S.T., M.Kom.
NIP. 197908312009121002

PERNYATAAN

**ANALISIS SENTIMEN BERBASIS ASPEK PADA APLIKASI LINKAJA
BERDASARKAN ULASAN PENGGUNA DENGAN MENGGUNAKAN
LATENT DIRECTIONAL ALLOCATION DAN
ATTENTION BASED-LSTM**

SKRIPSI

Saya mengakui bahwa skripsi ini adalah hasil karya saya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing telah disebutkan sumbernya.

Medan, 12 Juli 2024

Annisa Putri Daulay

191402030

UCAPAN TERIMA KASIH

Penulis mengucapkan rasa syukur kepada Allah SWT atas berkah, petunjuk, kesabaran, dan kelapangan hati yang telah diberikan, sehingga penulis mampu menyelesaikan skripsi ini sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara.

Skripsi ini tidak mungkin selesai tanpa adanya doa, bantuan, dan dukungan dari pihak-pihak di sekitar penulis. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada :

1. Keluarga penulis, Ayah, H. Muhamad Rizal Daulay, SE dan Ibu, Hj. Santi Harahap atas semua doa, semangat, dan cinta yang selalu diberikan kepada penulis sehingga mampu menjalani kehidupan hingga di titik ini. Kedua adik, Alya Najwa Hapsari Daulay dan Almira Izanti Kamilah Daulay atas kasih sayangnya.
2. Keluarga besar dari Kakek H. Darwin Daulay dan Nenek Hj. Latifah Hannum Nasution atas kasih sayangnya dan doanya kepada penulis serta keluarga besar dari Kakek Alm. H. Mei Ramlan Harahap dan Nenek Almh. Hj. Sangkot Nasution, kepada kakek dan nenek walaupun tidak dapat bertemu lagi tapi terima kasih untuk kasih sayangnya sampai akhir hayat kalian berdua.
3. Ibu Dr. Maya Silvi Lydia B.Sc., M.Sc. selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara.
4. Bapak Dedy Arisandi S.T.,M.Kom., selaku Ketua Program Studi S1 Teknologi Informasi Universitas Sumatera Utara dan Dosen Pembimbing 1 yang telah bersedia meluangkan waktu untuk memberikan bimbingan dan masukan kepada penulis selama pengerjaan skripsi ini.
5. Ibu Fanindia Purnamasari S.TI., M.IT, selaku Dosen Pembimbing 2 yang juga sudah banyak meluangkan waktu untuk memberikan bimbingan dan masukan kepada penulis selama pengerjaan skripsi ini.

6. Ibu Dr. Erna Budhiarti Nababan M.IT, selaku Dosen Pembimbing 1 saya dan Bapak Seniman S.Kom., M.Kom, selaku Dosen Pembimbing 2 saya yang sudah memberikan banyak masukan kepada penulis.
7. Segenap Dosen Fakultas Ilmu Komputer dan Teknologi Informasi serta staf akademik yang telah membantu penulis selama proses penyelesaian studi.
8. Para sahabat yang sudah banyak membantu dalam setiap proses perkuliahan, Rizka Annisa Hidayat, Sity Fadia Al Haya Maswin, Nurul Atiqah Siregar, Mhd. Arsyah Fikri, Huzaifah Lais Lubis, Haiqal Rizky Ramadhan, Nanda Ambiya, Irsyad Fauzi Hadenggan Munthe, dan Arafah Nur Ihza.
9. Jethro Vetric Sihotang, untuk semua dukungan yang diberikan sehingga penulis tidak merasa sendiri dalam penyelesaian masa tugas akhir ini.
10. Ananda Zahra Salsabila, sahabat penulis yang selalu ada dan membantu sejak masa SMP hingga perkuliahan saat ini.
11. Teman-teman BPH HIMATIF Periode 2022/2023, Aulia Rahman Partomuan Sihite, Geylfedra Matthew Panggabean, Fildzah Alifia Lubis, Nada Salsabila, dan Dhea Novianty Sitompul.
12. Angkatan 2019 Teknologi Informasi yang bersama-sama mengarungi masa perkuliahan.
13. Dan semua pihak yang tidak dapat disebutkan satu per satu, yang memberikan bantuan kepada penulis dalam menjalani masa perkuliahan ini.

Medan, 12 Juli 2024

Annisa Putri Daulay

191402030

**ANALISIS SENTIMEN BERBASIS ASPEK PADA APLIKASI LINKAJA
BERDASARKAN ULASAN PENGGUNA DENGAN MENGGUNAKAN
LATENT DIRECTIONAL ALLOCATION DAN
ATTENTION BASED-LSTM**

ABSTRAK

Financial Technology yang semakin marak digunakan di Indonesia salah satunya adalah Sistem Pembayaran Digital. LinkAja, sebagai salah satu aplikasi yang populer di kalangan masyarakat, masih memiliki berbagai kendala dalam penggunaannya, seperti layanan pelanggan yang lambat merespon keluhan, transaksi yang gagal, proses pengembalian dana yang lama, dan pembaruan aplikasi yang sering mengakibatkan error. Sangat penting untuk memahami opini pengguna terhadap aplikasi ini, dengan diperlukannya peningkatan suatu kinerja layanan yang dapat memberikan suatu pengalaman dengan hasil yang lebih baik bagi pengguna. Namun, banyaknya ulasan yang ada membuat proses analisis secara manual menjadi tidak efisien. Oleh karena itu, penelitian ini mengembangkan sistem otomatis untuk mengolah dan mengelompokkan ulasan pengguna aplikasi LinkAja berdasarkan aspek kemudahan, layanan pengguna, kualitas, dan kecepatan menggunakan metode Latent Dirichlet Allocation (LDA) dan Attention Based -Long Short Term Memory (LSTM). Penelitian ini menunjukkan bahwa model Attention-Based LSTM mencapai akurasi 90% untuk aspek kemudahan, 92% untuk aspek layanan pengguna, 85% untuk aspek kualitas, dan 92% untuk aspek kecepatan dengan rata-rata akurasi keseluruhan 89,75%.

Kata Kunci : *Financial Technology, LinkAja, Analisis Sentimen berbasis aspek, Attention-LSTM, LDA*

**ASPECT-BASED SENTIMENT ANALYSIS ON THE LINKAJA
APPLICATION BASED ON USER REVIEWS USING
LATENT DIRECTIONAL ALLOCATION AND
ATTENTION-BASED LSTM**

ABSTRACT

One of the financial technologies that is increasingly being used in Indonesia is the Digital Payment Systems. LinkAja, as one of the popular applications among the public, still faces various issues in its use, such as slow customer service response to complaints, failed transactions, long refund processes, and frequent errors due to application updates. It is crucial to understand user opinions about this application to improve service performance, providing a better experience for users. However, the large number of reviews makes manual analysis inefficient. Therefore, this study develops an automated system to process and categorize user reviews of the LinkAja application based on aspects of ease of use, customer service, quality, and speed using Latent Dirichlet Allocation (LDA) and Attention-Based Long Short Term Memory (LSTM) methods. This study shows that the Attention-Based LSTM model achieves an accuracy of 90% for ease of use, 92% for customer service, 85% for quality, and 92% for speed, with an overall average accuracy of 89,75%.

Keyword : *Financial Technology, LinkAja, Aspect Based Sentiment Analysis, Attention-LSTM, LDA*

DAFTAR ISI

PERSETUJUAN	Error! Bookmark not defined.
PERNYATAAN	iv
UCAPAN TERIMA KASIH	v
ABSTRAK	vi
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian	4
1.4. Batasan Masalah	4
1.5. Manfaat Penelitian	4
1.6. Metodologi Penelitian	5
1.7. Sistematika Penulisan	6
BAB 2 LANDASAN TEORI	7
2.1. Aspect Based Sentiment Analysis	7
2.2. <i>Financial Technology (Fintech)</i>	7
2.3. Latent Dirichlet Allocation (LDA)	7
2.4. Recurrent Neural Network	8
2.5. Long Short Term Memory (LSTM)	9
2.6. Attention Mechanism	11
2.7. Penelitian Terdahulu	12
2.8. Perbedaan Penelitian	16
BAB 3 ANALISIS DAN PERANCANGAN SISTEM	17
3.1. Dataset	17
3.2. Arsitektur Umum	18
3.2.1. Preprocessing	19
3.2.2. Topic Building	30

3.2.3. Model Building	30
3.2.4. Classification	31
3.3. Perancangan Aplikasi Sistem	31
3.3.1. Rancangan Halaman Utama (Home)	31
3.3.2. Rancangan Halaman Training	32
3.3.3. Rancangan Halaman Testing	33
3.4. Metode Evaluasi	35
BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM	38
4.1. Implementasi Sistem	38
4.1.1. Spesifikasi perangkat keras dan perangkat lunak	38
4.1.2. Implementasi perancangan tampilan antarmuka	38
4.2. Implementasi Model	43
4.2.1. Pelatihan Model	43
4.3. Pengujian Model	44
4.4. Hasil Pengujian Model	46
4.4.1. Pengujian Pemodelan Topik	46
4.4.2. Pengujian Model Prediksi	50
4.5. Evaluasi Model	51
BAB 5 KESIMPULAN DAN SARAN	58
5.1. Kesimpulan	58
5.2. Saran	58
DAFTAR PUSTAKA	59

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	14
Tabel 3.1 Hasil scrapping ulasan aplikasi	17
Tabel 3.2 Contoh Data Cleaning	21
Tabel 3.3 Contoh Case Folding	21
Tabel 3.4 Contoh Punctual Removal	22
Tabel 3.5 Contoh Normalization	24
Tabel 3.6 Contoh Stopword Removal	25
Tabel 3.7 Contoh Stemming	26
Tabel 3.8 Contoh Tokenizing	27
Tabel 3.9 Confussion Matrix	36
Tabel 4.1 Hyperparameter Tunning Akurasi	43
Tabel 4.2 Hyperparameter Tunning Loss	44
Tabel 4.3 Data Hasil Pelatihan Model	45
Tabel 4.4 Keyword topik LDA	49
Tabel 4.5 Tabel Pengujian Model	50
Tabel 4.6 Nilai Variabel Sentimen Negatif Aspek Kemudahan	52
Tabel 4.7 Nilai Variabel Sentimen Netral Aspek Kemudahan	52
Tabel 4.8 Nilai Variabel Sentimen Positif Aspek Kemudahan	53
Tabel 4.9 Nilai Variabel Sentimen Negatif Aspek Layanan Pelanggan	53
Tabel 4.10 Nilai Variabel Sentimen Netral Aspek Layanan Pelanggan	54
Tabel 4.11 Nilai Variabel Sentimen Positif Aspek Layanan Pelanggan	54
Tabel 4.12 Nilai Variabel Sentimen Negatif Aspek Kualitas	55
Tabel 4.13 Nilai Variabel Sentimen Netral Aspek Kualitas	55
Tabel 4.14 Nilai Variabel Sentimen Positif Aspek Kualitas	55
Tabel 4.15 Nilai Variabel Sentimen Negatif Aspek Kecepatan	56
Tabel 4.16 Nilai Variabel Sentimen Netral Aspek Kecepatan	56
Tabel 4.17 Nilai Variabel Sentimen Positif Aspek Kecepatan	56

DAFTAR GAMBAR

Gambar 2.1 Representasi Latent Dirichlet Allocation (Blei et al., 2003)	8
Gambar 2.2 Siklus perputaran informasi pada RNN (Olah, 2015)	9
Gambar 2.3 Garis horizontal cell state dalam LSTM (Olah, 2015)	9
Gambar 2.4 Arsitektur LSTM dengan Attention (Yu et al., 2018)	12
Gambar 3.1 Arsitektur Umum	19
Gambar 3.8 Embedding Matrix	29
Gambar 3.10 Halaman Utama (Home)	32
Gambar 3.11 Userflow diagram Halaman Home	32
Gambar 3.12 Rancangan Halaman Training	33
Gambar 3.13 Userflow diagram Halaman Training	33
Gambar 3.14 Rancangan Halaman Testing	34
Gambar 3.15 Userflow diagram Halaman Testing	34
Gambar 3.16 Rancangan Halaman Uji Coba	35
Gambar 3.17 Userflow diagram Halaman Testing User Input	35
Gambar 4.1 Tampilan Halaman Utama	39
Gambar 4.2 Tampilan Halaman Training	39
Gambar 4.3 Tampilan Hasil Preprocessing	40
Gambar 4.4 Tampilan Hasil Grafik per aspek	40
Gambar 4.5 Tampilan Halaman Testing	41
Gambar 4.6. Tampilan Hasil Prediksi	41
Gambar 4.7. Tampilan Hasil Classification Report dan Confusion Matrix	42
Gambar 4.9 Tampilan Halaman Input Ulasan	42
Gambar 4.10 Tampilan Hasil Prediksi	43
Gambar 4.11 Visualisasi LDA Topik 1	47
Gambar 4.12 Visualisasi LDA Topik 2	48
Gambar 4.13 Visualisasi LDA Topik 3	48
Gambar 4.14 Visualisasi LDA Topik 4	49
Gambar 4.15 Confussion Matrix aspek kemudahan	52
Gambar 4.16 Confussion Matrix aspek Layanan Pelanggan	53
Gambar 4.17 Confussion Matrix aspek Kualitas	54

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Financial Technology (Fintech) merupakan serangkaian layanan inovatif yang didukung oleh kemajuan teknologi informasi dan komunikasi. *Fintech* mencakup perusahaan-perusahaan yang menyediakan layanan keuangan dengan pendekatan berbasis teknologi. Istilah *fintech* berasal dari penggabungan dua konsep, yaitu "*financial*" dan "*technology*" yang membentuk istilah "*financial technology*" (Abad-Segura et al., 2020). Kehadiran *financial technology* bertujuan untuk memberikan akses global terhadap produk finansial, memudahkan bisnis serta memajukan literasi keuangan (Rofiah & Setiyadi, 2020).

Di Indonesia, sektor *Fintech* telah mengalami pertumbuhan sejak tahun 2018, dipicu oleh minat yang tinggi dari masyarakat Indonesia dalam memanfaatkan layanan *Fintech*. Sejumlah perusahaan *Fintech* kemudian muncul dan bersaing untuk menyediakan layanan terbaik bagi masyarakat Indonesia (Safitri, 2020). *Digital Payment System* merupakan salah satu jenis *Fintech* yang berkembang di Indonesia. Link Aja merupakan salah satu jenis layanan *Digital Payment Sistem* yang diresmikan pada tahun 2019 dan berfungsi sebagai penyedia layanan untuk pembayaran semua tagihan, seperti pulsa dan pascabayar, cicilan kredit, dan angsuran kesehatan seperti BPJS.

Berdasarkan laporan *E-Wallet Industry Outlook 2023* dari *Insight Asia*, dalam kelompok pengguna pembayaran digital didapatkan hasil bahwa Link Aja menempati posisi ke 5 sebagai aplikasi yang banyak digunakan dengan proporsi pengguna 27%. Terdapat jarak yang cukup besar antara LinkAja dengan platform lainnya seperti OVO dengan proporsi pengguna mencapai 70% dan pengguna DANA dengan persentase 61%. Berdasarkan Direktur Riset Center of Reforms on Economics (CORE) meskipun memiliki banyak keunggulan dan fitur, aplikasi LinkAja belum mampu untuk bersaing dengan aplikasi e-wallet lainnya dikarenakan kurangnya promosi yang dilakukan sehingga banyak masyarakat yang tidak mengenal aplikasi ini (Wilfan & Martini,

2021). Selain itu, aplikasi LinkAja juga sering mengalami masalah seperti layanan pelanggan yang lambat merespon keluhan, transaksi yang gagal, proses pengembalian dana yang memakan waktu, pembaruan aplikasi yang rentan mengakibatkan aplikasi eror, dan sebagainya (Chrisanta, 2023).

Berdasarkan Google Play Store per Oktober 2023, terdapat lebih dari 10 juta jumlah pengguna telah mengunduh aplikasi LinkAja dengan rating aplikasi sebesar 3,6 dan diulas oleh 786 ribu pengguna. Pengguna dapat memberikan ulasan mengenai aplikasi LinkAja melalui adanya fitur ulasan dan pemberian nilai dalam bentuk bintang dalam Play Store dan Apple Store. Namun, ulasan yang ada tidak memberikan informasi yang signifikan untuk pengambilan keputusan karena hanya akan menganalisis ulasan dalam level kalimat saja. Oleh karena itu, diperlukan suatu pendekatan analisis berbasis aspek untuk mengolah ulasan berdasarkan aspek dengan tujuan untuk meningkatkan suatu kinerja dan membuat inovasi lebih lanjut, sehingga layanan tersebut menjadi lebih unggul.

Penelitian yang sudah dilakukan sebelumnya, diantaranya adalah dari (Turjaman & Budi, 2022) yang melakukan penelitian untuk klasifikasi aspek marketing mix pada aplikasi LinkAja. Penelitian ini menggunakan string matching dan algoritma Support Vector Machine dengan data ulasan terkait aplikasi LinkAja di twitter dalam periode waktu Januari-Mei 2022. Penelitian ini menghasilkan nilai akurasi terbaik untuk aspek harga sebesar 80%, aspek produk 77,5%, aspek promosi 76,2%, dan aspek tempat 71,4%.

Selanjutnya (Fitriani et al., 2019) melakukan analisis sentimen untuk ulasan aplikasi LinkAja dari twitter dengan memanfaatkan tiga algoritma machine learning yaitu Decision Tree (C4.5), Naïve Bayes, dan K-NN (K-Nearest Neighbor). Pada penelitian ini menunjukkan bahwa algoritma K-NN lebih unggul dibandingkan algoritma Naïve Bayes dan Decision Tree (C4.5) yaitu dengan nilai akurasi 73.45% untuk algoritma K-NN, Naïve Bayes dengan nilai 60,55%, dan 68.49% untuk algoritma Decision Tree. Selain itu, untuk hasil analisa sentimen diperoleh kesimpulan bahwa pengguna twitter lebih condong merespon dengan komentar negatif yang berisikan komplain, keluhan, komentar untuk menuntut, dan rasa ketidaknyamanan terhadap layanan aplikasi LinkAja.

Berikutnya, (Khaffi, 2023) melakukan penelitian dengan algoritma Extreme Gradient Boosting untuk menganalisis sentimen berbasis aspek berdasarkan ulasan aplikasi LinkAja pada google play store. Penelitian ini memperoleh akurasi 90% dengan

2000 data ulasan. Aspek yang diteliti yaitu aspek *customer service*, *user experience*, fitur aplikasi, dan verifikasi.

Penggunaan Deep learning mampu untuk menangkap fitur semantik dan sintaksis dari text tanpa perlu menggunakan fitur lain seperti pada penelitian sebelumnya mengenai klasifikasi aspek. RNN merupakan algoritma deep learning yang performanya lebih baik daripada CNN dalam proses klasifikasi jika berhubungan dengan semantik kalimat (Do et al., 2019). Penerapan LSTM yang merupakan algoritma perbaikan dari RNN telah pernah digunakan bersama dengan LDA dalam pengukuran sentiment pengguna berdasarkan ulasan mobil. Pertama, model LDA digunakan untuk mengekstrak topik informasi komentar. Kedua, pemetaan topik dilakukan terhadap komentar untuk mendapatkan korelasi antara setiap informasi dengan setiap topik dimensi. Terakhir, model LSTM dibangun untuk menganalisa tingkatan emosional konsumen pada setiap dimensi dengan nilai akurasi untuk model training 95%, dan akurasi model testing 90% (Lv et al., 2021)

Berdasarkan latar belakang serta penelitian-penelitian sebelumnya, maka penulis mengajukan penelitian dengan judul “Analisis sentimen berbasis aspek Pada Aplikasi LinkAja Berdasarkan Ulasan Pengguna Dengan Menggunakan Latent Dirichlet Allocation Dan Attention Based - Long Short Term Memory “.

1.2. Rumusan Masalah

Digital Payment System merupakan salah satu jenis sistem pembayaran digital yang banyak digunakan di Indonesia. LinkAja sebagai salah satu aplikasi yang digunakan oleh masyarakat, dapat diakses melalui Play Store dan juga Apple Store. Pengguna dapat memberikan ulasan mengenai aplikasi LinkAja melalui adanya fitur ulasan dan pemberian nilai dalam bentuk bintang. Namun, ulasan yang ada tidak memberikan informasi yang signifikan untuk pengambilan keputusan karena hanya akan menganalisis ulasan dalam level kalimat saja. Oleh karena itu, diperlukan suatu pendekatan analisis berbasis aspek untuk mengolah ulasan berdasarkan aspek dengan tujuan untuk meningkatkan suatu kinerja dan membuat inovasi lebih lanjut, sehingga layanan tersebut menjadi lebih unggul.

1.3. Tujuan Penelitian

Tujuan dari penelitian ini untuk mengidentifikasi aspek pengalaman pengguna dalam menggunakan aplikasi LinkAja berdasarkan ulasan pengguna dengan pemodelan topik menggunakan Latent Dirichlet Allocation Dan analisis polaritas sentimen dengan menggunakan algoritma Attention Based - Long Short Term Memory

1.4. Batasan Masalah

Dalam penelitian ini, diterapkan batasan untuk menghindari penyimpangan dan perluasan cakupan masalah sehingga penelitian fokus terhadap pokok permasalahan. Maka batasan masalah yang diterapkan dalam penelitian ini yaitu :

1. Penelitian ini menggunakan data ulasan pengguna LinkAja dari Apple App Store dan Google Play Store dari tahun 2022-2024.
2. Teknik pengambilan data dilakukan dengan scrapping dari website Apple App Store dan Google Play Store.
3. Data ulasan yang digunakan dalam pembangunan model adalah Bahasa Indonesia.
4. Aspek yang diuji meliputi kemudahan, layanan pengguna, kualitas, dan kecepatan layanan yang diperoleh melalui pemodelan topik.
5. Output penelitian ini adalah sistem berbasis web yang memberi kesimpulan terhadap ulasan pengguna aplikasi termasuk positif, negatif, atau netral berdasarkan aspek yang telah ditentukan.

1.5. Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian ini adalah :

1. Membantu perusahaan mengetahui opini dari pengguna untuk meningkatkan performa layanan.
2. Mempelajari kemampuan algoritma Latent Dirichlet Allocation dan Attention Based - LSTM (Long Short Term Memory) dan menilai akurasi yang diperoleh dalam menentukan pengalaman pengguna dan analisa sentiment terhadap ulasan pengguna.
3. Menjadi referensi penelitian-penelitian berikutnya dalam melakukan pemodelan topik menggunakan Latent Dirichlet Allocation dan analisa polaritas sentiment menggunakan algoritma Attention Based-Long Short Term Memory.

1.6. Metodologi Penelitian

Rangkaian tahapan yang dilakukan dalam penelitian ini adalah:

1. Studi Literatur

Pada tahap ini, dikumpulkan informasi mengenai penerapan metode Latent Dirichlet Allocation (LDA), Attention Mechanism, dan Long Short Term Memory (LSTM) dari berbagai sumber seperti artikel, jurnal, skripsi, internet, dan sumber informasi lainnya.

2. Analisis Permasalahan

Setelah tahap studi literatur, langkah berikutnya adalah analisis permasalahan. Pada tahap ini, dilakukan pemahaman konsep mengenai penerapan metode LDA untuk pemodelan topik dan metode Attention Based-Long Short Term Memory untuk analisis polaritas sentimen..

3. Perancangan

Setelah analisis permasalahan, tahap selanjutnya adalah proses perancangan, yang meliputi desain arsitektur umum, pengumpulan data, penentuan data pelatihan dan data pengujian, serta desain tampilan antarmuka sistem.

4. Implementasi

Pada langkah ini, rancangan yang telah dibuat diimplementasikan dengan membangun sebuah sistem yang sesuai dengan analisa sebelumnya.

5. Pengujian Sistem

Tahap selanjutnya ialah tahap pengujian dari performa sistem yang telah dirancang untuk mendapatkan hasil akurasi dari pengimplementasian metode Latent Dirichlet Allocation (LDA) dan Attention based-Long Short Term Memory pada penelitian ini.

6. Penyusunan Laporan

Tahap ini ialah tahap penyusunan laporan setelah melewati tahap pengujian sistem. Peneliti melakukan penyusunan laporan sebagai dokumentasi penelitian secara keseluruhan dan sebagai hasil akhir dari penelitian.

1.7. Sistematika Penulisan

Sistematika penulisan dalam skripsi ini dibagi menjadi lima bagian yang akan dijelaskan sebagai berikut:

BAB 1: PENDAHULUAN

Pada bab ini menjelaskan tentang latar belakang penelitian, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, metodologi penelitian, dan sistematika penulisan.

BAB 2: LANDASAN TEORI

Bab ini mencakup penjelasan teori-teori yang relevan dengan penelitian ini, termasuk definisi *financial technology*, prinsip kerja algoritma Latent Dirichlet Allocation dan Attention Based - LSTM, metodologi penelitian, serta aspek-aspek lain yang terkait dengan penelitian ini.

BAB 3: ANALISIS DAN PERANCANGAN SISTEM

Bab ini menguraikan analisis dan implementasi algoritma Latent Dirichlet Allocation dalam mengidentifikasi aspek dalam teks serta algoritma Attention Based LSTM dalam menganalisis sentimen berdasarkan aspek dari ulasan pengguna. Juga dibahas arsitektur umum yang dirancang untuk implementasi kedua algoritma tersebut.

BAB 4: IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini menjelaskan tentang penerapan desain yang telah direncanakan sebelumnya. Pada bab ini juga terdapat penjelasan hasil pengujian sistem yang telah dibuat untuk mengevaluasi kinerja model.

BAB 5: KESIMPULAN DAN SARAN

Pada bab ini menyajikan kesimpulan dari penelitian yang telah dilakukan serta memberikan saran untuk pengembangan lebih lanjut dalam penelitian mendatang.

BAB 2

LANDASAN TEORI

2.1. Aspect Based Sentiment Analysis

Analisis sentimen berbasis aspek (ABSA) membantu memahami masalah sentimen analisis dengan lebih baik, karena berfokus langsung pada sentimen daripada struktur Bahasa (Nazir et al., 2022). Aspek adalah topik pembahasan, yaitu subjek dari kalimat tersebut. Pada analisis sentimen berbasis aspek, proses klasifikasi sentimen berfokus pada setiap aspek. Sebelum dianalisis secara sentimen, aspek-aspek tersebut pertama-tama diekstraksi dari kalimat. Proses ekstraksi aspek memperhatikan fitur-fitur dari frasa dalam kalimat terkait (Kang & Zhou, 2017).

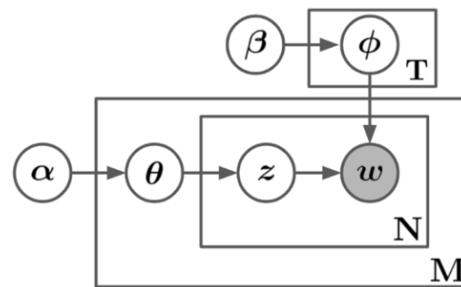
2.2. Financial Technology (Fintech)

Fintech berasal dari istilah financial technology atau diartikan sebagai teknologi finansial. Menurut The National Digital research Centre (NDRC), di Dublin, Irlandia mendefinisikan fintech sebagai inovasi dalam sektor finansial yang mendapat sentuhan teknologi modern (Rizal et al., 2018). Dalam hal ini, fintech memanfaatkan teknologi dalam memberikan pelayanan keuangan. Fintech merubah model bisnis atau bentuk layanan keuangan menjadi lebih modern, yang pada awalnya melakukan transaksi pembayaran harus dengan tatap muka langsung dan membawa uang tunai, kini bisa dilakukan dimanapun dan kapanpun tanpa adanya batasan waktu dan jarak. Dengan demikian, fintech dapat memberikan manfaat bagi para pelaku usaha maupun konsumen.

2.3. Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation merupakan algoritma yang biasa digunakan untuk pemodelan topik. Model LDA telah diakui dan digunakan dalam berbagai bidang penelitian. Algoritma ini telah banyak digunakan karena kemampuannya untuk menghasilkan suatu topik yang secara komprehensif menggambarkan korpus tertentu. Model ini mampu menangani kumpulan data berskala besar tanpa harus melakukan pelabelan teks secara manual. Setiap topik yang dihasilkan dapat ditafsirkan secara

mandiri, sehingga data yang dihasilkan dapat berguna dalam berbagai domain (Albalawi et al., 2020). Ide dasar dari metode LDA yaitu setiap dokumen direpresentasikan sebagai campuran acak atas topik yang tersembunyi, dimana setiap topik memiliki karakter yang ditentukan berdasarkan distribusi kata-kata yang terdapat didalamnya (Putra & Kusumawardani, 2017). Representasi dari model LDA dapat dilihat pada gambar dibawah ini:



Gambar 2.1 Representasi Latent Dirichlet Allocation (Blei et al., 2003)

Keterangan :

α = distribusi topik pada keseluruhan dokumen

β = distribusi kata pada keseluruhan topik

θ = distribusi topik pada keseluruhan dokumen M

ϕ = distribusi kata pada keseluruhan dokumen T

T = kumpulan topik

w = kata spesifik

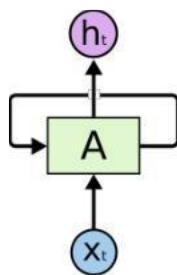
z = topik pada kata ke- N dalam dokumen M

N = kumpulan kata

M = kumpulan dokumen

2.4. Recurrent Neural Network

Recurrent Neural Network adalah jenis algoritma yang pemrosesannya dipanggil secara berulang untuk memproses masukan yang biasanya adalah data sekuensial. RNN termasuk kedalam kategori deep learning karena datanya diproses melalui banyak lapisan (layer) (Subagja, 2019). Siklus perulangan pada RNN digambarkan pada gambar



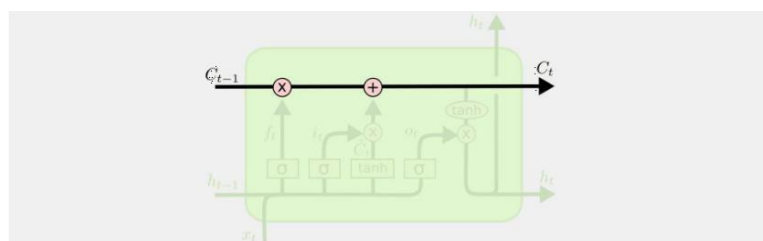
Gambar 2.2 Siklus perputaran informasi pada RNN (Olah, 2015)

Berdasarkan gambar, x_t adalah input dan h_t adalah output. Dan terdapat mekanisme loop yang menangkap informasi dari satu tahap ke tahapan berikutnya. RNN juga mampu menangani ketergantungan jangka panjang. Namun, dalam pengaplikasiannya RNN tidak dapat menyimpan informasi sebelumnya dengan baik dalam waktu yang lama karena masalah gradien menghilang. Untuk mengatasi kelemahan ini, maka diusulkan suatu algoritma perbaikan dari RNN yang disebut dengan LSTM (Tian et al., 2018)

2.5. Long Short Term Memory (LSTM)

Long Short Term Memory (LSTM) merupakan algoritma yang diusulkan pada tahun 1997 oleh Sepp Hochreiter dan Jurgen Schmidhuber. Algoritma ini muncul sebagai pengembangan dari Recurrent Neural Network (RNN) yang mampu mengatasi kekurangan RNN yaitu pengolahan informasi dalam waktu yang lama (Zahara et al., 2019).

Komponen dasar LSTM adalah cell state. Sebuah baris yang berjalan dari memory cell sebelumnya (C_{t-1}) ke memory cell saat ini (C_t), seperti yang ditunjukkan pada Gambar 2.3



Gambar 2.3 Garis horizontal cell state dalam LSTM (Olah, 2015)

Memory cell adalah garis horizontal yang menghubungkan semua output layer pada LSTM. Dengan adanya jalur tersebut, suatu nilai memory cell yang lama akan dengan mudah diteruskan ke memory cell yang baru dengan sedikit sekali modifikasi.

Ini memungkinkan informasi mengalir lurus ke bawah. Jaringan dapat menentukan jumlah informasi sebelumnya mengalir. Itu dikendalikan melalui lapisan pertama (σ_1) (Kumar et al., 2018).

LSTM mempunyai 3 jenis gates diantaranya yaitu input gate, output gate, dan forget gate. Forget gate merupakan gerbang yang menetapkan apakah informasi tersebut digunakan atau dihilangkan dari sel. Input gate merupakan gerbang yang menetapkan untuk memperbarui nilai dari input dalam state memory. Output gate merupakan gerbang yang menetapkan apakah keluaran identik dengan masukan dan memori dalam sel.

Formula forget gate dijabarkan pada persamaan 2.1

$$ft = \sigma(Wf \cdot [ht-1, xt] + bf) \quad (2.1)$$

Dengan keterangan :

ft = forget gate

σ = sigmoid function

Wf = weights for forget gate

$ht-1$ = nilai hidden state sebelum langkah t

xt = input saat langkah t

bf = forget gate bias

Formula mencari input gate dijabarkan pada persamaan 2.2.

$$it = \sigma(Wi \cdot [ht-1, xt] + bi) \quad (2.2)$$

Dengan keterangan :

it = input gate

σ = sigmoid function

W_i = weights for input gate

h_{t-1} = nilai hidden state sebelum langkah t

x_t = input saat langkah t

b_i = input gate bias

Formula mencari output gate dijabarkan pada persamaan 2.3

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2.3)$$

Dengan keterangan :

o_t = output gate

σ = sigmoid function

W_o = weights for output gate

h_{t-1} = nilai hidden layer sebelum langkah t

x_t = input pada langkah t

b_o = output gate bias

2.6. Attention Mechanism

Attention Mechanism adalah algoritma yang memfokuskan setiap output agar kata yang dihasilkan memiliki hubungan yang lebih baik sesuai konteks kata (Bahdanau et al., 2015). *Attention mechanism* telah terbukti mampu mendapatkan hasil yang baik dalam memilih informasi yang penting. Saat ini, attention mechanism telah menunjukkan hasil yang baik dalam bidang NLP, dengan meningkatkan kinerja algoritma deep neural network (DNN) dan membiarkan DNN belajar sendiri untuk fokus pada konteks (Yadav et al., 2021).

Konsep dasar *attention mechanism* adalah banyaknya decoder yang berfokus pada hidden state tertentu dari encoder agar memperoleh output saat waktu tertentu sehingga dapat membantu memperkuat model dalam memproses data berurutan (Nugraha, 2021). Penelitian ini menggunakan *attention mechanism* untuk tugas klasifikasi. Misalkan H adalah matriks yang terdiri dari vektor output $[h_1, h_2, \dots, h_T]$ LSTM yang dihasilkan, dimana T adalah panjang kalimat. Representasi r dari kalimat tersebut dibentuk oleh

jumlah tertimbang dari vektor output tersebut. Persamaan perhitungan attention mechanism dalam tugas klasifikasi diuraikan pada persamaan 2.4, 2.5, dan 2.6.

$$M = \tanh(H) \quad (2.4)$$

$$\alpha = \text{softmax}(w^T M) \quad (2.5)$$

$$r = H\alpha^T \quad (2.6)$$

Dengan keterangan :

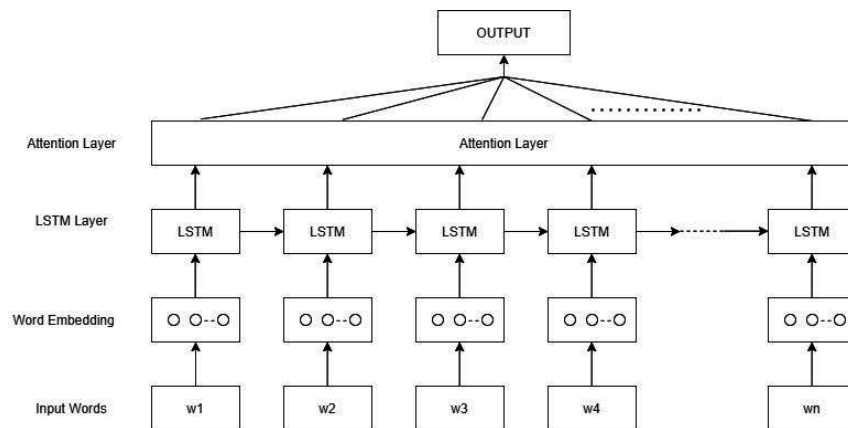
H = Output vektor

w = Trained parameter vektor

w^T = Transpose

Kemudian didapatkan representasi pasangan kalimat terakhir yang digunakan pada klasifikasi dari

$$h^* = \tanh(r) \quad (2.7)$$



Gambar 2.4 Arsitektur LSTM dengan Attention (Yu et al., 2018)

2.7. Penelitian Terdahulu

Dalam penelitian, disarankan untuk melakukan peninjauan literatur dari sejumlah penelitian sebelumnya. Hal ini bertujuan untuk memahami keterkaitan antara penelitian yang sedang dilakukan dengan kajian-kajian sebelumnya. Dengan adanya referensi tersebut, duplikasi dapat dihindari dalam penelitian yang bersangkutan. Penelitian

mengenai ulasan pengguna LinkAja telah pernah dilakukan. (Turjaman & Budi, 2022) yang melakukan penelitian untuk klasifikasi aspek marketing mix pada aplikasi LinkAja. Penelitian ini menggunakan string matching dan algoritma Support Vector Machine dengan data ulasan terkait aplikasi LinkAja di Twitter dalam periode waktu Januari-Mei 2022. Penelitian ini menghasilkan nilai akurasi terbaik untuk aspek harga dengan skor 80,0%. disusul oleh aspek produk 77,5%, aspek promosi 76,2%, dan aspek tempat 71,4%. Selanjutnya (Fitriani et al., 2019) melakukan analisis sentimen untuk ulasan aplikasi LinkAja dari twitter dengan menggunakan tiga algoritma machine learning yaitu Naïve Bayes, Decision Tree (C4.5), dan K-NN (K-Nearest Neighbor). Penelitian ini menunjukkan bahwa algoritma K-NN lebih unggul dibandingkan algoritma Naïve Bayes dan Decision Tree (C4.5) yaitu dengan nilai akurasi 73.45% untuk algoritma K-NN, Naïve Bayes dengan nilai 60,55%, dan 68.49% untuk algoritma Decision Tree. Selain itu untuk hasil sentiment pengguna diperoleh hasil bahwa pengguna twitter cenderung merespon dengan komentar negatif yang berisi keluhan, komplain, komentar menuntut, dan ketidaknyamanan pengguna terhadap produk LinkAja.

Dengan menggunakan algoritma Naïve Bayes dan TF-IDF, (Helmayanti et al., 2023) melakukan penelitian analisis sentiment berbasis aspek ulasan aplikasi Flip pada google play store. Penelitian ini memperoleh akurasi 84% dengan hasil sentiment netral paling banyak, diikuti dengan sentiment positif dan negatif. Sentiment negatif paling banyak untuk aspek kecepatan dan keamanan, dan sentiment positif paling banyak untuk aspek biaya yang menjelaskan bahwa aplikasi Flip perlu untuk meningkatkan aspek keamanan dan kecepatannya. Kemudian, (Fattahila et al., 2021) melakukan analisis sentiment berbasis aspek untuk ulasan pengguna aplikasi DANA, OVO, LinkAja, dan Sakuku melalui Google Play Store dengan menggunakan algoritma CNN dan LSTM. Ulasan yang diperoleh akan diklasifikasikan berdasarkan 5 aspek, yaitu transaksi, akun, layanan, aksesibilitas dan performa. Hasil penelitian menunjukkan akurasi untuk training sebesar 87% dan untuk validasi sebesar 83%. Berdasarkan data dan analisis dari keempat aplikasi tersebut, sentimen yang banyak diperoleh adalah negatif untuk masing-masing aspek yang telah dibuat. Dengan menggunakan gabungan metode Latent Dirichlet Allocation dan Long Short Term Memory, (Lv et al., 2021) melakukan analisa sentimen pengguna untuk ulasan mobil. Pertama, model LDA digunakan untuk mengekstrak topik pada komentar. Kedua, pemetaan topik dilakukan

terhadap komentar untuk mendapatkan korelasi antara setiap informasi dengan setiap topik dimensi. Terakhir, model LSTM dibangun untuk menganalisa tingkatan emosional konsumen pada setiap aspek yang diperoleh dengan nilai akurasi untuk model training 95%, dan akurasi model testing 90%. Selanjutnya, (Priyantina & Sarno, 2019) melakukan penelitian analisis sentimen berbasis aspek untuk ulasan hotel dengan menggunakan LDA dan LSTM. Dengan pemetaan topik LDA, didapatkan 5 aspek yaitu lokasi, layanan, kebersihan, kenyamanan, dan makanan. Kemudian klasifikasi sentimen pelanggan (puas atau tidak puas) dilakukan dengan menggunakan kombinasi Word Embedding dan Long-short Term Memory (LSTM) dan memperoleh akurasi sebesar 93%.

Tabel 2.1 Penelitian Terdahulu

No.	Penulis	Judul	Tahun	Keterangan
1.	Rifqy Mikoriza Turjaman dan Indra Budi	Analisis Sentimen Berbasis Aspek Marketing Mix Terhadap Ulasan Aplikasi Dompot Digital (Studi Kasus: Aplikasi Linkaja Pada Twitter)	2022	Penelitian ini melakukan klasifikasi aspek <i>marketing mix</i> dengan menggunakan <i>string matching</i> dan algoritma Support Vector Machine. Hasil akurasi untuk aspek harga dengan skor 80,0%. disusul oleh aspek produk 77,5%, aspek promosi 76,2%, dan aspek tempat 71,4%.
2.	Ayu Fitriani, Rachmadita Andreswari, dan Irfan Darmawan	Analisis Sentimen Pengguna Dompot Digital Pada Media Sosial Twitter Menggunakan Perbandingan Algoritma Klasifikasi (Studi Kasus Linkaja)	2019	Penelitian ini melakukan analisis sentimen untuk ulasan aplikasi LinkAja dengan menggunakan tiga algoritma yaitu Naïve Bayes, Decision Tree (C4.5), dan K-NN (K-Nearest Neighbor). Dengan akurasi 73.45% untuk algoritma K-NN, Naïve Bayes dengan nilai 60,55%, dan 68.49% untuk algoritma Decision Tree.

Tabel 2.2 Penelitian Terdahulu (Lanjutan)

3.	Sheva Aditya Helmayanti, Faqih Hamami, dan Riska Yanu Fa'rifah	Penerapan Algoritma TF-IDF Dan Naïve Bayes Untuk Analisis Sentimen Berbasis Aspek Ulasan Aplikasi Flip Pada Google Play Store	2023	Penelitian ini melakukan analisis sentimen berbasis aspek untuk ulasan Flip dengan algoritma Naïve Bayes dan TF-IDF. Penelitian ini memperoleh akurasi 84% dengan hasil sentiment netral paling banyak, diikuti dengan sentiment positif dan negatif.
4.	Ananda Affan Fattahila, Fendi Irfan Amorokhman, Kaenova Mahendra Auditama, Kurniadi Ahmad Wijaya, dan Ade Romadhony	Indonesian Digital Wallet Sentiment Analysis Using CNN And LSTM Method	2021	Penelitian ini melakukan analisis sentimen berbasis aspek untuk ulasan aplikasi DANA, OVO, LinkAja, dan Sakuku dan melakukan klasifikasi ulasan kedalam lima aspek yaitu transaksi, akun, layanan, aksesibilitas dan performa. Hasil penelitian menunjukkan akurasi untuk training sebesar 87% dan untuk validasi sebesar 83%.
5.	Yi-hang Lv, Qin Liu, dan Chang-bin Jiang	Sentiment analysis measure of topic mapping in LDA-LSTM combination model	2021	Penelitian ini menggunakan algoritma LDA untuk mengekstrak topik dari ulasan pengguna dan LSTM untuk klasifikasi sentiment pengguna. Penelitian ini memperoleh akurasi untuk model training 95%, dan akurasi model testing 90%.

Tabel 2.3 Penelitian Terdahulu (Lanjutan)

6.	Reza Amalia Priyantina dan Riyanarto Sarno	Sentiment Analysis of Hotel Reviews Using Latent Dirichlet Allocation, Semantic Similarity and LSTM	2019	Penelitian ini menggunakan algoritma LDA untuk mendapatkan aspek berdasarkan ulasan pengguna hotel dan algoritma LSTM untuk klasifikasi ulasan dengan akurasi 93%.
----	--------------------------------------------	-----------------------------------------------------------------------------------------------------	------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------

2.8. Perbedaan Penelitian

Terdapat perbedaan pada penelitian ini dengan penelitian-penelitian terdahulu. Pada penelitian ini dilakukan identifikasi aspek pengalaman pengguna pada ulasan aplikasi LinkAja dengan data ulasan dari Google Play Store dan Apple App Store. Algoritma yang digunakan pada penelitian ini adalah Latent Dirichlet Allocation untuk memperoleh aspek dan Attention based Long Short Term Memory untuk klasifikasi sentiment ulasan pengguna. Perbedaan penelitian dengan yang dilakukan oleh (Turjaman & Budi, 2022) adalah pada penelitian tersebut dilakukan analisa sentimen berdasarkan aspek Marketing Mix pada ulasan aplikasi LinkAja dengan menggunakan algoritma Support Vector Machine. Kemudian, perbedaan penelitian dengan (Fitriani et al., 2019) adalah pada penelitian tersebut menggunakan data ulasan pengguna aplikasi LinkAja dari twitter dan melakukan analisa sentimen ulasan dengan membandingkan tiga algoritma yaitu Naïve Bayes, Decision Tree (C4.5), dan K-NN (K-Nearest Neighbor). Dan perbedaan penelitian dengan (Fattahila et al., 2021) adalah penelitian tersebut melakukan analisis sentimen berbasis aspek untuk empat aplikasi digital wallet yaitu DANA, OVO, LinkAja, dan Sakuku. Untuk aspek yang diperoleh dalam penelitian tersebut adalah transaksi, akun, layanan, aksesibilitas dan performa dengan analisa sentimen menggunakan algoritma CNN-LSTM.

BAB 3

ANALISIS DAN PERANCANGAN SISTEM

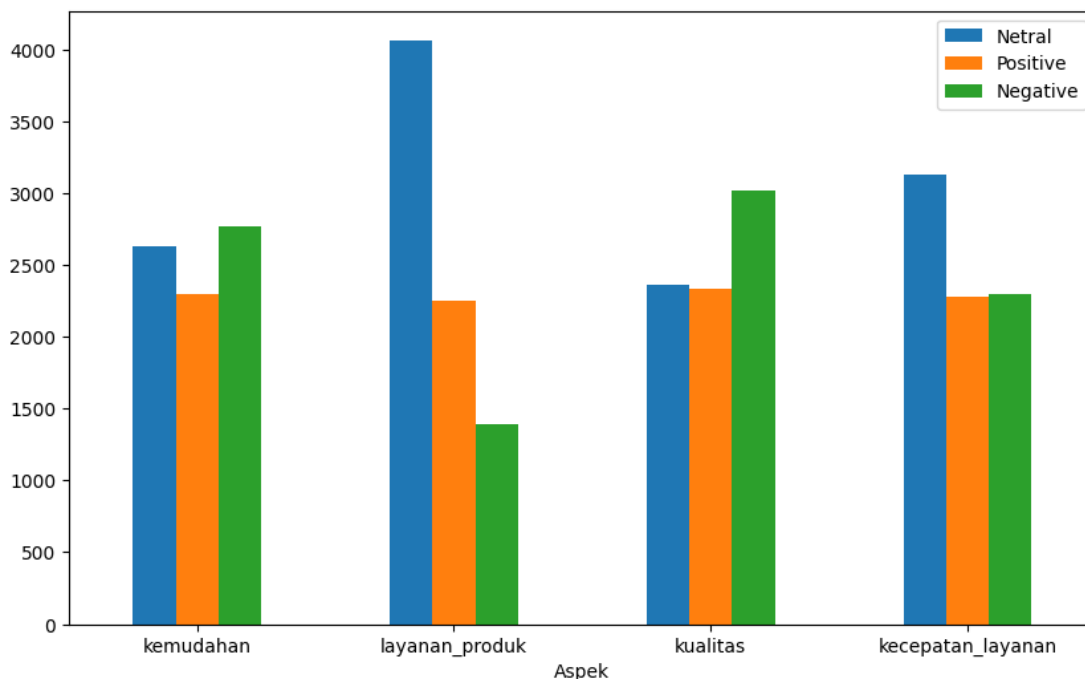
3.1. Dataset

Data dalam penelitian ini diperoleh melalui ulasan pengguna aplikasi LinkAja pada Google Play Store dan Apple App Store dengan teknik scrapping dalam format csv. Data yang discrapping merupakan data ulasan sepanjang tahun 2022-2024 dengan total 8779. Di bawah ini dalam tabel 3.1 merupakan ulasan pengguna aplikasi.

Tabel 3.1 Hasil scrapping ulasan aplikasi

No	Ulasan
1.	Aplikasinya sering error, akun sudah lama ngga di buka eh langsung terblokir dan tidak bisa dipulihkan padahal saldo masih lumayan banyak. Rata-rata konsumen sudah banyak yang komplain tapi kok ngga ada tindak lanjut ya, emang aneh ni aplikasi
2.	Alpikasi mbok, e ancok. Ws kadong tranfer uang mau melakukan transaksi selalu tidak di proses. Udah hubungi live chat sama agen. Aduh parah parah sangat merugikan bagi pengguna.aplikasi bener bener bosok ni. Ada tulisa menu klik.di.live chat tapi gk bisa di klik.mohon tanggapannya sesegera mungkin. Trims
3.	Tidak bisa reset PIN, ketika mau hapus akun dan mau buat akun baru kasih keterangan akun lama masih aktif, tapi ga bisa reset PIN. Aplikasi apa ini sih. Udah di chat langsung di website resmi, ga ada solusi. Call center nya ga bisa dihubungi. Kocak kocak
4.	Mantap sampai saat ini tidak ada kendala dan sangat mudah digunakan untuk transaksi besar atau kecil.. sekedar saran kalo bisa bikin fitur yang untuk menghapus riwayat transaksi.. terimakasih
5.	Tolong kembalikan saldo saya...topup etoll saldo sudah kepotong tapi tidak masuk setelah di update kartunya....tolong ini perusahaan besar jangan maling...tlp call center tidak ada solusi

Data ulasan yang sudah dikumpulkan dibagi dengan rasio 90:10 sehingga jumlahnya menjadi 8779 untuk data latih dan 439 untuk data uji. Kemudian dari data latih dibagi lagi untuk validasi sehingga menghasilkan 439 data. Berikut merupakan grafik persebaran jumlah sentimen dalam setiap aspek.

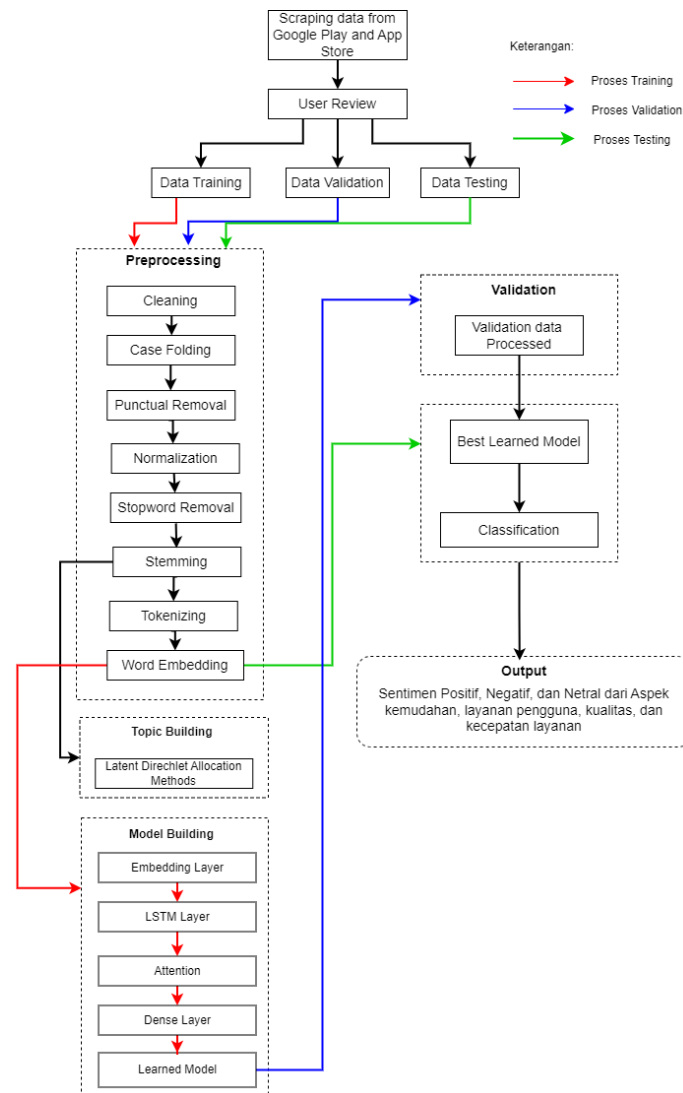


Gambar 3.1 Grafik persebaran sentimen per aspek

3.2. Arsitektur Umum

Terdapat beberapa tahapan yang dilakukan dalam penelitian ini yaitu pengumpulan data ulasan dengan scrapping yang bersumber dari Google Play Store dan Apple App Store kemudian disimpan kedalam file dengan ekstensi csv. Kemudian data yang sudah diperoleh dibagi menjadi data training dan data testing dengan rasio pembagian 90 : 10. Kemudian data training akan dibagi lagi menjadi dua yaitu data training utama dan data validasi dengan rasio pembagian 90 : 10. Selanjutnya dilakukan preprocessing data dimana dalam tahap ini terdapat proses seperti case folding, stopwords removal, stemming, dan tokenizing. Selanjutnya hasil dari tokenizing akan memasuki tahap word embedding untuk mengubah token tersebut menjadi vektor. Selanjutnya vektor tersebut akan menjadi input dalam pembangunan model klasifikasi. Berikutnya merupakan tahap untuk mencari aspek menggunakan metode Latent Dirichlet Allocation. Kemudian dibangun sebuah model klasifikasi dengan algoritma Long Short Term Memory dengan Attention Based berdasarkan parameter-parameter yang sudah ditentukan melalui proses tuning parameter. Setelah model dibangun maka model ditraining dengan data training yang sudah disiapkan sebelumnya. Model yang sudah melalui proses training juga akan melakukan proses validasi dengan data validasi yang

bertujuan untuk mengevaluasi apakah model yang dibangun sudah mencapai performa yang baik selama proses training. Hasil akhir dari proses training ini adalah sebuah model yang mampu melakukan klasifikasi sentimen ulasan pengguna aplikasi berdasarkan aspek yang sudah diperoleh. Sistem ini akan diuji dengan data testing yang sudah disiapkan untuk menampilkan hasil klasifikasinya. Untuk lebih jelasnya dapat dilihat pada gambar 3.2.



Gambar 3.2 Arsitektur Umum

3.2.1. Preprocessing

Preprocessing merupakan tahapan yang dilakukan dengan tujuan untuk mengubah data mentah untuk diolah menjadi data yang lebih terstruktur Adapun tahapan tersebut yaitu:

3.2.1.1. Data Cleaning

Data Cleaning adalah proses pembersihan atribut yang tidak berhubungan dengan informasi yang ada pada data seperti karakter, simbol dan emoticon sehingga data yang dihasilkan menjadi lebih bersih. Proses data cleaning digambarkan pada *pseudocode* dibawah ini

```
FUNCTION cleaning(data):
    # Step 1: mengganti karakter '\r' dengan spasi
    data['ulasan_1']<-data['content'].replace("\\\\r", " ")
    # Step 2: menghapus emoji
    data['ulasan_1'] <- APPLY function(x):
        x <- x.encode('ascii', 'ignore')
        x <- x.decode('ascii')
        RETURN x
    ON data['ulasan_1']

    # Step 3: menghapus username
    data['ulasan_1']<-data['ulasan_1'].replace('@[\w]+', '')
    # Step 4: Ganti karakter '\n' dengan spasi
    data['ulasan_1'] <- data['ulasan_1'].replace("\\\\n", " ")
    # Step 5: Ganti karakter baris baru '\n' dengan spasi
    data['ulasan_1'] <- data['ulasan_1'].replace("\n", " ")
    # Step 6: Ganti karakter carriage return '\r' dengan spasi
    data['ulasan_1'] <- data['ulasan_1'].replace("\r", " ")
    # Step 7: Ganti beberapa spasi dengan satu spasi
    data['ulasan_1'] <- data['ulasan_1'].replace("    ", " ")
    # Step 8: menghapus link
    data['ulasan_1']<-data['ulasan_1'].replace(URL_PATTERN, " ")
    # Step 9: Hapus tanda kutip
    data['ulasan_1'] <- data['ulasan_1'].replace("'", '')
    # Step 10: Hapus karakter heksadesimal
    data['ulasan_1'] <- data['ulasan_1'].replace("\\\\x[a-zA-z0-9][a-zA-z0-9]", "")
```



```

RETURN data
END FUNCTION

```

Untuk lebih jelasnya mengenai proses data cleaning bagaimana sebelum dilakukan dan sesudah dilakukan data cleaning dapat dilihat pada tabel 3.2.

Tabel 3.2 Contoh Data Cleaning

Sebelum Data Cleaning	Sesudah Data Cleaning
Ini gimana linkaja,, perbaikan setengah hari gak selesai ² , perbaikan pas jam kerja pula,, ganggu kerjaan aja rewel gak selesai ²	Ini gimana linkaja,, perbaikan setengah hari gak selesai,, perbaikan pas jam kerja pula,, ganggu kerjaan aja rewel gak selesai

3.2.1.2. Case Folding

Case Folding adalah proses mengubah semua huruf yang ada menjadi huruf kecil. Hal ini bertujuan untuk membuat proses lebih sederhana dan mengurangi sensitifitas pada kalimat. Proses *case folding* digambarkan pada *pseudocode* dibawah ini

```

FUNCTION lowerCase(data) :

    # Step 1: Ubah teks di kolom 'ulasan_1' menjadi huruf kecil
    data['ulasan_2'] <- data['ulasan_1'].to_lower()

    # Step 2: Kembalikan data yang diubah

    RETURN data

END FUNCTION

```

Untuk lebih jelasnya mengenai proses case folding bagaimana sebelum dilakukan dan sesudah dilakukan case folding dapat dilihat pada tabel 3.3.

Tabel 3.3 Contoh Case Folding

Sebelum Case Folding	Sesudah Case Folding
Ini gimana linkaja,, perbaikan setengah hari gak selesai,, perbaikan pas jam kerja pula,, ganggu kerjaan aja rewel gak selesai	ini gimana linkaja,, perbaikan setengah hari gak selesai,, perbaikan pas jam kerja pula,, ganggu kerjaan aja rewel gak selesai

3.2.1.3. Punctual Removal

Punctual Removal adalah proses untuk menghapus atau menghilangkan karakter yang tidak memiliki makna seperti tanda baca, angka atau simbol serta atribut tidak penting. Karena di penelitian ini akan melakukan penelitian terhadap kalimat bukan karakter. Proses *punctual removal* digambarkan pada *pseudocode* dibawah ini

```
FUNCTION removePunct(data):
    # Step 1: salin teks dari kolom 'ulasan_2' ke kolom
    'ulasan_3'
    data['ulasan_3'] <- data['ulasan_2']
    # Step 2: Ganti semua karakter non-abjad dengan spasi di
    kolom 'ulasan_3'
    data['ulasan_3'] <- data['ulasan_3'].replace('[^a-zA-
Z]+', ' ')
    # Step 3: Kembalikan data hasil modifikasi
    RETURN data
END FUNCTION
```

Untuk lebih jelasnya mengenai proses punctual removal bagaimana sebelum dilakukan dan sesudah dilakukan punctual removal dapat dilihat pada tabel 3.4.

Tabel 3.4 Contoh Punctual Removal

Sebelum <i>Punctual Removal</i>	Sesudah <i>Punctual Removal</i>
ini gimana linkaja,, perbaikan setengah hari gak selesai,, perbaikan pas jam kerja pula,, ganggu kerjaan aja rewel gak selesai	ini gimana linkaja perbaikan setengah hari gak selesai perbaikan pas jam kerja pula ganggu kerjaan aja rewel gak selesai

3.2.1.4. Normalization

Normalization adalah proses untuk mengubah kalimat yang berisi kata tidak baku, seperti singkatan dan kesalahan ketik, menjadi kata-kata yang sesuai dengan bentuk baku dan memiliki makna yang jelas. Peneliti menggunakan kamus yang dibuat sendiri

yang sesuai dengan data ulasan. Proses *normalization* digambarkan pada *pseudocode* dibawah ini.

```

FUNCTION normalize_text(text, stdword_, nonstdword_):
    # Step 1: Pisahkan teks menjadi kata-kata individual
    words <- split text by " "

    # Step 2: Ulangi setiap kata dalam teks
    FOR i FROM 0 TO length(words) - 1:

        # Step 3: Periksa apakah kata tersebut ada dalam daftar kata
        # tidak baku
        IF words[i] IS IN nonstdword_:

            # Step 4: Temukan indeks kata tidak baku
            index <- index of words[i] in nonstdword_

            # Step 5: Ganti kata yang tidak baku dengan kata baku yang
            # sesuai
            words[i] <- stdword_[index]

    END FOR

    # Step 6: Gabungkan kata-kata kembali menjadi satu string
    normalized_text <- join words with " "

    # Step 7: Kembalikan teks yang dinormalisasi
    RETURN normalized_text
END FUNCTION

FUNCTION normalizeData(data, stdword_, nonstdword_):
    # Step 1: Salin teks dari kolom 'ulasan_3' ke kolom
    # 'ulasan _4'
    data['ulasan_4'] <- data['ulasan_3']

    # Step 2: Terapkan fungsi normalize_text ke setiap teks di
    # kolom 'ulasan_4'
    data['ulasan_4'] <- MAP each text IN data['ulasan_4'] WITH
    FUNCTION(com):

        RETURN normalize_text(com, stdword_, nonstdword_)

    END FUNCTION

    # Step 3: Kembalikan data yang diubah

```

```

RETURN data
END FUNCTION

```

Untuk lebih jelasnya mengenai proses normalization bagaimana sebelum dilakukan dan sesudah dilakukan normalization dapat dilihat pada tabel 3.5.

Tabel 3.5 Contoh Normalization

Sebelum <i>Normalization</i>	Sesudah <i>Normalization</i>
ini gimana linkaja perbaikan setengah hari gak selesai perbaikan pas jam kerja pula ganggu kerjaan aja rewel gak selesai	ini bagaimana linkaja perbaikan setengah hari tidak selesai perbaikan pas jam kerja pula ganggu pekerjaan saja rewel tidak selesai

3.2.1.5. Stopword Removal

Stopword Removal adalah proses untuk menghapus kata-kata yang sering ditemukan tanpa menghilangkan makna dari kalimat tersebut. Hal ini dilakukan untuk menyisakan kalimat penting saja. Proses stopwords removal digambarkan pada *pseudocode* dibawah ini.

```

FUNCTION removeMeaningless(data, meaningless_):
    FOR each word IN meaningless_:
        # Step 1: Buat pola regex agar cocok dengan keseluruhan
        kata
        regex_meaningless <- r"\b" + word + r"\b"
        # Step 2: Hapus kata dari kolom 'ulasan_4' menggunakan
        pola regex
        data['ulasan_4'] <-
        data['ulasan_4'].replace(regex_meaningless, '')
    END FOR
    # Step 3: Kembalikan data hasil modifikasi
    RETURN data
END FUNCTION

```

Untuk lebih jelasnya mengenai proses stopwords removal bagaimana sebelum dilakukan dan sesudah dilakukan stopwords removal dapat dilihat pada tabel 3.6.

Tabel 3.6 Contoh Stopword Removal

<i>Sebelum Stopword Removal</i>	<i>Sesudah Stopword Removal</i>
ini bagaimana linkaja perbaikan setengah	linkaja perbaikan tidak selesai perbaikan
hari tidak selesai perbaikan pas jam kerja	pas jam kerja ganggu pekerjaan rewel
pula ganggu pekerjaan saja rewel tidak	tidak selesai
selesai	

3.2.1.6. Stemming

Stemming adalah proses untuk mengubah kalimat yang berimbuhan menjadi kata dasar. Hal ini dilakukan untuk mempermudah model dalam mengenali makna dari kata tersebut. Proses stemming digambarkan pada pseudocode dibawah ini.

```

FUNCTION stemming(text, stemmer_id):

# Step 1: Pisahkan teks menjadi kata-kata individual
    text_split <- split text by " "

# Step 2: Inisialisasi daftar kosong untuk menyimpan kata-kata
yang telah di stem
    stemmed_list <- empty list

# Step 3: Iterasi setiap kata dalam teks.
    FOR each word IN text_split:

# Step 4: Terapkan proses stemming pada kata tersebut
        stemmed_word <- stemmer_id.stem(word)

# Step 5: Tambahkan kata yang telah di-stem ke dalam daftar
        ADD stemmed_word TO stemmed_list

    END FOR

# Step 6: Gabungkan kembali kata-kata yang telah di-stem menjadi
sebuah string
    stemmed_text <- join stemmed_list with " "
```

```

# Step 7: Kembalikan teks yang telah di-stem
    RETURN stemmed_text
END FUNCTION

FUNCTION stemData(data, stemmer_id):
# Step 1: Salin teks dari 'kolom ulasan_5' ke kolom baru
'ulasan_6'.

    data['ulasan_6'] <- data['ulasan_5']

# Step 2: Terapkan fungsi stemming pada setiap teks di kolom
ulasan_6

    data['ulasan_6'] <- MAP each text IN data['ulasan_6'] WITH
FUNCTION(com):

        RETURN stemming(com, stemmer_id)

    END FUNCTION

# Step 3: Kembalikan dataset yang telah dimodifikasi

    RETURN data
END FUNCTION

```

Untuk lebih jelasnya mengenai proses stemming bagaimana sebelum dilakukan dan sesudah dilakukan stemming dapat dilihat pada tabel 3.7.

Tabel 3.7 Contoh Stemming

Sebelum dilakukan <i>Stemming</i>	Sesudah dilakukan <i>Stemming</i>
linkaja perbaikan tidak selesai perbaikan	linkaja baik tidak selesai baik pas jam
pas jam kerja ganggu pekerjaan rewel	kerja ganggu kerja rewel tidak selesai
tidak selesai	

3.2.1.7. Tokenizing

Tokenizing adalah proses membagi kalimat menjadi bagian yang lebih kecil berupa token-token. Proses tokenisasi digambarkan dengan *pseudocode* dibawah ini

```

Inisialisasi tokenizer dengan parameter yang ditentukan
num_words=MAX_NB_WORDS, lower=True, char_level=False

```

```

# Step 1: Menerapkan tokenizer pada teks gabungan dari daftar
pelatihan dan pengujian untuk membuat indeks kata

CALL tokenizer.fit_on_texts WITH PARAMETERS (train_list +
test_list)

# Step 2: Mengonversi teks pelatihan dan pengujian menjadi
urutan indeks kata.

train_seq <- tokenizer.texts_to_sequences(train_list)

test_seq <- tokenizer.texts_to_sequences(test_list)

# Step 3: Mendapatkan indeks kata dari tokenizer

word_index <- tokenizer.word_index

# Step 4: Memastikan semua urutan memiliki panjang yang sama
dengan menambahkan padding jika diperlukan

train_seq_pad <- CALL pad_sequences WITH PARAMETERS train_seq,
maxlen=max_len

test_seq_pad <- CALL pad_sequences WITH PARAMETERS test_seq,
maxlen=max_len

# Step 5: Mengonversi urutan yang telah di-pad menjadi array
numpy

train_seq_pad <- CONVERT train_seq_pad TO np.array
test_seq_pad <- CONVERT test_seq_pad TO np.array

```

Untuk lebih jelasnya mengenai proses tokenizing bagaimana sebelum dilakukan dan sesudah dilakukan tokenizing dapat dilihat pada tabel 3.8.

Tabel 3.8 Contoh Tokenizing

Sebelum dilakukan <i>Tokenizing</i>	Sesudah dilakukan <i>Tokenizing</i>
linkaja baik tidak selesai baik pas jam	'linkaja', 'baik', 'tidak', 'selesai', 'baik',
kerja ganggu kerja rewel tidak selesai	'pas', 'jam', 'kerja', 'ganggu', 'kerja',
	'rewel', 'tidak', 'selesai'

3.2.1.8. Word Embedding

Word Embedding adalah tahapan yang bertujuan untuk mengubah data yang sudah berbentuk token menjadi bentuk vektor. Proses ini dilakukan agar model bisa mengolah data karena model hanya mampu mengolah data yang berbentuk angka atau numerik. Untuk proses ini digunakan library dari FastText. FastText merupakan pengembangan dari library Word2Vec yang lebih dahulu dikenal dalam proses word embedding. FastText tidak hanya memakai satu kata secara utuh untuk diproses tapi menggunakan n-gram, kelebihan lainnya adalah waktu proses yang relatif cepat dan mampu menangani kata yang tidak pernah muncul dalam kamus (Ariwibowo & Girsang, 2021).

Setelah korpus terbentuk, maka setiap kata akan diubah menjadi vektor yang disebut embedding matrix yang berdimensi 300 dimana akan digunakan sebagai weight pada pembuatan model di tahap selanjutnya. Proses tokenisasi digambarkan dengan *pseudocode* dibawah ini

```
INITIALIZE words_not_found AS empty list

# Step 1: Menentukan jumlah kata (nb_words) yang akan digunakan
dalam matriks

nb_words <- MIN(MAX_NB_WORDS, length(word_index) + 1)

# Step 2: Inisialisasi matriks embedding dengan nol.

embedding_matrix <- INITIALIZE matrix WITH SHAPE (nb_words,
embed_dim) FILLED WITH zeros

# Step 3: Melakukan iterasi pada setiap kata dan indeksinya

FOR each word, i IN word_index:

# Step 4: Melewati kata-kata dengan indeks yang lebih besar
atau sama dengan nb_words.

    IF i >= nb_words:

        CONTINUE

# Step 5: Mendapatkan vektor embedding untuk kata dari model
FastText

    embedding_vector <- fasttext_word_to_index[word]

# Step 6: Jika vektor embedding ditemukan dan panjangnya lebih
dari 0, tambahkan ke matriks embedding
```



```

IF (embedding_vector IS NOT None) AND
(length(embedding_vector) > 0):

    embedding_matrix[i] <- embedding_vector

ELSE:

    # Step 7: Jika tidak, tambahkan kata tersebut ke daftar
    words_not_found.

    ADD word TO words_not_found

END FOR

# Step 8: Mencetak jumlah embedding kata yang bernilai nol
dalam matriks embedding.

PRINT 'number of null word embeddings: ',
SUM(SUM(embedding_matrix, axis=1) == 0)

```

Berikut merupakan contoh dari embedding matrix yang merupakan vektor dari kata 'aplikasi' pada Gambar 3.3.

```

[ 0.0155  0.0517 -0.013  0.0371 -0.0123 .
-0.0428  0.061   0.0341  0.0157 -0.0137
-0.0423 -0.022  -0.0517 -0.0122 -0.0027
-0.0205  0.0248 -0.0048  0.0259 -0.0072
 0.0101 -0.0255 -0.0033  0.0228  0.0357 .
 0.0301 -0.0159  0.0197  0.0085 -0.0064 .
-0.0052  0.044   0.0217  0.011   0.0393 .
 0.0313  0.0149 -0.0191  0.0635  0.0082
 0.0046 -0.0442  0.0299  0.0126  0.015 .
 0.0247 -0.0088  0.0005 -0.0126  0.0577 .
-0.0071 -0.0037 -0.0142  0.037  -0.0334 .
 0.0146  0.0077 -0.016  -0.0138 -0.0177
-0.0053  0.0183 -0.021  -0.0174  0.0089
 0.0104  0.0578 -0.0096 -0.0727 -0.0019
 0.0718  0.0085  0.0535 -0.0139 -0.0114
-0.0157 -0.0212 -0.0136  0.0042  0.0464
-0.0099  0.0223  0.0309  0.0186 -0.0196
-0.0122  0.0676 -0.0298  0.0264 -0.0431
 0.0226  0.0093  0.0121 -0.0185  0.0377 .
 0.0124  0.0124  0.0205  0.0321 -0.0031 .
-0.013  -0.0167 -0.0189 -0.0144 -0.0248
-0.0053 -0.0182  0.0513 -0.0165  0.0169 .
 0.0313 -0.0443 -0.0048 -0.0111  0.0061 .
-0.0018  0.0176  0.0183  0.0549 -0.0249 .
-0.0089  0.0043 -0.0004 -0.0443 -0.0057
-0.0052 -0.029   0.019   0.0017 -0.0096 .
-0.0376  0.0225  0.0106  0.0297 -0.0104
 0.0063 -0.0168 -0.0069 -0.0265 -0.0085

```

Gambar 3.3 Embedding Matrix

3.2.2. Topic Building

Tahapan ini merupakan proses untuk menemukan topik yang tersembunyi pada dokumen teks dengan menggunakan Latent Dirichlet Allocation atau disingkat dengan LDA. Adapun tahapan untuk membangun model ini adalah :

1. Tahap pertama yaitu mengumpulkan data teks hasil preprocessing agar data lebih mudah dipahami oleh sistem.
2. Tahap kedua yaitu membuat dictionary atau korpus kata yang akan digunakan untuk pelatihan model LDA.
3. Tahap ketiga yaitu menentukan jumlah topik yang akan diekstrak dan jumlah kata teratas yang akan ditampilkan untuk setiap topik.
4. Tahap keempat yaitu melatih model LDA dengan menggunakan library gensim dari python berdasarkan parameter dan korpus kata yang sudah ada. Parameter yang diperlukan dalam pembangunan model ini seperti num_topics yang menunjukkan jumlah topik yang akan diekstrak, nilai alpha yaitu parameter untuk distribusi topik per dokumen dan nilai eta yaitu parameter untuk distribusi kata per topik.
5. Tahap kelima yaitu melakukan evaluasi model berdasarkan parameter yang sudah ditetapkan dengan menggunakan nilai koheren. Jika model sudah mampu mendapatkan nilai koheren yang mendekati 1, maka diperoleh hasil dari model LDA berupa visualisasi grafik berdasarkan jumlah distribusi topik dan kemudian digunakan menjadi topik – topik terhadap konteks data. Namun, jika belum mendapatkan nilai koheren yang mendekati 1, maka dilakukan iterasi atau perulangan dengan mengubah nilai parameter yang sesuai.

3.2.3. Model Building

Adapun tahapan yang dilakukan dalam membangun model LSTM dengan mekanisme atensi yaitu :

1. Tahap pertama yaitu menggunakan *embedding layer*, yang merupakan vektor sepanjang 300 atau berdimensi 300 yang merepresentasikan setiap kata pada kalimat ulasan. Input dari *embedding layer* ini adalah matrix yang dihasilkan pada tahap *word embedding* lalu memberikan hasil berupa vektor berdimensi 300 dengan panjang yang sama.

2. Tahap kedua yaitu masuk ke layer LSTM atau *long short term memory*, dimana pada layer ini dilakukan pelatihan model dengan jumlah neuron sebanyak 64 unit.
3. Tahap ketiga yaitu output dari pemodelan LSTM akan masuk ke attention layer.
4. Selanjutnya menuju dense layer, pada layer ini akan ditentukan output melalui fungsi aktivasi. Pada penelitian ini, digunakan fungsi aktivasi softmax yang dapat mengklasifikasikan objek yang memiliki jumlah kelas lebih dari dua. *Softmax* dapat menghitung probabilitas untuk setiap kelas yang diklasifikasikan dengan cara mengambil vektor yang bernilai riil dari semua label kelas yang ada (Ilahiyah S & Nilogiri A, 2018).

Dalam penelitian ini digunakan batch size 32 dimana penentuan jumlah ini ditentukan setelah adanya proses *tunning* serta fungsi optimasi yang digunakan adalah *adam* (*Adaptive Moment Estimation*) untuk mendapatkan parameter yang menghasilkan nilai loss minimum.

3.2.4. Classification

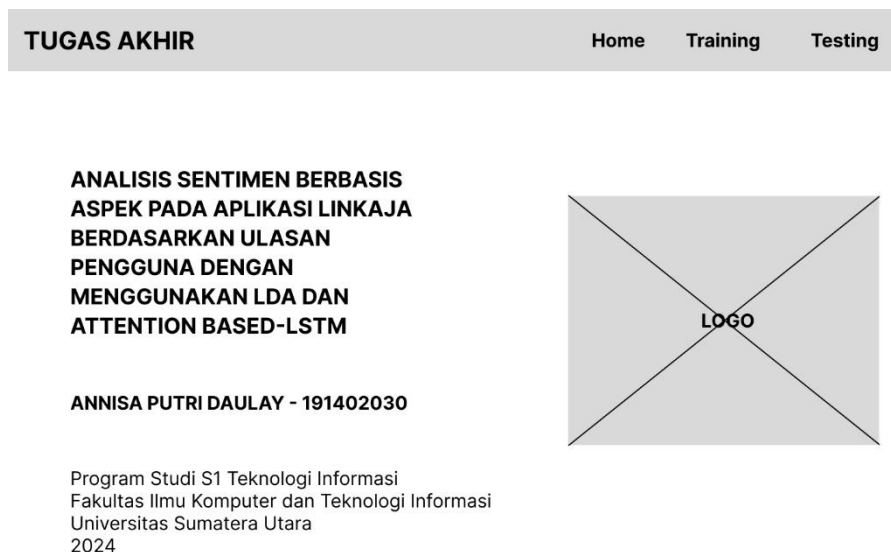
Setelah model dibangun, maka tahap klasifikasi sudah bisa dilakukan untuk mengklasifikasi sentimen berdasarkan aspek yang didapatkan melalui proses pemodelan topik. Dataset testing dimasukkan oleh user dengan format csv ataupun dengan kalimat secara langsung. Selanjutnya dataset tersebut melewati tahap preprocessing kemudian akan dilakukan klasifikasi berdasarkan model yang telah dibuat pada tahap training.

3.3. Perancangan Aplikasi Sistem

Tahapan perancangan ini bertujuan untuk mempermudah pengguna untuk menjalankan sistem, peneliti menggunakan sistem berbasis web sebagai antarmuka sistem dengan pengguna.

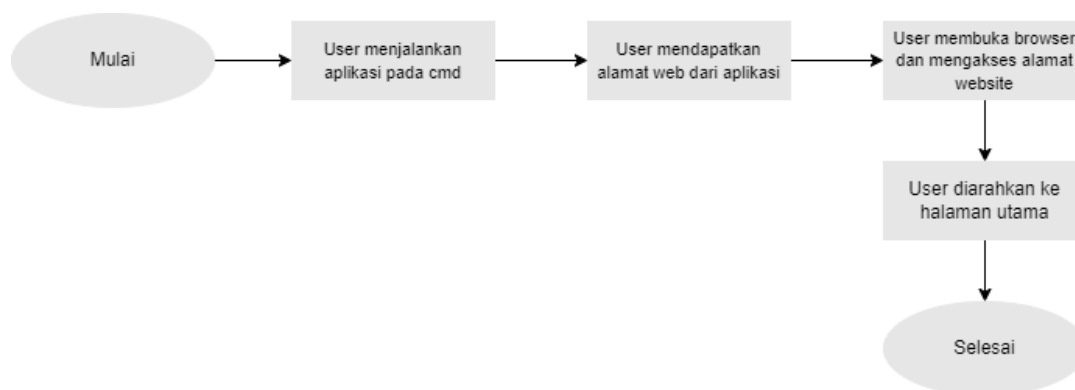
3.3.1. Rancangan Halaman Utama (Home)

Halaman Utama (Home) merupakan halaman yang ditampilkan pertama kali saat user mengoperasikan sistem. Di halaman ini ditampilkan judul penelitian dan informasi mengenai identitas penulis. Untuk lebih jelas mengenai tampilan halaman utama dapat dilihat pada Gambar 3.4.



Gambar 3.4 Halaman Utama (Home)

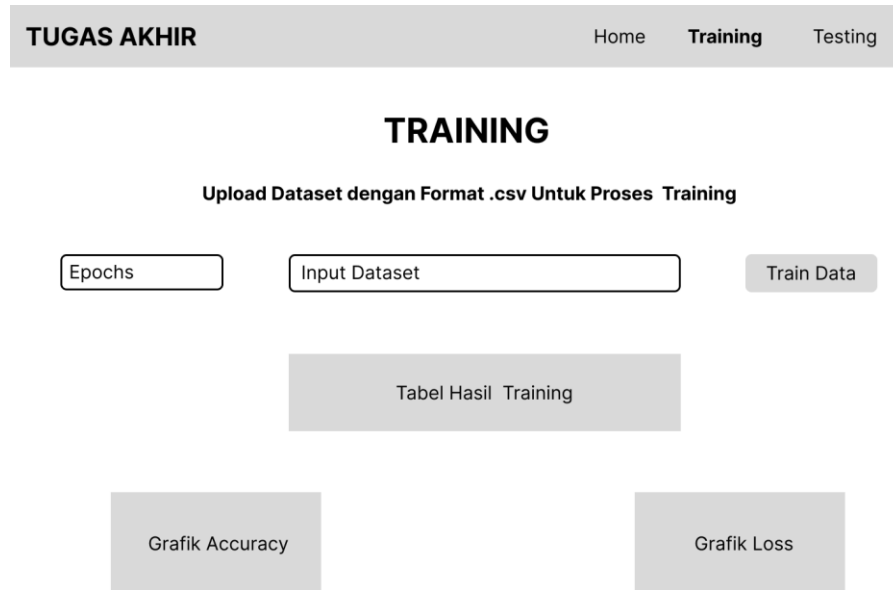
Untuk mengakses halaman utama dapat dilihat pada userflow diagram pada Gambar 3.5.



Gambar 3.5 Userflow diagram Halaman Home

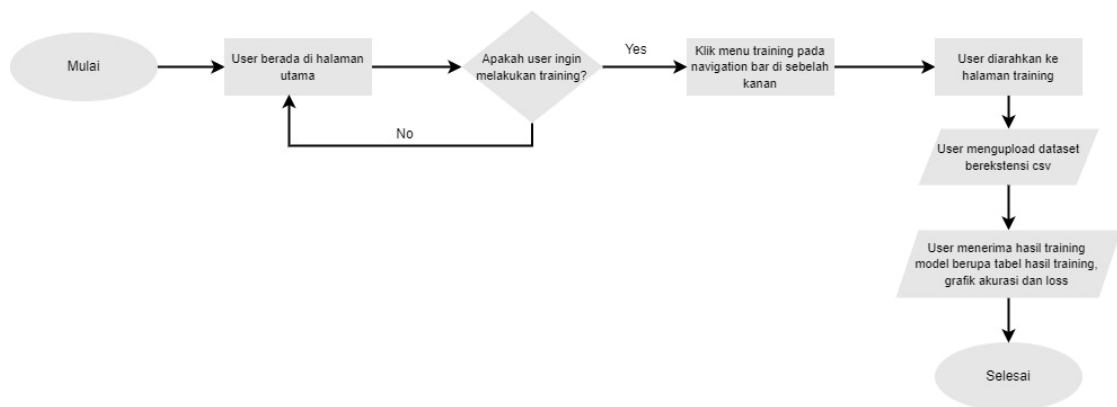
3.3.2. Rancangan Halaman Training

Halaman ini digunakan oleh user untuk melakukan proses pelatihan model. User harus mengunggah file dengan format csv, kemudian sistem akan memberikan output berupa tabel yang menampilkan hasil preprosesing dan plot grafik dari nilai akurasi dan loss. Rancangan halaman training ini dapat dilihat pada gambar 3.6.



Gambar 3.6 Rancangan Halaman Training

Proses mengenai bagaimana halaman training dijalankan dapat dilihat pada userflow diagram pada Gambar 3.7.

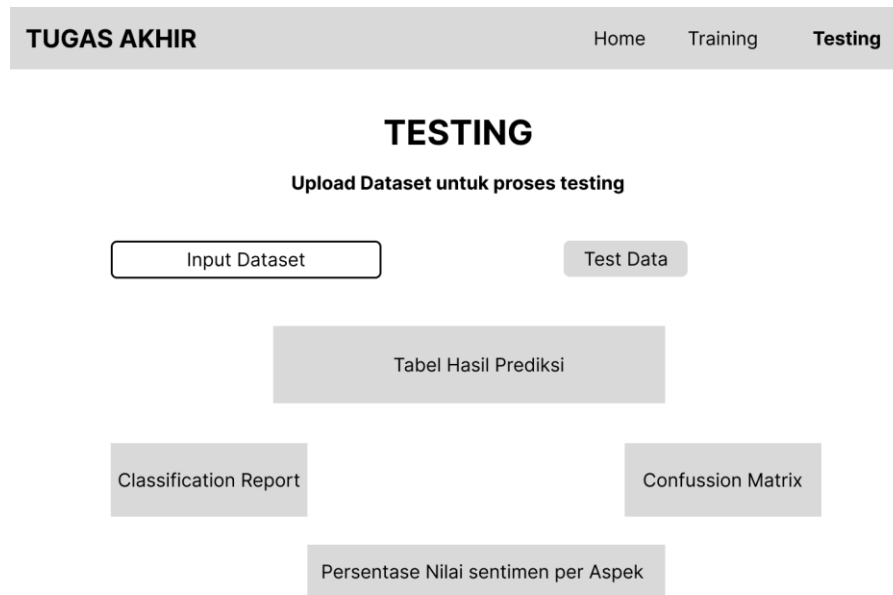


Gambar 3.7 Userflow diagram Halaman Training

3.3.3. Rancangan Halaman Testing

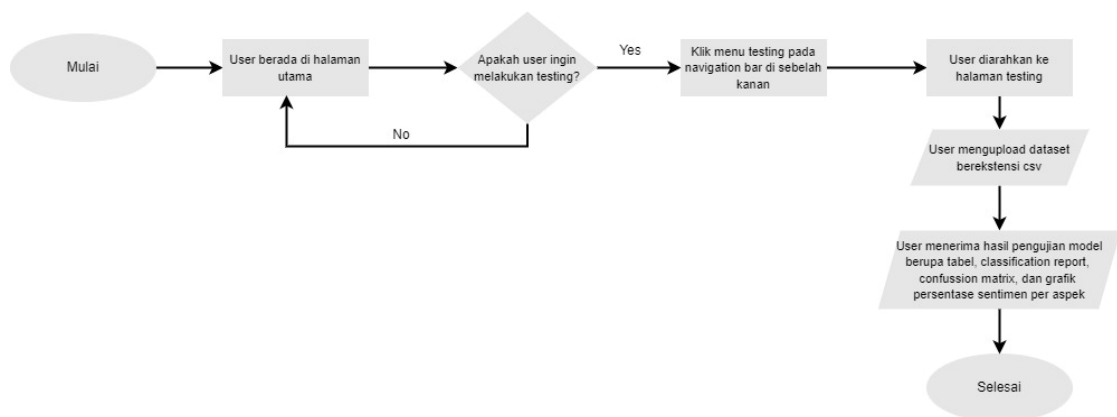
Halaman ini digunakan oleh user untuk melakukan pengujian terhadap data uji dimana user dapat memasukkan file ulasan pengguna berekstensi csv yang kemudian akan memberikan output berupa tabel yang menunjukkan hasil prediksi model, classification report dan confusion matrix dari masing-masing aspek untuk mengevaluasi kinerja model. Untuk menambah informasi yang diperoleh, ditampilkan juga grafik yang

menampilkan persentase dari sentimen masing-masing aspek. Rancangan halaman testing ini dapat dilihat pada Gambar 3.8.



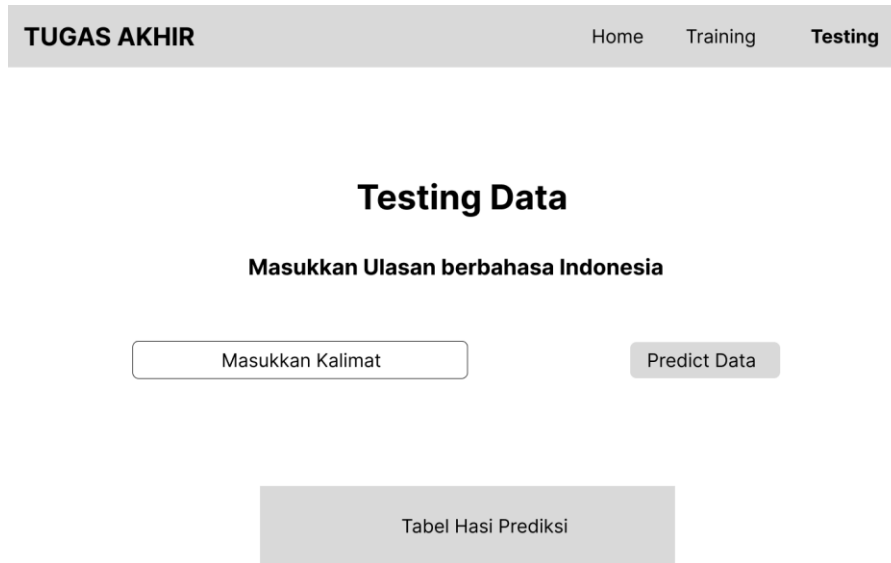
Gambar 3.8 Rancangan Halaman Testing

Proses mengenai bagaimana halaman testing dijalankan dapat dilihat pada userflow diagram pada Gambar 3.9.



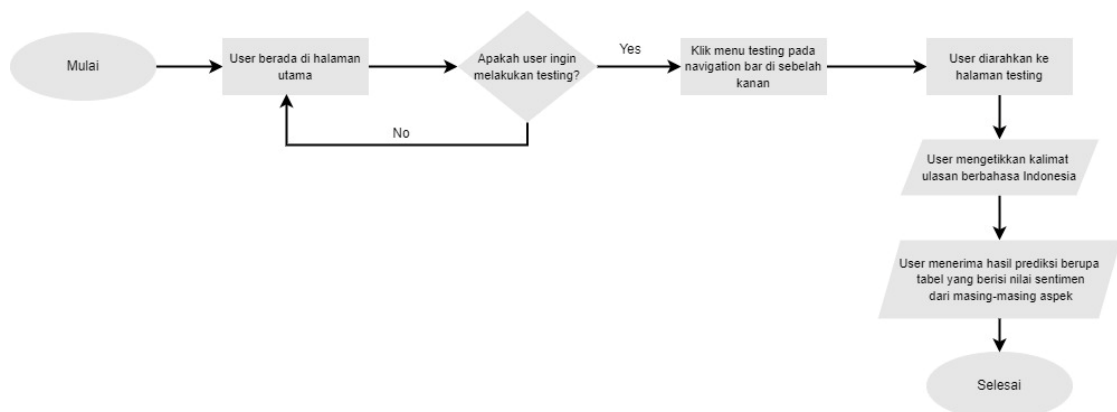
Gambar 3.9 Userflow diagram Halaman Testing

User juga dapat melakukan uji coba terhadap model dengan memasukkan ulasan berbahasa Indonesia untuk dapat diprediksi dengan output berupa sentimen dari masing-masing aspek. Rancangan halaman testing ini dapat dilihat pada Gambar 3.10.



Gambar 3.10 Rancangan Halaman Uji Coba

Proses mengenai bagaimana halaman testing user input ini dijalankan dapat dilihat pada userflow diagram pada Gambar 3.11.



Gambar 3.11 Userflow diagram Halaman Testing User Input

3.4. Metode Evaluasi

Metode evaluasi dilakukan untuk menilai kinerja suatu model dalam melakukan klasifikasi polaritas sentimen ulasan pengguna aplikasi LinkAja. Confussion Matrix adalah metode yang digunakan dalam penelitian ini. Evaluasi yang diterapkan adalah

perhitungan terhadap variabel *accuracy*, *precision*, *recall*, dan *F1-Score*. Confussion Matrix diterapkan dalam tabel 3.9 di bawah ini.

Tabel 3.9 Confussion Matrix

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

Terdapat 4 istilah yang merupakan representasi hasil proses klasifikasi pada *confussion matrix* yaitu :

- True Posive (TP) menunjukkan banyaknya hasil prediksi positif dengan data aktual positif
- True Negative (TN) menunjukkan banyaknya hasil prediksi negatif dengan data aktual negatif
- False Positive (FP) menunjukkan banyaknya hasil prediksi positif dengan data aktual negatif
- False Negative (FN) menunjukkan banyaknya hasil prediksi negatif dengan data aktual positif

Evaluasi yang diterapkan adalah perhitungan terhadap variabel *accuracy*, *precision*, *recall*, dan *F1-Score*. Untuk lebih jelasnya sebagai berikut :

a. Accuracy

Accuracy merupakan metrik evaluasi untuk mengukur kebenaran dalam melakukan prediksi terhadap setiap data berdasarkan nilai aktualnya. Semakin tinggi nilai *accuracy*, maka semakin akurat model dalam melakukan klasifikasi. Untuk mendapatkan nilai *accuracy* dapat dilihat pada persamaan 3.1 dibawah ini

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.1)$$

b. Precision

Precision merupakan metrik evaluasi untuk mengukur hasil prediksi yang benar untuk data aktual positif dari total prediksi positif yang dilakukan. Untuk mendapatkan nilai *precision* dapat dilihat pada persamaan 3.2 dibawah ini

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3.2)$$

c. Recall

Recall merupakan metrik evaluasi untuk mengidentifikasi data aktual positif dengan benar dari total prediksi yang dilakukan. Untuk mendapatkan nilai *recall* dapat dilihat pada persamaan 3.3 dibawah ini

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3.3)$$

d. F1-Score

F1-Score merupakan metrik evaluasi untuk menggambarkan keseimbangan antara *precision* dan *recall*. Untuk mendapatkan nilai *F1-Score* dapat dilihat pada persamaan 3.4 dibawah ini

$$\text{F1-Score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (3.4)$$

BAB 4

IMPLEMENTASI DAN PENGUJIAN SISTEM

4.1. Implementasi Sistem

Sistem yang dirancang dalam penelitian ini dibangun dengan menggunakan perangkat keras dan perangkat lunak yang mendukung, yaitu :

4.1.1. Spesifikasi perangkat keras dan perangkat lunak

Adapun perangkat keras yang digunakan dalam implementasi sistem memiliki spesifikasi sebagai berikut :

1. Processor Intel Core(TM) i7-8565U CPU @ 1.80GHz
2. Memory (RAM) 8.00 GB
3. SSD 512 GB

Adapun perangkat lunak yang digunakan dalam implementasi sistem memiliki spesifikasi sebagai berikut :

1. Operating System : Windows 10 Home Single Language
2. Python versi 3.8 dengan library diantaranya Tensorflow, Keras, Pandas, Matplotlib, Numpy, Sklearn, dan NLTK
3. Google Colab
4. Microsoft Visual Studio Code

4.1.2. Implementasi perancangan tampilan antarmuka

Adapun implementasi dari rancangan tampilan antarmuka yang telah dibuat yaitu :

1. Tampilan Halaman Utama

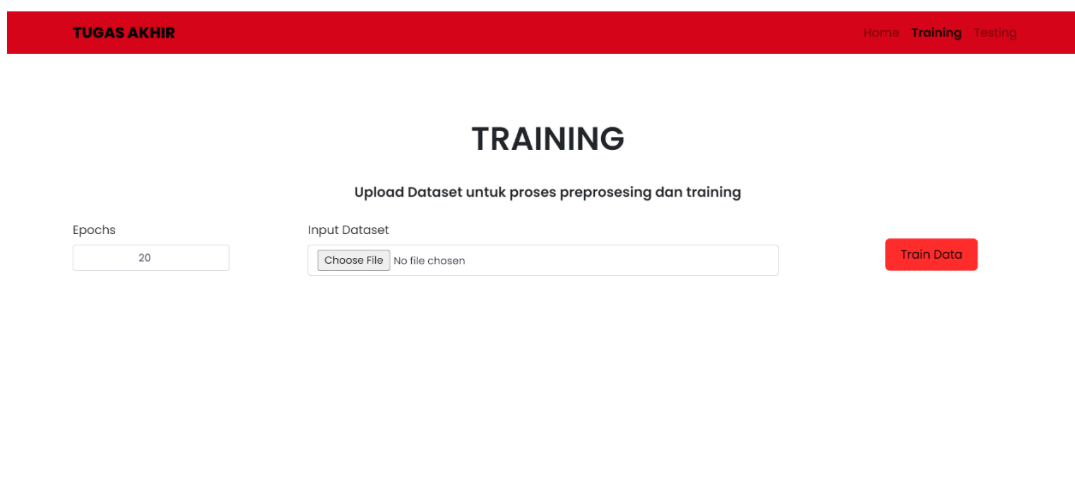
Halaman utama merupakan tampilan yang pertama kali dilihat oleh pengguna ketika menjalankan sistem. Halaman ini berisi judul penelitian dan identitas peneliti. Untuk lebih jelasnya tampilan halaman utama dapat dilihat pada Gambar 4.1.



Gambar 4.1 Tampilan Halaman Utama

2. Tampilan Halaman Training

Halaman ini merupakan halaman yang muncul ketika pengguna menjalankan fitur training. Pada halaman ini, pengguna akan melakukan *training* untuk melatih model dengan melakukan input dataset berupa file dengan ekstensi csv. Selain itu, pengguna dapat menentukan *epochs* agar menghasilkan hasil yang lebih baik. Dalam sistem ini, *epochs* yang ditetapkan adalah 40. Untuk lebih jelasnya tampilan halaman training dapat dilihat pada Gambar 4.2.



Gambar 4.2 Tampilan Halaman Training

Selanjutnya, setelah dataset diunggah dan melalui proses training, maka hasil dari proses tersebut akan ditampilkan oleh sistem diantaranya hasil preprosesing, grafik nilai akurasi dan grafik nilai loss dari masing-masing aspek. Pada halaman ini juga

disediakan fitur untuk menentukan jumlah tampilan per-halaman dan kolom *search* untuk mencari ulasan. Fitur *pagination* juga disediakan untuk memudahkan melihat ulasan dengan jumlah yang cukup banyak. Untuk lebih jelasnya mengenai tampilannya, dapat dilihat pada Gambar 4.3 dan Gambar 4.4.

TUGAS AKHIR

Home

Training

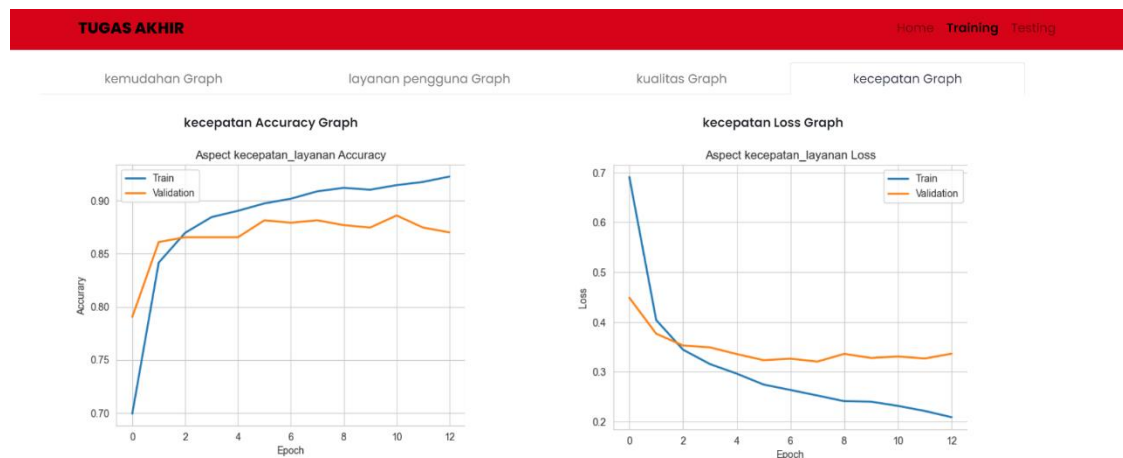
Testing

Show10entries

Search:

Review	Preprocessing	kemudahan	layanan pengguna	kualitas	kecepatan
Ada apa dgn aplikasi ini...Isi saldo tp gagal trus...	aplikasi isi saldo gagal	0	0	-1	0
Ada fitur link kartu debit tp tak bisa digunakan. GOUBLOK!!! Butuh brp dekade itu fitur baru aktif. TOLOL developernya!!!	fitur link kartu debit tidak goublok butuh dekade fitur aktif tolo developer	-1	0	-1	0
Ada fitur Top Up e-money. Ya itu bagus, tapi setelah Top Up, saldo LinkAja dah kepotong tapi saldo e-moneynya ga masuk. KECEWA Saya rasa banyak diluar sana yang mengalami hal yang sama. Tolong dong, ini pertanggung jawabannya gimana??	fitur top up e money bagus setelah top up saldo linkaja potong saldo e moneynya tidak masuk kecewa luar alami tolong tanggung jawab gimana	0	0	-1	0
Ada kendala Udah coba tlpn cs tapi gak bisa nyambung !	kendala udah coba telepon customer service tidak nyambung	0	-1	0	0
Ada kendala. Aplikasi tdk bisa digunakan	kendala aplikasi tidak transaksi saldo	0	-1	-1	0

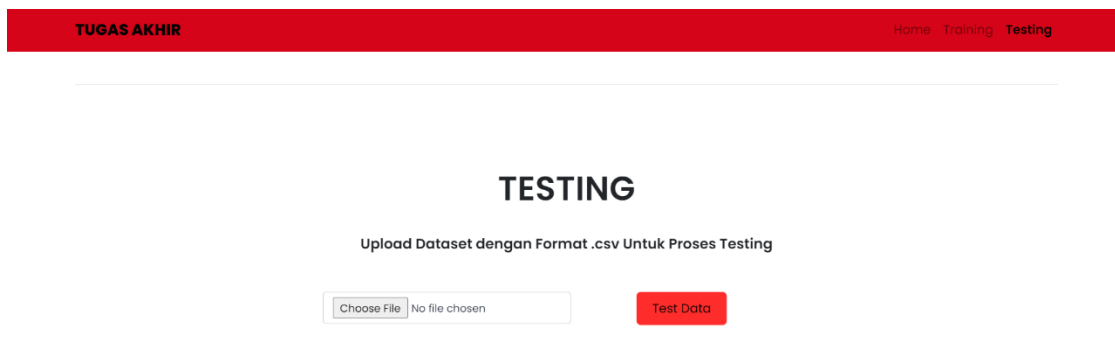
Gambar 4.3 Tampilan Hasil Preprocessing



Gambar 4.4 Tampilan Hasil Grafik per aspek

3. Tampilan Halaman Testing

Halaman ini merupakan yang muncul ketika pengguna menjalankan fitur testing. Pada halaman ini, pengguna akan melakukan *testing* untuk melakukan prediksi dengan melakukan input dataset berupa file dengan ekstensi csv. Untuk lebih jelasnya tampilan halaman testing dapat dilihat pada Gambar 4.5.



Gambar 4.5 Tampilan Halaman Testing

Selanjutnya, setelah dataset diunggah dan melalui proses testing maka sistem akan menampilkan hasil prediksi, *classification report* dan *confusion matrix* dari masing-masing aspek. Untuk lebih jelasnya mengenai tampilannya dapat dilihat pada Gambar 4.6, dan Gambar 4.7.

TUGAS AKHIR		Home Training Testing							
Show 10 entries		Search:							
Review	Preprocessing	Actual				Prediction			
		kemudahan	layanan pengguna	kualitas	kecepatan	kemudahan	layanan pengguna	kualitas	kecepatan
Tidak bisa masuk aplikasi error	tidak masuk aplikasi error	-1	0	0	0	-1	0	-1	0
tidak bisa sign up, berkali2 masukkan kode otp ga muncul2 ke tahap berikutnya. so bad	tidak sign up kali masuk kode otp tidak muncul tahap so bad	-1	0	0	0	-1	0	0	0

Gambar 4.6. Tampilan Hasil Prediksi



Gambar 4.7. Tampilan Hasil *Classification Report* dan *Confusion Matrix*

Selain dengan melakukan input dataset, pengguna juga dapat melakukan testing dengan memasukkan ulasan berbahasa Indonesia pada kolom yang sudah disediakan untuk kemudian diprediksi berdasarkan model yang telah dibangun. Untuk lebih jelasnya tampilan halaman testing ketika user dapat memasukkan ulasan dapat dilihat pada Gambar 4.9.

TUGAS AKHIR Home Training **Testing**

Testing Data

Masukkan Ulasan Berbahasa Indonesia

Masukkan Kalimat

Mulai Testing

Gambar 4.8 Tampilan Halaman Input Ulasan

TUGAS AKHIR
Home Training **Testing**

Testing Data

Masukkan Ulasan Berbahasa Indonesia

ini gimana sih aplikasinya kok susah banget ya dipake, masa saya udah top up tapi saldonya gak masuk. udah hubungin cs juga gak berhasil. hadeh.

Input	kemudahan	layanan pengguna	kualitas	kecepatan
ini gimana sih aplikasinya kok susah banget ya dipake, masa saya udah top up tapi saldonya gak masuk. udah hubungin cs juga gak berhasil. hadeh	Netral	Negatif	Negatif	Netral

Gambar 4.9 Tampilan Hasil Prediksi

4.2. Implementasi Model

4.2.1. Pelatihan Model

Dalam melakukan pelatihan model, penulis melakukan *hyperparameter tuning* atau mencari model yang memiliki performa yang terbaik untuk mengolah data. Penulis melakukan percobaan dengan mengubah nilai unit neuron dan batch size. Selain itu penulis juga menambahkan *callback function*. *Callback function* berguna untuk menghentikan proses training jika model tidak mengalami peningkatan performa lagi. Berikut percobaan *hyperparameter tuning* yang dilakukan dapat dilihat pada tabel 4.1.

Tabel 4.1 *Hyperparameter Tuning* Akurasi

Unit Neuron	Batch Size	Akurasi Validation Aspek			
		Kemudahan	Layanan Pengguna	Kualitas	Kecepatan
32	32	0.88	0.93	0.87	0.90
32	64	0.89	0.93	0.87	0.92
32	128	0.87	0.91	0.84	0.89
64	32	0.90	0.94	0.88	0.92
64	64	0.89	0.93	0.86	0.91
64	128	0.89	0.92	0.87	0.90
128	32	0.89	0.93	0.87	0.91
128	64	0.90	0.93	0.88	0.92

Tabel 4.2 *Hyperparameter Tunning Akurasi (Lanjutan)*

128	128	0.88	0.93	0.88	0.91
-----	-----	------	------	------	------

Tabel 4.3 *Hyperparameter Tunning Loss*

Unit Neuron	Batch Size	Loss Validation Aspek			
		Kemudahan	Layanan Pengguna	Kualitas	Kecepatan
32	32	0.29	0.20	0.30	0.28
32	64	0.29	0.18	0.29	0.21
32	128	0.31	0.22	0.34	0.28
64	32	0.28	0.16	0.28	0.23
64	64	0.27	0.20	0.30	0.21
64	128	0.28	0.21	0.29	0.25
128	32	0.30	0.18	0.30	0.22
128	64	0.27	0.18	0.29	0.24
128	128	0.28	0.20	0.28	0.22

Berdasarkan Tabel 4.1 dan 4.2 performa terbaik dicapai pada model dengan batch size 32, dan unit neuron 64 dengan rata-rata akurasi model training 91% dan loss rata-rata 0.95%.

4.3. Pengujian Model

Setelah proses pembangunan model selesai, maka akan dilakukan pelatihan terhadap model dengan dataset berupa ulasan pengguna aplikasi LinkAja yang telah dibagi untuk proses *training* dengan rasio 90% dan 10% untuk data validasi. Dalam proses pelatihan ini dilakukan *preprocessing* berupa *data cleaning*, *case folding*, *stopword removal*, *stemming*, dan *normalization* yang hasilnya kemudian dapat dilihat pada tabel 4.3.

Tabel 4.4 Data Hasil Pelatihan Model

		Kategori Aspek			
		Kemudahan	Layanan Pengguna	Kualitas	Kecepatan
Review	Preprocessing				
2 transaksi terpending saldo pun di sedot laporan sudah via email pun tidak ada balasan bingung mau komplain kemana	transaksi pending saldo sedot laporan via email tidak balas bingung komplain mana	0	-1	-1	0
Aplikasi aneh, reset PIN tapi selalu jaringan internet dan error. Padahal buat aplikasi lain internet stabil. Hadeuuuuh!!!	aplikasi aneh suruh reset pin cek jaring internet error aplikasi internet stabil hadeuuuuh	0	0	-1	0
Aplikasi buruk ada keluhan tidak ada respon sama sekali	aplikasi buruk keluhan tidak respon	0	-1	0	0

Tabel 4.5 Data Hasil Pelatihan Model (Lanjutan)

Aplikasi	gak	aplikasi	tidak	-1	0	0	0
berfungsi		fungsi	susah				
nyusahin	aja,	susah					
ada	brp						
tempat	yg	hrs					
menggunakan							
ini	yg	ada					
malahan							
tmbh	susah						
□ □ □							
Terimakasih	terima	kasih	1	1	1	1	
kepada		milik aplikasi					
pemilik		linkaja	bantu				
aplikasi		tranfer	semua				
linkaja	bisa						
membantu							
tranfer							
kesemua.							

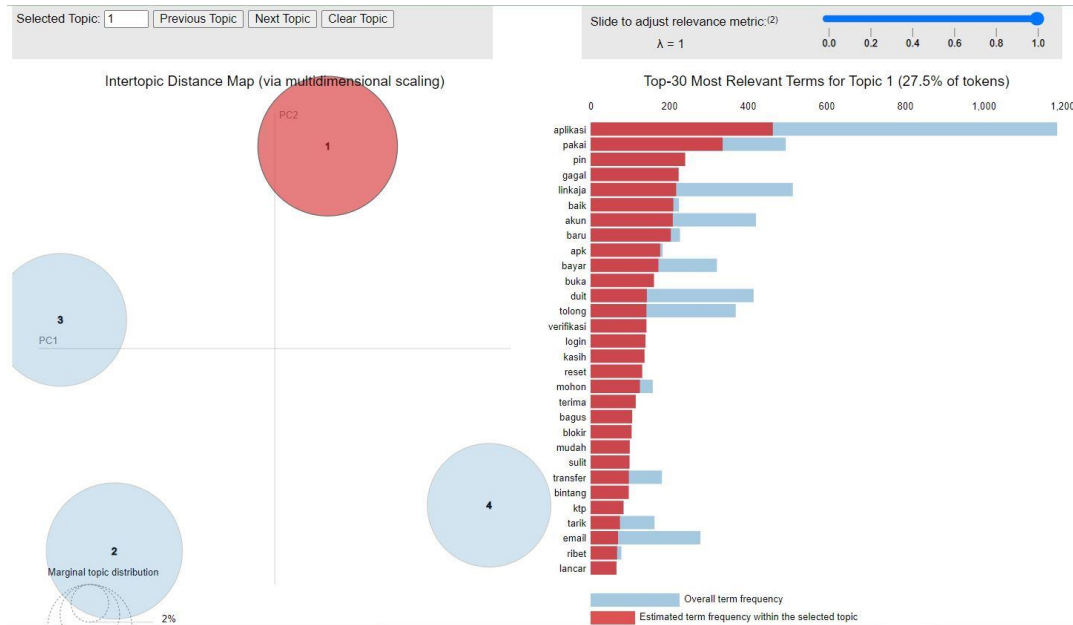
Pada Tabel 4.3 dapat dilihat hasil dari pelatihan data ditampilkan dalam bentuk tabel yang berisi ulasan pengguna, hasil dari preprocessing teks, dan sentimen dari masing-masing aspek yang dilambangkan dengan -1 untuk sentimen negatif, 0 untuk sentimen netral, dan 1 untuk sentimen positif.

4.4. Hasil Pengujian Model

4.4.1. Pengujian Pemodelan Topik

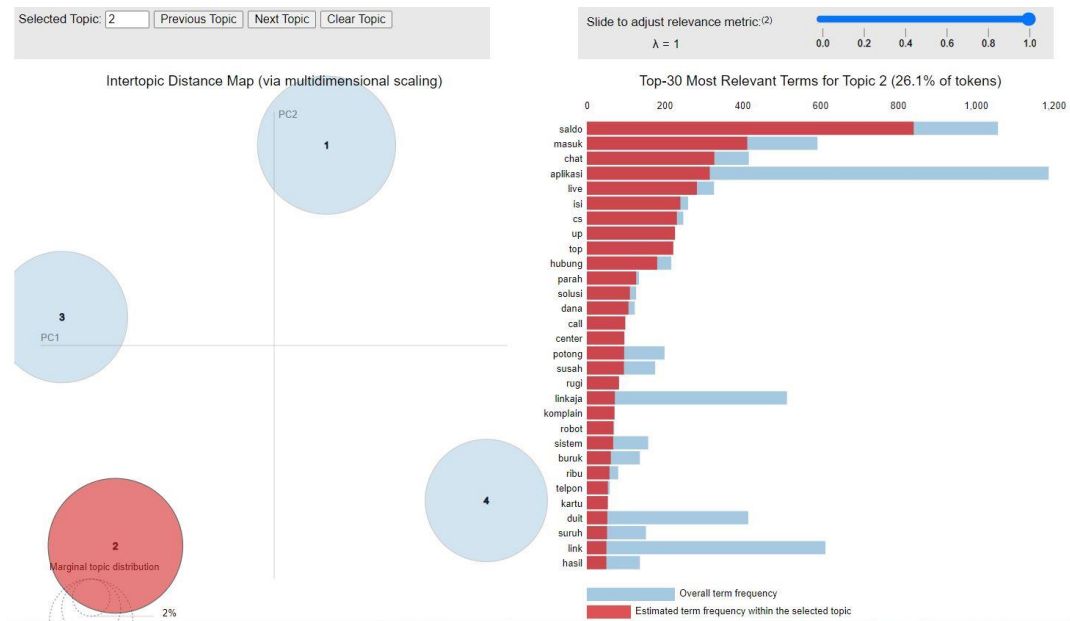
Pemodelan topik dibentuk dengan menggunakan algoritma Latent Dirichlet Allocation yang hasilnya disajikan dalam bentuk grafik visualisasi yang menampilkan keyword yang berkaitan dengan topik. Topik-topik yang diperoleh berdasarkan pemodelan topik ini akan menjadi aspek yang akan digunakan dalam analisis sentimen.

Terdapat 4 topik yang diperoleh, topik pertama yaitu kemudahan penggunaan aplikasi, topik kedua yaitu layanan pengguna, topik ketiga yaitu kualitas produk aplikasi, dan topik keempat yaitu kecepatan layanan aplikasi. Visualisasi dapat dilihat pada gambar dibawah ini.



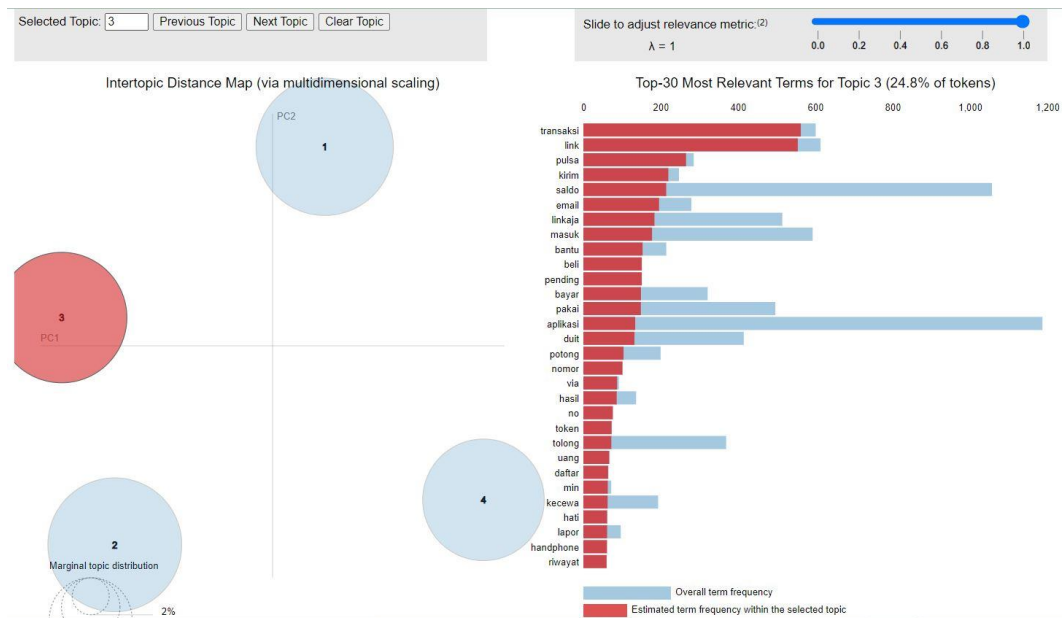
Gambar 4.10 Visualisasi LDA Topik 1

Berdasarkan grafik visualisasi di atas, topik 1 memperoleh persentase 27.5% untuk relevansi terhadap keseluruhan topik yang dibahas dalam ulasan-ulasan tersebut. Disimpulkan bahwa setiap kata-kata yang ada dalam topik 1 ini relevan untuk topik kemudahan penggunaan aplikasi karena terdapat kata sulit, ribet, lancar, gagal, dan lain-lain.



Gambar 4.11 Visualisasi LDA Topik 2

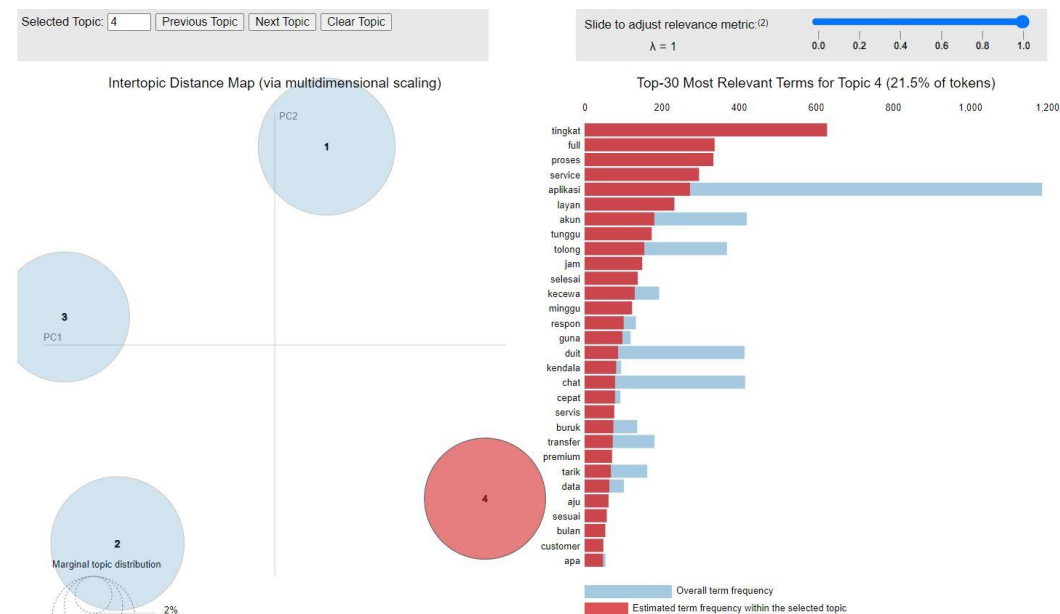
Berdasarkan grafik visualisasi di atas, topik 2 memperoleh persentase 26.1% untuk relevansi terhadap keseluruhan topik yang dibahas dalam ulasan-ulasan tersebut. Disimpulkan bahwa setiap kata-kata yang ada dalam topik 2 ini relevan untuk topik layanan pengguna karena terdapat kata cs, komplain, chat dan lain-lain.



Gambar 4.12 Visualisasi LDA Topik 3

Berdasarkan grafik visualisasi di atas, topik 3 memperoleh persentase 24.8% untuk relevansi terhadap keseluruhan topik yang dibahas dalam ulasan-ulasan tersebut.

Disimpulkan bahwa setiap kata-kata yang ada dalam topik 3 ini relevan untuk kualitas produk aplikasi karena terdapat kata pulsa, pending, transaksi dan lain-lain.



Gambar 4.13 Visualisasi LDA Topik 4

Berdasarkan grafik visualisasi di atas, topik 4 memperoleh persentase 21.5% untuk relevansi terhadap keseluruhan topik yang dibahas dalam ulasan-ulasan tersebut. Disimpulkan bahwa setiap kata-kata yang ada dalam topik 4 ini relevan untuk kecepatan layanan aplikasi karena terdapat kata service, cepat, kendala dan lain-lain.

Penentuan topik dilakukan dengan dilakukan dengan memilih 10 kata utama dari 30 kata yang memiliki kecocokan lebih besar daripada kata lainnya. Penamaan topik ini juga disesuaikan dengan jurnal aspek pengalaman pengguna pada aplikasi Fintech (Barbu et al, 2021). Pemilihan kata dapat dilihat pada tabel dibawah ini.

Tabel 4.6 Keyword topik LDA

Topik	Aspek	Kata Kunci
1	Kemudahan	“aplikasi”, “pakai”, “bagus”, “mudah”, “baik”, “sulit”, “ribet”, “gagal”, “lancar”, “blokir”
2	Layanan Pengguna	“chat”, “cs”, “solusi”, “call”, “center”, “komplain”, “hubung” “parah”, “susah”, “telpon”
3	Kualitas	“transaksi”, “pulsa”, “kirim”, “beli”, “bayar”, “pending”, “poto ng”, “saldo”, “email”, “token”

Tabel 4.7 Keyword topik LDA (Lanjutan)

4	Kecepatan	“proses”, “service”, “kecewa”, “tunggu”, “cepat”, “buruk”, “kendala”, “lambat”, “bulan”, “minggu”
---	-----------	------------------------------------------------------------------------------------------------------

4.4.2. Pengujian Model Prediksi

Setelah model dilatih menggunakan data latih, maka dilakukan pengujian terhadap data uji yang belum dikenal sebelumnya untuk melihat bagaimana model memprediksi hasil dengan baik. Hasil pengujian dapat dilihat pada tabel 4.5.

Tabel 4.8 Tabel Pengujian Model

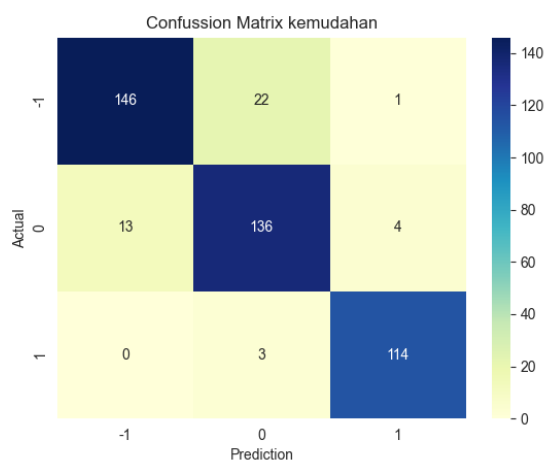
Review	Pre processing	Actual				Prediction			
		Kem udah an	Laya nan Peng guna	Kua litas	Kec epat an	Kem udah an	Laya nan Peng guna	Kua litas	Kece patan
Ada apa dgn aplikasi ini...isi saldo tp gagal trus...	aplikasi isi saldo gagal	0	0	-1	0	0	0	-1	0
Tolong saya beli pulsa udah dua hari gax masuk masuk.. Kendala ini bukan sekali dua kali terjadi.. Pulasa gax masuk.. Saldo tidak kembali lagi.. Ngadu sama cs nya sama sekali tidak ada bantuan nyah.. Apk ewalet paling busuk..	tolong beli pulsa gax masuk masuk kendala bukan kali pulasa gax masuk saldo tidak adu rekan tidak bantu ewalet busuk	0	-1	-1	0	0	-1	-1	0

Tabel 4.9 Tabel Pengujian Model (Lanjutan)

Status sudah full servis kenapa kirim uang tf tidak bisa, hubungi email. Giliran sudah di email, email tdk di pakai pusat bantuan live chat tdk membantu robot tolong perbaiki akun saya	status full servis kirim duit transfer tidak hubungi email gilir email email tidak pakai pusat bantu live chat tidak bantu robot tolong baik akun	0	-1	-1	0	0	-1	-1	0
Saya udah usulan upgrade ke full service tapi udh 3 hari sedang dalam proses aja lama bangeut gunain live chat yang bales robot h□♂	usul tingkat full service proses lama bangeut live chat balas robot	0	-1	0	-1	0	-1	0	-1
Kenapa setelah updet saya susah transaksi kirim uang dan pembayaran	updet susah transaksi kirim duit bayar	-1	0	-1	0	0	0	-1	0

4.5. Evaluasi Model

Untuk mengukur tingkat kinerja model dalam melakukan proses klasifikasi, Pada penelitian ini digunakan tabel perbandingan yaitu *confussion matrix* yang membandingkan prediksi model dengan nilai data sebenarnya. Dengan menggunakan perhitungan *precision*, *recall*, dan *f1-score* maka akan diperoleh perbandingan nilai prediksi dan aktual pada model. Berikut merupakan *confussion matrix* dari masing-masing aspek pada penelitian ini.



Gambar 4.14 Confussion Matrix aspek kemudahan

Berdasarkan Gambar 4.15, pada aspek kemudahan didapatkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) dari setiap sentimen yang ditunjukkan pada tabel 4.6, tabel 4.7, dan tabel 4.8 dibawah ini.

Tabel 4.10 Nilai Variabel Sentimen Negatif Aspek Kemudahan

No	Sentimen Negatif	Total
1.	True Positive (TP)	146
2.	True Negative (TN)	257
3.	False Positive (FP)	13
4.	False Negative (FN)	23

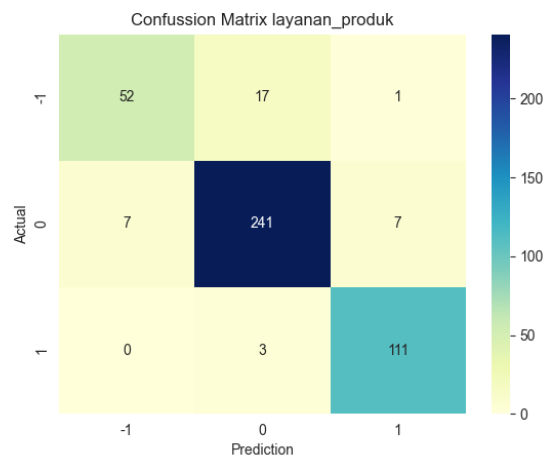
Tabel 4.11 Nilai Variabel Sentimen Netral Aspek Kemudahan

No	Sentimen Netral	Total
1.	True Positive (TP)	136
2.	True Negative (TN)	261
3.	False Positive (FP)	25
4.	False Negative (FN)	17

Tabel 4.12 Nilai Variabel Sentimen Positif Aspek Kemudahan

No	Sentimen Positif	Total
1.	True Positive (TP)	114
2.	True Negative (TN)	317
3.	False Positive (FP)	5
4.	False Negative (FN)	3

Selanjutnya *confussion matrix* aspek layanan pengguna yang ditunjukkan pada Gambar 4.16 di bawah ini.

**Gambar 4.15** Confussion Matrix aspek Layanan Pelanggan

Berdasarkan Gambar 4.16, pada aspek layanan pelanggan didapatkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) dari setiap sentimen yang ditunjukkan pada tabel 4.9, tabel 4.10, dan tabel 4.11 dibawah ini.

Tabel 4.13 Nilai Variabel Sentimen Negatif Aspek Layanan Pelanggan

No	Sentimen Negatif	Total
1.	True Positive (TP)	52
2.	True Negative (TN)	362
3.	False Positive (FP)	7
4.	False Negative (FN)	18

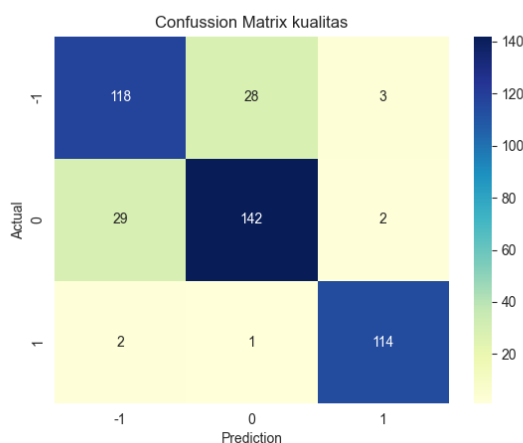
Tabel 4.14 Nilai Variabel Sentimen Netral Aspek Layanan Pelanggan

No	Sentimen Netral	Total
1.	True Positive (TP)	241
2.	True Negative (TN)	164
3.	False Positive (FP)	20
4.	False Negative (FN)	14

Tabel 4.15 Nilai Variabel Sentimen Positif Aspek Layanan Pelanggan

No	Sentimen Positif	Total
1.	True Positive (TP)	111
2.	True Negative (TN)	317
3.	False Positive (FP)	3
4.	False Negative (FN)	8

Selanjutnya *confussion matrix* aspek kualitas yang ditunjukkan pada Gambar 4.17 di bawah ini.

**Gambar 4.16** Confussion Matrix aspek Kualitas

Berdasarkan Gambar 4.17, pada aspek kualitas didapatkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) dari setiap sentimen yang ditunjukkan pada tabel 4.12, tabel 4.13, dan tabel 4.14 dibawah ini.

Tabel 4.16 Nilai Variable Sentimen Negatif Aspek Kualitas

No	Sentimen Negatif	Total
1.	True Positive (TP)	118
2.	True Negative (TN)	259
3.	False Positive (FP)	31
4.	False Negative (FN)	31

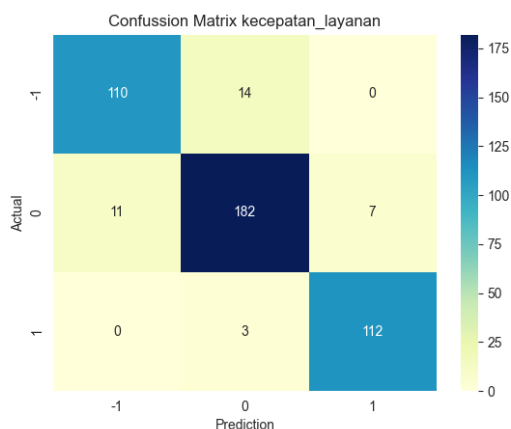
Tabel 4.17 Nilai Variabel Sentimen Netral Aspek Kualitas

No	Sentimen Netral	Total
1.	True Positive (TP)	142
2.	True Negative (TN)	237
3.	False Positive (FP)	29
4.	False Negative (FN)	31

Tabel 4.18 Nilai Variabel Sentimen Positif Aspek Kualitas

No	Sentimen Positif	Total
1.	True Positive (TP)	114
2.	True Negative (TN)	317
3.	False Positive (FP)	5
4.	False Negative (FN)	3

Selanjutnya *confussion matrix* aspek kecepatan yang ditunjukkan pada Gambar 4.18 di bawah ini.



Selanjutnya, Berdasarkan Gambar 4.18, pada aspek kecepatan didapatkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) dari setiap sentimen yang ditunjukkan pada tabel 4.15, tabel 4.16, dan tabel 4.17 dibawah ini.

Tabel 4.19 Nilai Variabel Sentimen Negatif Aspek Kecepatan

No	Sentimen Negatif	Total
1.	True Positive (TP)	110
2.	True Negative (TN)	307
3.	False Positive (FP)	11
4.	False Negative (FN)	14

Tabel 4.20 Nilai Variabel Sentimen Netral Aspek Kecepatan

No	Sentimen Netral	Total
1.	True Positive (TP)	182
2.	True Negative (TN)	222
3.	False Positive (FP)	17
4.	False Negative (FN)	18

Tabel 4.21 Nilai Variabel Sentimen Positif Aspek Kecepatan

No	Sentimen Positif	Total
1.	True Positive (TP)	112
2.	True Negative (TN)	317
3.	False Positive (FP)	7
4.	False Negative (FN)	3

Berdasarkan nilai variable dari masing-masing aspek pada tabel 4.6 sampai tabel 4.17 maka akan dihitung akurasi dari masing-masing aspek yang ditunjukkan sebagai berikut

$$\text{Akurasi Aspek Kemudahan} = \frac{TP+TN}{TP+FP+TN+FN} = \frac{396}{439} = 0.90$$

$$\text{Akurasi Aspek Layanan Pengguna} = \frac{TP+TN}{TP+FP+TN+FN} = \frac{404}{439} = 0.92$$

$$\text{Akurasi Aspek Kualitas} = \frac{TP+TN}{TP+FP+TN+FN} = \frac{374}{439} = 0.85$$

$$\text{Akurasi Aspek Kecepatan} = \frac{TP+TN}{TP+FP+TN+FN} = \frac{404}{439} = 0.92$$

Berdasarkan perhitungan akurasi dari setiap aspek, maka total akurasi dari 4 aspek adalah sebagai berikut

$$\text{Akurasi} = \frac{0.90+0.92+0.85+0.92}{4} = \frac{3.59}{4} = 0.8975 \times 100\% = 89,75\%$$

Berdasarkan rata-rata hasil akurasi yang diperoleh, hasil akurasi untuk dataset pengujian 439 ulasan dengan 4 aspek dan 3 jenis kategori sentimen dengan mencapai hasil 89,75%. Hasil akurasi pengujian yang diperoleh dari masing-masing aspek masih mengalami kesalahan dalam prediksinya sehingga menyebabkan adanya *loss*. Kesalahan model dalam melakukan pengujian dikarenakan jumlah data ulasan untuk masing-masing aspeknya masih terbatas dan kurang berpola sehingga membuat persebaran polaritas data kurang dan model mempelajari data yang terbatas.

BAB 5

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan penelitian yang telah dilakukan, maka penulis menarik kesimpulan sebagai berikut :

1. Dengan menggunakan algoritma Latent Dirichlet Allocation sebagai pemodelan topik, diperoleh empat aspek yang digunakan dalam pelabelan aspek ulasan penelitian ini, yaitu aspek kemudahan, aspek layanan, aspek kualitas, dan aspek kecepatan.
2. Pemodelan dengan menggunakan algoritma Attention-Based LSTM dalam melakukan analisis sentimen berbasis aspek mencapai performa terbaik pada model dengan batch size 32, dan unit neuron 64 dengan rata-rata akurasi model training 91% dan loss rata-rata 0.95%.
3. Pemodelan dengan menggunakan algoritma Attention-Based LSTM dalam melakukan analisis sentimen berbasis aspek menghasilkan akurasi sebesar 90% untuk aspek kemudahan, 92% untuk aspek layanan pengguna, 85% untuk aspek kualitas, dan 92% untuk aspek kecepatan dengan rata-rata akurasi keseluruhan 89,75%.

5.2. Saran

Untuk pengembangan penelitian yang lebih baik lagi, berikut merupakan beberapa saran yang dapat dipertimbangkan :

1. Keterbatasan data dalam masing-masing aspek mempengaruhi performa pengujian model dalam proses klasifikasi, sehingga dibutuhkan jumlah data yang lebih seimbang untuk mendapatkan hasil akurasi yang lebih baik.
2. Melakukan perbaikan pada kamus normalisasi untuk mendapatkan data yang lebih bersih.
3. Berkonsultasi dengan ahli Bahasa untuk membantu pelabelan data agar hasil klasifikasi sentimen lebih sesuai.

DAFTAR PUSTAKA

- Abad-Segura, E., González-Zamar, M. D., López-Meneses, E., & Vázquez-Cano, E. (2020). Financial Technology: Review of trends, approaches and management. *Mathematics*, 8(6), 1–36. <https://doi.org/10.3390/math8060951>
- Albalawi, R., Yeap, T. H., & Benyoucef, M. (2020). Using Topic Modeling Methods for Short-Text Data: A Comparative Analysis. *Frontiers in Artificial Intelligence*, 3(July), 1–14. <https://doi.org/10.3389/frai.2020.00042>
- Bahdanau, D., Cho, K. H., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15.
- Barbu, C. M., Florea, D. L., Dabija, D. C., & Barbu, M. C. R. (2021). Customer experience in fintech. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(5), 1415–1433. <https://doi.org/10.3390/jtaer16050080>
- Blei, D. M., & Stroulia, E. (2015). Latent Dirichlet Allocation: Extracting Topics from Software Engineering Data. *The Art and Science of Analyzing Software Data*, 3, 139–159. <https://doi.org/10.1016/B978-0-12-411519-4.00006-9>
- Fattahila, A. A., Irfan Amorokhman, F., Auditama, K. M., Ahmad Wijaya, K., & Romadhony, A. (2021). Indonesian Digital Wallet Sentiment Analysis Using CNN and LSTM Method. *2021 International Conference on Artificial Intelligence and Big Data Analytics, ICAIBDA 2021*, 173–178. <https://doi.org/10.1109/ICAIBDA53487.2021.9689712>
- Fitriani, A., Andreswari, R., & Darmawan, I. (2019). *Analisis Sentimen Pengguna Dompot Digital Pada Media Sosial Twitter Menggunakan Perbandingan Algoritma Klasifikasi (Studi Kasus Linkaja) Sentiment Analysis With Comparison Of Classification Algorithm Case Study : E-Wallet (Linkaja) From Social Media T.*
- Helmayanti, S. A., Hamami, F., & Fa'rifah, R. Y. (2023). Penerapan Algoritma Tf-Idf Dan Naïve Bayes Untuk Analisis Sentimen Berbasis Aspek Ulasan Aplikasi Flip Pada Google Play Store. *Jurnal Indonesia : Manajemen Informatika Dan*

- Komunikasi*, 4(3), 1822–1834. <https://doi.org/10.35870/jimik.v4i3.415>
- Ilahiyah S, & Nilogiri A. (2018). Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network _ Ilahiyah _ JUSTINDO (Jurnal Sistem dan Teknologi Informasi Indonesia). *JUSTINDO(Jurnal Sistem & Teknologi Informasi Indonesia)*, 3(2), 49–56.
- Kang, Y., & Zhou, L. (2017). RubE: Rule-based methods for extracting product features from online consumer reviews. *Information and Management*, 54(2), 166–176. <https://doi.org/10.1016/j.im.2016.05.007>
- Khaffi, M. (2023). *Sentimen Analisis Berbasis Aspek terhadap Review Aplikasi Digital Wallet Menggunakan Metode Extreme Gradient Boosting*. <https://repositori.usu.ac.id/handle/123456789/91276>
- Kumar, J., Goomer, R., & Singh, A. K. (2018). Long Short Term Memory Recurrent Neural Network (LSTM-RNN) Based Workload Forecasting Model for Cloud Datacenters. *Procedia Computer Science*, 125, 676–682. <https://doi.org/10.1016/j.procs.2017.12.087>
- Lv, Y. H., Liu, Q., & Jiang, C. Bin. (2021). Sentiment analysis measure of topic mapping in LDA-LSTM combination model. *IOP Conference Series: Earth and Environmental Science*, 702(1). <https://doi.org/10.1088/1755-1315/702/1/012015>
- Nazir, A., Rao, Y., Wu, L., & Sun, L. (2022). Issues and Challenges of Aspect-based Sentiment Analysis: A Comprehensive Survey. *IEEE Transactions on Affective Computing*, 13(2), 845–863. <https://doi.org/10.1109/TAFFC.2020.2970399>
- Nugraha, A. (2021). *Aspect Based Sentiment Analysis pada Ulasan Mobil Menggunakan Algoritma Attention-Based LSTM (Long Short Term Memory)*. <https://repositori.usu.ac.id/handle/123456789/47293>
- Priyantina, R. A., & Sarno, R. (2019). Sentiment analysis of hotel reviews using Latent Dirichlet Allocation, semantic similarity and LSTM. *International Journal of Intelligent Engineering and Systems*, 12(4), 142–155. <https://doi.org/10.22266/ijies2019.0831.14>

- Putra, K. B., & Kusumawardani, R. P. (2017). Analisis Topik Informasi Publik Media Sosial di Surabaya Menggunakan Pemodelan Latent Dirichlet Allocation (LDA). *Jurnal Teknik ITS*, 6(2), 4–9. <https://doi.org/10.12962/j23373539.v6i2.23205>
- Rofiah, S., & Setiyadi, D. (2020). Pemilihan FinTech Payment untuk Mahasiswa menggunakan Fuzzy Technique For Order Preference by Similarity of Ideal Solution. *Bina Insani Ict Journal*, 7(1), 13. <https://doi.org/10.51211/biict.v7i1.1341>
- Safitri, T. A. (2020). *The Development of Fintech in Indonesia*. 436, 666–670. <https://doi.org/10.2991/assehr.k.200529.139>
- Tian, C., Ma, J., Zhang, C., & Zhan, P. (2018). A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network. *Energies*, 11(12). <https://doi.org/10.3390/en11123493>
- Turjaman, R. M., & Budi, I. (2022). Analisis Sentimen Berbasis Aspek Marketing Mix Terhadap Ulasan Aplikasi Dompot Digital (Studi Kasus: Aplikasi Linkaja Pada Twitter). *Jurnal Darma Agung*, 30(2), 266. <https://doi.org/10.46930/ojsuda.v30i2.1672>
- Wilfan, A. F., & Martini, E. (2021). Faktor-faktor Yang Mempengaruhi Penggunaan Linkaja Berdasarkan Model Teori Utaut2 (unified Theory Of Acceptance And Use Of Technology2). *EProceedings ...*, 8(6), 7729–7741. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/management/article/view/16895%0Ahttps://openlibrarypublications.telkomuniversity.ac.id/index.php/management/article/view/16895/16612>
- Yadav, R. K., Jiao, L., Goodwin, M., & Granmo, O. C. (2021). Positionless aspect based sentiment analysis using attention mechanism[Formula presented]. *Knowledge-Based Systems*, 226, 107136. <https://doi.org/10.1016/j.knosys.2021.107136>
- Yu, Y., Liu, G., Yan, H., Li, H., & Guan, H. (2018). Attention-Based Bi-LSTM Model for Anomalous HTTP Traffic Detection. *2018 15th International Conference on Service Systems and Service Management, ICSSSM 2018*, 1–6. <https://doi.org/10.1109/ICSSSM.2018.8465034>

Zahara, S., Sugianto, & M. Bahril Ilmiddafiq. (2019). Prediksi Indeks Harga Konsumen Menggunakan Metode Long Short Term Memory (LSTM) Berbasis Cloud Computing. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 3(3), 357–363. <https://doi.org/10.29207/resti.v3i3.1086>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN,
RISET, DAN TEKNOLOGI
UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
Jalan Universitas No. 9A Gedung A, Kampus USU Medan 20155, Telepon: (061) 821007
Laman: <http://Fasilkomti.usu.ac.id>

KEPUTUSAN
DEKAN FAKULTAS ILMU KOMPUTER
DAN TEKNOLOGI INFORMASI
NOMOR : 2750/UN5.2.14.D/SK/SPB/2024
DEKAN FAKULTAS ILMU KOMPUTER
DAN TEKNOLOGI INFORMASI UNIVERSITAS SUMATERA UTARA

- Membaca** : Surat Permohonan Mahasiswa Fasilkom-TI USU tanggal 11 Juli 2024 perihal permohonan ujian skripsi:
Nama : ANNISA PUTRI DAULAY
NIM : 191402030
Program Studi : Sarjana (S-1) Teknologi Informasi
Judul Skripsi : Analisis Sentiment Berbasis Aspek Pada Aplikasi Link Aja Berdasarkan Ulasan Pengguna Dengan Menggunakan Latent Dirichlet Allocation dan Attention Based LSTM
- Memperhatikan** : Bahwa Mahasiswa tersebut telah memenuhi kewajiban untuk ikut dalam pelaksanaan Meja Hijau Skripsi Mahasiswa pada Program Studi Sarjana (S-1) Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara TA 2023/2024.
- Menimbang** : Bahwa permohonan tersebut diatas dapat disetujui dan perlu ditetapkan dengan surat keputusan
- Mengingat** : 1. Undang-undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional.
2. Peraturan Pemerintah Nomor 17 tahun 2010 tentang pengelolaan dan penyelenggara pendidikan.
3. Keputusan Rektor USU Nomor 03/UN5.1.R/SK/SPB/2021 tentang Peraturan Akademik Program Sarjana Universitas Sumatera Utara.
4. Surat Keputusan Rektor USU Nomor 1876/UN5.1.R/SK/SDM/2021 tentang pengangkatan Dekan Fasilkom-TI USU Periode 2021-2026
- MEMUTUSKAN**
- Menetapkan** :
- Pertama** : Membentuk dan mengangkat Tim Penguji Skripsi mahasiswa sebagai berikut:
- Ketua : Dr. Erna Budhiarti Nababan M.IT
NIP: 196210262017042001
- Sekretaris : Seniman S.Kom., M.Kom.
NIP: 198705252014041001
- Anggota Penguji : Dedy Arisandi ST., M.Kom.
NIP: 197908312009121002
- Anggota Penguji : Fanindia Purnamasari S.TLM.IT
NIP: 198908172019032023
- Moderator : -
- Panitera : -
- Kedua** : Segala biaya yang diperlukan untuk pelaksanaan kegiatan ini dibebankan pada Dana Penerimaan Bukan Pajak (PNPB) Fasilkom-TI USU Tahun 2024.
- Ketiga** : Keputusan ini berlaku sejak tanggal ditetapkan dengan ketentuan bahwa segala sesuatunya akan diperbaiki sebagaimana mestinya apabila dikemudian hari terdapat kekeliruan dalam surat keputusan ini.

Tembusan :

1. Ketua Program Studi Sarjana (S-1) Teknologi Informasi
2. Yang bersangkutan
3. Arsip

Medan
Ditandatangani secara elektronik oleh:
Dekan



Maya Silvi Lydia
NIP 197401272002122001

