

**ANALISIS SENTIMEN BERBASIS ASPEK PADA
ULASAN E-WALLET DANA DENGAN
ALGORITMA CNN-BiLSTM**

SKRIPSI

RINI HARYATI

181402088



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

MEDAN

2024

**ANALISIS SENTIMEN BERBASIS ASPEK PADA
ULASAN E-WALLET DANA DENGAN
ALGORITMA CNN-BiLSTM**

SKRIPSI

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

RINI HARYATI

181402088



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

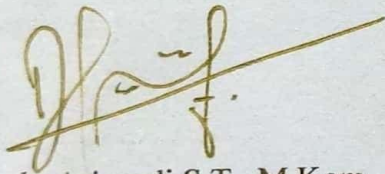
PERSETUJUAN

Judul : Analisis Sentimen Berbasis Aspek Pada Ulasan E-Wallet Dana Dengan Algoritma CNN-BiLSTM
Kategori : Skripsi
Nama : Rini Haryati
Nomor Induk Mahasiswa : 181402088
Program Studi : Teknologi Informasi
Fakultas : Fakultas Ilmu Komputer Dan Teknologi Informasi
Universitas Sumatera Utara

Medan, 10 Januari 2024

Komisi Pembimbing:

Pembimbing 2



Dedy Arisandi S.T., M.Kom.

NIP. 197908312009121002

Pembimbing 1



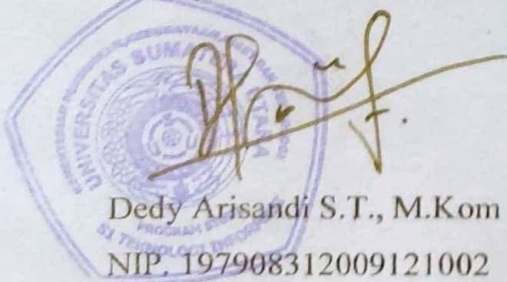
Dr. Erna Budhiarti Nababan M.IT

NIP. 196210262017042001

Diketahui/disetujui oleh

Program Studi SI Teknologi Informasi

Ketua,



Dedy Arisandi S.T., M.Kom

NIP. 197908312009121002

PERNYATAAN**ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN E-WALLET DANA
DENGAN ALGORITMA CNN-BiLSTM****SKRIPSI**

Saya menyatakan dengan tegas bahwa skripsi yang telah saya susun merupakan hasil orisinalitas dan karya intelektual pribadi saya. Meskipun demikian, untuk beberapa kutipan dan ringkasan yang terdapat dalam skripsi ini, setiap sumbernya telah saya nyatakan dengan jelas.

Medan, 10 Januari 2024



Rini Haryati

181402088

UCAPAN TERIMA KASIH

Segala puji dan syukur penulis haturkan kepada Allah SWT, yang dengan limpahan rahmat-Nya, telah memberikan kemampuan kepada penulis untuk menyelesaikan tugas akhir ini. Tugas akhir ini disusun sebagai syarat untuk memperoleh gelar Sarjana Komputer dari Program Studi S1 Teknologi Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara..

Penulisan skripsi ini berhasil diselesaikan berkat dukungan dan doa yang diberikan tanpa henti oleh keluarga penulis, yaitu ayah Azhar, ibu Yetmawati, serta saudara dan saudari penulis Yeza Haryati dan Muhammad Fauzan, serta saudara ipar penulis Arif Fauzi Hidayatulah. Penulis juga ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada semua pihak yang terlibat dan memberikan bantuan selama perjalanan perkuliahan hingga penyelesaian penyusunan skripsi ini.


Adapun pada kesempatan ini penulis ingin menyampaikan terima kasih kepada:

1. Bapak Prof. Dr. Mulyanto Amin, M.Si., selaku Rektor Universitas Sumatera Utara.
2. Ibu Dr. Maya Silvi Lydia, B.Sc., M.Sc., selaku Dekan Fasilkom-TI USU.
3. Bapak Dedy Arisandi, S.T., M.Kom., selaku Ketua Program Studi S1 Teknologi Informasi Universitas Sumatera Utara.
4. Ibu Dr. Erna Budhiarti Nababan M.IT. selaku Dosen Pembimbing I dan juga Bapak Dedy Arisandi S.T., M.Kom., selaku Dosen Pembimbing II, yang telah memberikan bimbingan, kritik, saran serta arahan kepada penulis.
5. Ibu Sarah Purnamawati S.T., M.Sc, selaku dosen pembimbing I saya dan Bapak Niskarto Zendrato S.Kom., M.Kom selaku dosen pembimbing II saya, yang telah memberikan bimbingan, kritik, saran serta arahan kepada penulis.
6. Seluruh Dosen, Staff, dan Pegawai di Program Studi S1 Teknologi Informasi yang tidak hanya berbagi pengetahuan yang sangat bermanfaat kepada penulis, tetapi juga memberikan dukungan dalam semua aspek administratif selama masa perkuliahan dan proses penulisan skripsi.
7. Sahabat seperjuangan penulis, Indriyani br Sembiring, Indah, dan Suci Khairiah yang telah mendampingi dan menjadi penyemangat sepanjang perjalanan penyelesaian tugas akhir, dari awal proses penulisan hingga tahap akhir.

8. Arinda Bella Putri Manik, Fitri Khairina Nst, Vina Yusdianty, Ade Ruqayah Lubis, Ummi Nadrah Rangkuti, Siti Khadijah dan Sri Wahyuni selaku sahabat penulis semasa Madrasah Aliyah (MA) yang telah memberikan semangat penuh kepada penulis sehingga penulis mampu menyelesaikan skripsi.
9. Tasya Arifin Tanjung dan Adriati Nur Azimmi sebagai sahabat sejak Sekolah Menengah Pertama (SMP) hingga saat ini, mereka selalu memberikan dukungan penuh kepada penulis, mendorong agar skripsi dapat terselesaikan dengan cepat.
10. Sahabat penulis Putri Andesta yang senantiasa memberikan semangat dan doa kepada penulis.
11. Teman-teman dari program studi Teknologi Informasi khususnya angkatan 2018 yang membantu penulis dalam menyelesaikan perkuliahan.

Kiranya Allah SWT. sentiasa bersama dan menyalurkan limpahan berkah kepada semua orang yang turut serta, memberikan bantuan, serta memberikan dukungan kepada penulis dalam menyelesaikan penulisan skripsi ini.

Medan, 10 Januari 2024



Penulis

ABSTRAK

E-Wallet merupakan bentuk kemudahan di era digital saat ini. Dengan munculnya e-wallet, peran uang tunai sebagai sarana pembayaran mengalami penggantian oleh metode pembayaran non-tunai yang dinilai lebih efektif dan efisien. Teknologi e-wallet lebih mudah digunakan karena pengguna tidak perlu melakukan kontak langsung untuk melakukan pembayaran. Salah satu e-wallet yang cukup populer di Indonesia saat ini adalah aplikasi dompet digital DANA Indonesia. Tingkat kepuasan pengguna layanan DANA bervariasi, hal ini dipengaruhi oleh pengalaman yang dirasakan pengguna selama menggunakan aplikasi. Dibalik kemudahan yang ditawarkan, teknologi dompet digital memiliki beberapa kasus seperti tidak bertambahnya saldo setelah melakukan pengisian saldo, performa aplikasi yang terkesan lambat, sulitnya proses registrasi dan verifikasi akun. Untuk mengatasi permasalahan ini diperlukan satu pendekatan untuk membantu perusahaan dalam meningkatkan pelayanannya, penelitian ini menerapkan pendekatan menggunakan model *Convolutional Neural Network Bi-Directional Long Short Term Memory* (CNN-BiLSTM) dengan *word embedding fastText*. Hasil pemodelan CNN-BiLSTM dalam mengklasifikasi ulasan pengguna E-Wallet Dana berdasarkan aspek Transaksi, Pelayanan, Account, Access dan performa dan aspek positif, negative dan netral model CNN-BiLSTM memperoleh hasil rata rata akurasi sebesar 88.7%

Kata Kunci : *Dompot Digital, Analisis Sentimen Berbasis Aspect, Dana, CNN-BiLSTM.*

ASPECT-BASED SENTIMENT ANALYSIS ON DANA E-WALLET REVIEWS USING CNN-BiLSTM ALGORITHM

ABSTRACT

E-Wallet represents a convenience in the current digital era. With the emergence of e-wallets, the role of cash as a means of payment has been replaced by non-cash payment methods that are considered more effective and efficient. E-wallet technology is easier to use because users do not need to make direct contact to make payments. One of the popular e-wallet applications in Indonesia is the DANA digital wallet. User satisfaction levels with the DANA service vary, influenced by the user's experiences while using the application. Despite the convenience offered, digital wallet technology has encountered several issues such as the failure to update the balance after a top-up, perceived slow application performance, and difficulties in the registration and verification process. To address these challenges, an approach is needed to assist the company in improving its services. This research applies a Convolutional Neural Network Bi-Directional Long Short Term Memory (CNN-BiLSTM) model with fastText word embeddings. The results of the CNN-BiLSTM modeling in classifying DANA E-Wallet user reviews based on Transaction, Service, Account, Access, and Performance aspects, as well as positive, negative, and neutral aspects, show an average accuracy of 88.7%.

Keywords : *E-Wallet, Aspect Based Sentiment Analysis, Dana, CNN-BiLSTM*

DAFTAR ISI

PERSETUJUAN	iii
PERNYATAAN	iv
UCAPAN TERIMA KASIH	v
ABSTRAK	vii
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 Text Mining	7
2.2 Sentiment Analysis	7
2.3 Aspect Based Sentiment Analysis	8
2.4 Word Embedding	9
2.5 Convolutional Neural Network Bi-Directional Long Short Term Memory (CNN-BiLSTM)	9
2.6 Penelitian Terdahulu	11
2.7 Perbedaan Penelitian	14
BAB 3 ANALISIS DAN PERANCANGAN SISTEM	15
3.1 Dataset	15
3.2 Arsitektur Umum	16
3.2.1 Input	16
3.2.2 Pre-processing	19

3.2.3 Word Embedding	29
3.2.4 Model Building	31
3.3 Perancangan Antarmuka Sistem	32
3.3.1 Rancangan Halaman Beranda	32
3.3.2 Rancangan Tampilan Halaman Training Data	33
3.3.3 Rancangan Tampilan Halaman Testing	34
3.3.4 Rancangan Tampilan Halaman User Input	35
3.4 Metode Evaluasi	36
BAB 4 IMPLEMENTASI DAN PENGUJIAN	39
4.1 Implementasi Sistem	39
4.1.1 Spesifikasi Perangkat Keras dan Perangkat Lunak	39
4.1.2 Implementasi Perancangan Tampilan Antarmuka	39
4.2 Hasil Pengujian Sistem	46
4.3 Evaluasi	52
BAB 5 KESIMPULAN DAN SARAN	60
5.1 Kesimpulan	60
5.2 Saran	60
DAFTAR PUSTAKA	61

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	13
Tabel 2. 1 Penelitian Terdahulu (Lanjutan)	14
Tabel 3. 1 Hasil scrapping ulasan pengguna Dana	15
Tabel 3. 2 Pembagian Dataset	16
Tabel 3. 3 Kata kunci dari setiap aspek	18
Tabel 3. 4 Contoh data ulasan pengguna dan label	18
Tabel 3. 4 Contoh data ulasan pengguna dan label (Lanjutan)	19
Tabel 3. 5 Confussion Matrix	37
Tabel 4. 1 Data Training	47
Tabel 4. 2 Hasil Preprocessing	48
Tabel 4. 3 Hasil Pengujian	50
Tabel 4. 4 Nilai Variabel Sentimen Negatif dari Aspek Transaksi	53
Tabel 4. 5 Nilai Variabel Sentimen Positif dari Aspek Transaksi	53
Tabel 4. 6 Nilai Variabel Sentimen Netral dari Aspek Transaksi	53
Tabel 4. 7 Nilai Variabel Sentimen Negatif dari Aspek Pelayanan	54
Tabel 4. 8 Nilai Variabel Sentimen Positif dari Aspek Pelayanan	54
Tabel 4. 9 Nilai Variabel Sentimen Netral dari Aspek Pelayanan	55
Tabel 4. 10 Nilai Variabel Sentimen Negatif dari Aspek Account	55
Tabel 4. 11 Nilai Variabel Sentimen Positif dari Aspek Account	56
Tabel 4. 12 Nilai Variabel Sentimen Netral dari Aspek Account	56
Tabel 4. 13 Nilai Variabel Sentimen Negatif dari Aspek Access	57
Tabel 4. 14 Nilai Variabel Sentimen Positif dari Aspek Access	57
Tabel 4. 15 Nilai Variabel Sentimen Netral dari Aspek Access	57
Tabel 4. 16 Nilai Variabel Sentimen Negatif dari Aspek Performa	58
Tabel 4. 17 Nilai Variabel Sentimen Positif dari Aspek Performa	58
Tabel 4. 18 Nilai Variabel Sentimen Netral dari Aspek Performa	59

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur Umum CNN (Ari Bangsa et al., 2020)	10
Gambar 2. 2 Arsitektur umum BiLSTM (Puteri, Dian., 2023).	10
Gambar 2. 3 Arsitektur umum CNN-BiLSTM.	11
Gambar 3. 1 Arsitektur Umum	17
Gambar 3. 2 Embedding Matrix	30
Gambar 3. 3 Gambar Rancangan Halaman Beranda	33
Gambar 3. 4 Rancangan Tampilan Halaman Training Data	34
Gambar 3. 5 Rancangan Tampilan Halaman Testing	35
Gambar 3. 6 Rancangan Tampilan Halaman User Input	36
Gambar 4. 1 Tampilan Halaman Utama	40
Gambar 4. 2. Tampilan Halaman Training	41
Gambar 4. 3. Tampilan Halaman Training (preprocessing)	41
Gambar 4. 4. Tampilan Hasil Training (accuration dan loss aspek transaksi)	42
Gambar 4. 5. Tampilan Hasil Training (accuration dan loss aspek pelayanan)	42
Gambar 4. 6. Tampilan Hasil Training (accuration dan loss aspek account)	43
Gambar 4. 7. Tampilan Hasil Training (accuration dan loss aspek access)	43
Gambar 4. 8. Tampilan Hasil Training (accuration dan loss aspek performa)	44
Gambar 4. 9 Tampilan Halaman Testing	44
Gambar 4. 10. Tampilan Halaman Testing (hasil prediksi)	45
Gambar 4. 11. Tampilan Halaman User Input	45
Gambar 4. 12. Tampilan Halaman User Input (sentimen negatif dan netral)	46
Gambar 4. 13. Tampilan Halaman User Input (sentiment positif dan netral)	46
Gambar 4. 14 Confussion Matrix Aspek Transaksi	52
Gambar 4. 15 Confussion Matrix Aspek Pelayanan	54
Gambar 4. 16 Confussion Matrix Aspek Account	55
Gambar 4. 17 Confussion Matrix Aspek access	56
Gambar 4. 18 Confussion Matrix Aspek Performa.	58

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Dompot digital atau electronic wallet (e-wallet) merupakan salah satu bentuk kemudahan di era digital saat ini. Dengan kemunculan e-wallet peran uang tunai sebagai alat pembayaran tergantikan dengan bentuk pembayaran non-tunai yang lebih efektif dan efisien.

Berdasarkan data bank Indonesia, pada tahun 2018 tercatat transaksi e-wallet di Indonesia mencapai Rp 21,3 triliun dan nilai transaksi diperkirakan akan terus tumbuh menjadi Rp 355,7 triliun pada tahun 2023. Teknologi e-wallet lebih mudah digunakan karena pengguna tidak perlu melakukan kontak langsung untuk melakukan pembayaran. Namun, dibalik kemudahan yang ditawarkan, teknologi dompet digital memiliki beberapa kasus seperti tidak bertambahnya saldo setelah melakukan pengisian saldo, performa aplikasi yang terkesan lambat, sulitnya proses registrasi dan verifikasi akun.

Salah satu e-wallet yang cukup populer di Indonesia saat ini adalah aplikasi dompet digital DANA Indonesia. Tingkat kepuasan pengguna layanan DANA bervariasi, hal ini dipengaruhi oleh pengalaman yang dirasakan pengguna selama menggunakan aplikasi. Pengalaman pengguna dalam menggunakan layanan DANA dapat ditemukan dalam kolom komentar yang tersedia di Google Playstore, cuitan pengguna di twitter, status pengguna facebook dan masih banyak lainnya. Opini pengguna ini menjadi bahan bagi perusahaan DANA untuk di evaluasi demi peningkatan kualitas layanan. Perusahaan diminta mengakses, menganalisis dan memahami opini pengguna dalam jumlah yang banyak untuk mendapatkan hasil akhir permasalahan yang dialami oleh pengguna. proses analisis ini akan membutuhkan waktu dan tenaga jika dilakukan secara manual. Oleh karena itu, metode yang tepat untuk mengatasi hal tersebut adalah dengan melakukan analisis sentimen berbasis aspek terhadap ulasan pengguna. Penelitian tentang analisis sentimen berbasis aspek pada e-wallet sebelumnya telah dilakukan oleh Rifqy

Mikoriza Turjaman dan Indra Budi dengan objek penelitian aplikasi dompet digital LinkAja pada twitter. Penelitian ini menggunakan teori marketing mix 4P dimana faktor yang mempengaruhi keberhasilan pemasaran barang atau jasa terdiri dari produk, harga, tempat dan promosi. Factor ini digunakan sebagai dasar dalam menentukan aspek apa saja yang dapat ditingkatkan dan diperbaiki oleh LinkAja untuk mendapatkan kenyamanan pelanggan dalam menggunakan layanan. Data cuitan yang digunakan sebanyak 4057 data dan pemodelan yang digunakan adalah algoritma *Support Vector Machine* (Turjaman & Indonesia, 2022).

Penelitian selanjutnya dilakukan oleh (Saraswita et al., 2021) menggunakan *Support Vector Machine* dan *Recursive Feature Elimination* untuk analisis sentiment e-wallet di twitter. Hasil evaluasi dari metode SVM didapatkan akurasi sebesar 74% dan meningkat saat metode SVM dan RFE digabungkan. Pada RFE seleksi fitur dilakukan dengan membuang peringkat terendah dari fitur yaitu sebanyak 500 fitur sampai dengan 4000 fitur. Sehingga didapatkan akurasi sebesar 81%.

Kemudian (Fattahila, n.d.) melakukan penelitian analisis sentiment berbasis aspek terhadap beberapa dompet digital Indonesia menggunakan metode CNNLSTM. Objek yang digunakan adalah e-wallet OVO, DANA, LinkAja dan Sakuku. Untuk melakukan klasifikasi pada data yang telah melalui preprocessing dilakukan metode Word Cloud dimana metode ini digunakan untuk melakukan highlight pada kata dan frasa populer. Dalam penentuan aspek digunakan kata dengan jumlah kemunculan terbanyak dari hasil crawling data dan hasil kata dari Word Cloud sehingga didapatkan aspek berupa Pengaksesan, Transaksi, Akun, Pelayanan dan Performa. Berdasarkan hasil prediksi dengan menggunakan CNN-LSTM aplikasi e-wallet DANA memiliki 90.7% ulasan negatif dan 9.29% ulasan positif dari total seluruh ulasan berjumlah 3164 data. Aplikasi e-wallet OVO memiliki 91.5% ulasan negatif dan 8.51% ulasan positif dari total seluruh ulasan yang berjumlah 2009 data. Aplikasi e-wallet LinkAja memiliki 81% ulasan bersifat negatif dan 19% ulasan positif dari total ulasan berjumlah 1580 data. Kemudian aplikasi Sakuku memiliki 89% ulasan negatif dan 11% ulasan positif dari total seluruh ulasan berjumlah 844 data. Pemodelan CNN-LSTM yang dirancang menghasilkan akurasi 81%.

Penelitian tentang sentimen analisis teks menggunakan model CNN-BiLSTM sebelumnya telah dilakukan oleh (Zhou & Long, 2018) melakukan penelitian analisis sentimen teks berbahasa mandarin menggunakan model CNN-BiLSTM. Penelitian ini menunjukkan bahwa model yang diusulkan dapat menyelesaikan klasifikasi sentimen dengan sangat baik. CNN mampu mengekstraksi fitur urutan dari informasi dan dapat mempertimbangkan hubungan antar fitur. BiLSTM tidak hanya menyelesaikan masalah ketergantungan jangka panjang, tetapi juga mempertimbangkan konteks teks pada saat yang bersamaan. Dataset yang digunakan sebanyak 21.105 ulasan yang dibagi menjadi dua kategori yaitu emosi positif dan emosi negatif, dengan jumlah dataset positif sebanyak 10.677 dan dataset negative sebanyak 10.428.

Penelitian ini juga melakukan perbandingan dengan 4 algoritma yaitu CNN, LSTM, CNN-LSTM, dan CNN-BiLSTM. Untuk teks positif nilai p terlihat unggul pada model CNN-BiLSTM dengan akurasi 89.98% teks positif nilai r unggul pada model CNN-LSTM dengan akurasi 92.94% dan untuk teks positif nilai f1 unggul pada model CNN-BiLSTM dengan akurasi 91.29% sedangkan untuk teks negative nilai p unggul pada model LSTM dengan skor 93.06% untuk teks negative nilai r unggul pada model CNN-BiLSTM dengan akurasi sebesar 89.43% dan untuk teks negative nilai f1 unggul pada model CNN-BiLSTM dengan akurasi sebesar 90.81%.

Berdasarkan latar belakang yang telah diuraikan, peneliti memutuskan untuk melakukan penelitian dengan metode algoritma *Convolutional Neural Network – Bidirectional Long Short Term Memory* yang berjudul **“Analisis Sentimen Berbasis Aspek Pada Ulasan E-Wallet DANA Dengan Algoritma (CNN-BiLSTM)”**.

1.2 Rumusan Masalah

Penggunaan e-wallet sebagai salah satu bentuk pembayaran cashless semakin marak seiring dengan perkembangan di era digital. Aplikasi DANA sebagai salah satu e-wallet yang diminati oleh masyarakat masih memiliki banyak kekurangan diantaranya tidak bertambahnya saldo setelah pengguna melakukan pengisian saldo, proses login dan registrasi yang sering gagal, aplikasi yang terkesan lamban, dan proses upgrade akun yang sering gagal. Hal ini dapat ditemukan dalam ulasan

pengguna aplikasi DANA yang terdapat pada aplikasi *Google Play Store*. Oleh karena itu diperlukan satu pendekatan untuk mengklasifikasi ulasan pengguna.

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk menganalisis keluhan pengguna DANA serta mengimplementasikan algoritma CNN-BiLSTM untuk mengklasifikasi sentiment masyarakat berdasarkan aspek yang telah ditentukan terhadap layanan DANA.

1.4 Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan masalah dilakukan untuk mempersempit cakupan pembahasan pada pokok permasalahan penelitian saja. Adapun batasan masalah dalam penelitian ini sebagai berikut:

1. Dataset yang digunakan untuk proses *training* dan *testing* berasal dari hasil *scraping* di *Google Play Store*.
2. Dataset yang digunakan merupakan komentar berbahasa Indonesia
3. Identifikasi ulasan dilakukan terhadap *semantic* kata, tidak terhadap emoji, angka, dan karakter.
4. Aspek yang diuji meliputi *access*, *transaction*, *user account*, *customer service*, dan *performance*.
5. Sentimen yang diuji meliputi positif, negatif, dan netral.

1.5 Manfaat Penelitian

Adapun manfaat penelitian ini adalah:

1. Menganalisis dan mengklasifikasi sentiment dari layanan aplikasi DANA berdasarkan aspek *access*, *transaction*, *user account*, *customer service*, dan *performance*.
2. Mengetahui tingkat kemampuan algoritma CNN-BiLSTM dalam mengklasifikasi sentiment berbasis aspek dari layanan DANA Dompet Digital Indonesia.

1.6 Metodologi Penelitian

Yang dilakukan pada proses penelitian ini yaitu:

1. Studi Literatur

Pada tahapan ini dilakukan studi literatur dari beberapa sumber untuk memperoleh teori yang relevan dengan penelitian dan mendapatkan gambaran dari penelitian yang telah dilakukan sebelumnya. Referensi penelitian diperoleh dari jurnal, buku, dan sumber referensi lainnya yang berkaitan dengan *Aspect Based Sentiment Analysis*, *FastText* dan algoritma CNN-BiLSTM.

2. Analisis Permasalahan

Pada tahapan ini dilakukan analisis terhadap hasil dari tahapan sebelumnya, bahan yang telah dikumpulkan akan dianalisis untuk mendapatkan gambaran mengenai metode CNN-BiLSTM dalam menganalisis sentiment berbasis aspek terhadap ulasan pengguna melalui Google Play Store tentang layanan DANA.

3. Perancangan Sistem

Pada tahapan ini dilakukan perancangan sistem berdasarkan hasil analisis literatur yang telah dipelajari. Kemudian dilanjutkan dengan merancang arsitektur umum, mengumpulkan data, dan merancang antarmuka sistem.

4. Implementasi Sistem

Pada tahapan ini melakukan penerapan rancangan yang telah ditetapkan sebelumnya untuk membangun sistem dengan menggunakan CNN-BiLSTM.

5. Pengujian Sistem

Pada tahap ini dilakukan uji kinerja sistem yang telah dikembangkan guna memastikan sistem berjalan dengan baik.

6. Penyusunan Laporan

Pada tahap ini dilakukan penyusunan laporan menyeluruh dari seluruh rangkaian penelitian yang telah dilaksanakan, disertai dengan dokumentasi yang diperoleh selama proses penelitian.

1.7 Sistematika Penulisan

Adapun sistematika penulisan dari penelitian ini di rincikan sebagai berikut:

BAB 1: PENDAHULUAN

Bab pendahuluan membahas mengenai latar belakang pemilihan judul skripsi “Analisis Sentimen Berbasis Aspek pada Ulasan E-Wallet Dana dengan Algoritma CNN-BiLSTM”, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, metodologi penelitian dan sistematika penulisan.

BAB 2: LANDASAN TEORI

Bab landasan teori membahas mengenai teori terkait untuk memahami permasalahan yang dibahas dalam penelitian ini. Bab ini menjelaskan mengenai *Sentiment Analysis*, *Word Embedding*, *Convolutional Neural Network BiDirectional Long Short Term Memory* (CNN-BiLSTM)

BAB 3: ANALISIS DAN PERANCANGAN SISTEM

Bab analisis dan perancangan sistem membahas mengenai proses arsitektur umum, analisis dan penerapan metode CNN-BiLSTM dalam menganalisis sentimen berbasis aspek layanan dompet digital DANA.

BAB 4: IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab implementasi dan pengujian sistem menjelaskan tentang penerapan analisis dan perancangan yang telah disusun pada Bab 3 dan memaparkan hasil dari pengujian sistem yang telah dibangun.

BAB 5: KESIMPULAN DAN SARAN

Bab kesimpulan dan saran membahas mengenai kesimpulan yang diperoleh dari penelitian yang telah dilakukan dan saran untuk penelitian yang akan dilakukan selanjutnya.

BAB 2

LANDASAN TEORI

2.1 Text Mining

Text Mining merupakan proses menemukan atau penggalian informasi yang bermanfaat dari data teks yang tidak terstruktur. Proses penambangan teks sama dengan penambangan data, kecuali alat penambangan data dirancang untuk menangani data terstruktur sedangkan penambangan teks dapat menangani kumpulan data tidak terstruktur atau semi-terstruktur. *Text mining* digunakan untuk menemukan informasi baru yang sebelumnya tidak teridentifikasi dari berbagai sumber tertulis. Penambangan teks adalah bidang penelitian ilmu komputer yang mencoba memecahkan masalah yang terjadi di bidang penambangan data, pembelajaran mesin, ekstraksi informasi, pemrosesan bahasa alami, pencarian informasi, manajemen pengetahuan, dan klasifikasi. (Dr. S. Vijayarani, Ms. J. Ilamathi, 2015). Data teks yang tidak terstruktur memerlukan tahapan awal untuk mempersiapkan teks dapat diubah menjadi lebih terstruktur. Salah satu implementasi dari *text mining* adalah text preprocessing. Pada tahapan text preprocessing dilakukan seleksi data yang akan diproses pada setiap data.

2.2 Sentiment Analysis

Analisis sentimen teks merupakan teknologi yang signifikan dalam bidang *Natural Language Processing* (NLP), yang bertujuan untuk mengidentifikasi dan mengekstrak informasi subyektif dari data teks, khususnya dalam konteks ulasan. (Zhou & Long, 2018). Pertumbuhan pesat bidang ini bertepatan dengan perkembangan media sosial Web untuk review, forum diskusi, blog, mikroblog, twitter, dan jejaring sosial. Sejak awal tahun 2000, analisis sentimen telah berkembang menjadi salah satu bidang penelitian yang paling dinamis dalam pemrosesan bahasa alami.

Hal ini juga dipelajari secara luas dalam data penambangan, penambangan Web, penambangan teks, dan pengambilan informasi. Bahkan, telah menyebar dari

ilmu komputer ke ilmu manajemen dan ilmu sosial seperti pemasaran, keuangan, ilmu politik, komunikasi, ilmu kesehatan, dan bahkan sejarah, karena pentingnya bagi bisnis dan masyarakat secara keseluruhan. Hal ini disebabkan oleh fakta bahwa opini adalah pusat untuk hampir semua aktivitas manusia dan merupakan pengaruh utama dari perilaku manusia. Keyakinan dan persepsi tentang realitas, dan pilihan yang di buat, sebagian besar dikondisikan pada bagaimana orang lain melihat dan menilai dunia. Hal ini tidak hanya berlaku untuk individu tetapi juga berlaku untuk organisasi. (Zhang et al., 2018)

Sentiment analysis atau yang bisa juga disebut dengan penambangan opini adalah bidang studi yang menganalisis pendapat orang, sentimen, evaluasi, penilaian, sikap dan emosi terhadap entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik serta atributnya. Objek dari sentiment analysis biasanya berupa layanan atau produk yang sangat diminati oleh orang-orang sehingga mereka menaruh sentiment terhadapnya. Umumnya sentiment analysis dianggap sebagai polaritas pendapat apakah seseorang telah mengungkapkan sentimen positif, negative atau netral tentang suatu peristiwa (Liu, 2012).

Pengaplikasian sentiment analysis telah tersebar luas hampir di setiap area. Kelompok berpengaruh dan organisasi bisnis seperti Google, Microsoft, SAP dan SAS telah merancang aplikasi sentiment analysis untuk mendukung mereka dalam mengambil keputusan, membantu dalam mengembangkan bisnis yang lebih baik serta untuk melacak dan memprediksi tren pasar yang sedang berkembang. Menurut penelitian, sentiment analysis secara umum dikategorikan menjadi tiga tingkat yaitu, tingkat dokumen, tingkat kalimat, tingkat aspect sentiment analysis. (Nazir et al., 2022).

2.3 Aspect Based Sentiment Analysis

Analisis sentimen berbasis aspek (ABSA) membantu memahami permasalahan sentimen analisis lebih baik, karena langsung berfokus pada sentimen daripada struktur bahasa, dimana, suatu aspek terkait dengan entitas dan dasar konsep suatu aspek yang tidak hanya terbatas pada penilaian tetapi juga meluas ke arah pemikiran, sudut pandang, cara berikir, serta tema yang mendasari atau pengaruh sosial terhadap suatu kejadian. (Nazir et al., 2022).

Analisis sentimen berbasis aspek memiliki tujuan untuk mengidentifikasi polaritas sentimen seperti positif, negatif, dan netral dari satu kalimat terhadap kata sasaran. *Aspect Based Sentiment Analysis* (ABSA) mencakup dua sub tugas yaitu deteksi aspek dan klasifikasi sentimen. Penelitian ini menggunakan aspek identifikasi untuk mendefinisikan penyematan aspek, karena aspek yang digunakan sudah ditentukan sebelumnya. ABSA adalah tugas mendasar dalam pemrosesan bahasa alami. ABSA bertujuan untuk mengidentifikasi polaritas sentimen dari kalimat yang diekspresikan terhadap target, yang merupakan salah satu aspek tertentu dari entitas tertentu. Meskipun metode ini telah mencapai hasil yang lebih baik, rekayasa fitur memerlukan banyak pelabelan manual, yang memakan waktu dan tenaga.

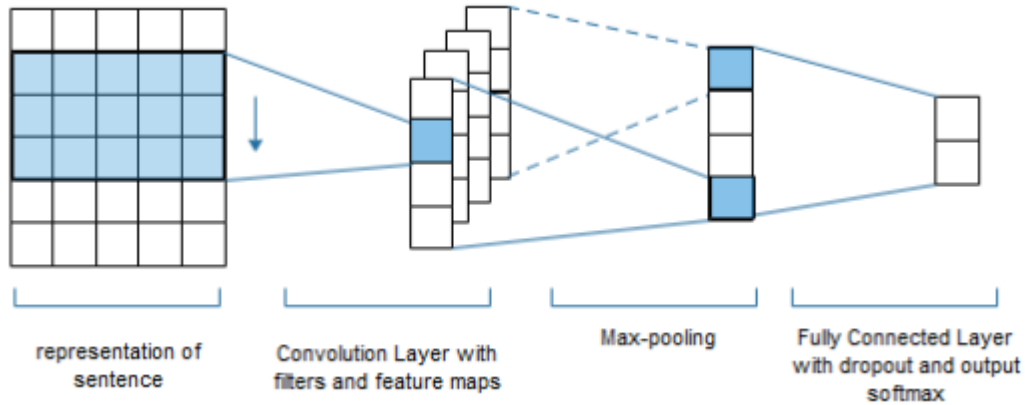
2.4 Word Embedding

Pembentukan representasi vektor dari data berbentuk teks, dikenal sebagai *Word Embedding*, menjadi langkah krusial dalam mengkonversi informasi ke dalam bentuk numerik guna memfasilitasi pemrosesan oleh model-model pemodelan, yang umumnya hanya dapat memproses data bersifat numerik. Salah satu *library* yang saat ini diterapkan dalam tahapan *Word Embedding* adalah *FastText*, sebuah perpustakaan (*library*) yang dikembangkan oleh *Facebook*. *FastText* mengembangkan konsep *Word2Vec*, suatu metode yang telah terlebih dahulu diakui dalam domain *Word Embedding*. Perbedaan fitur dari *FastText* terletak pada penggunaannya terhadap n-gram, di mana setiap kata tidak hanya diproses sebagai satu entitas keseluruhan, melainkan menggunakan potongan kata berukuran n, contohnya trigram (n=3) seperti "pin", "int", "nta", "tar" untuk kata "pintar". Keunggulan utama *FastText* melibatkan kecepatan relatif proses yang dijalankannya. Metode ini mampu menangani kata-kata yang tidak pernah muncul dalam kamus.

2.5 Convolutional Neural Network Bi-Directional Long Short Term Memory (CNN-BiLSTM)

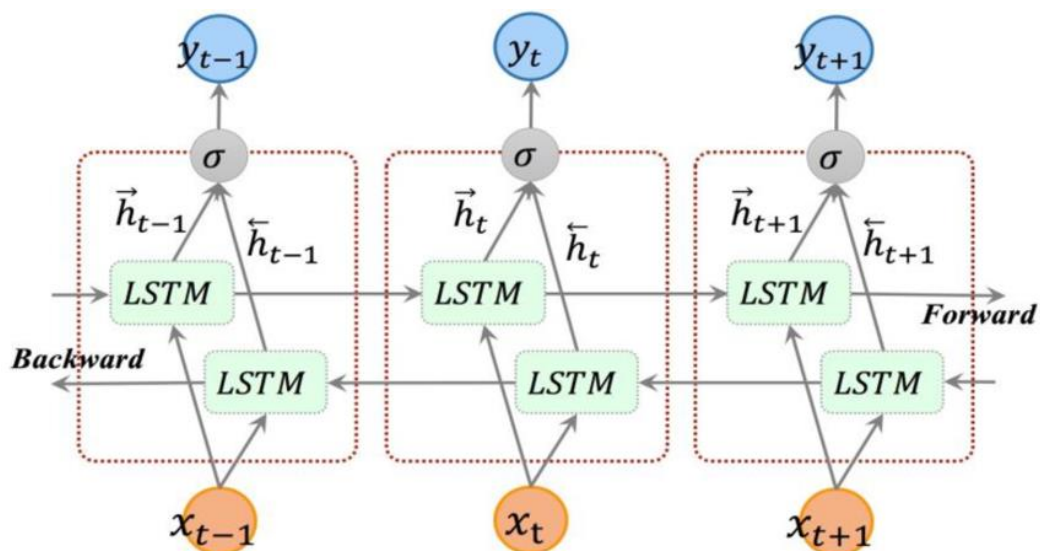
Convolutional Neural Network (CNN) merupakan bagian dari *deep neural network* yaitu berupa jaringan saraf tiruan berlapis dengan kemampuan untuk mendeteksi

fitur kompleks dalam data, seperti mengekstraksi fitur dalam data gambar dan teks. Arsitektur CNN yang digunakan memiliki tiga layer yaitu : convolutional layer, pooling layer dan fully connected layer.



Gambar 2. 1 Arsitektur Umum CNN (Ari Bangsa et al., 2020)

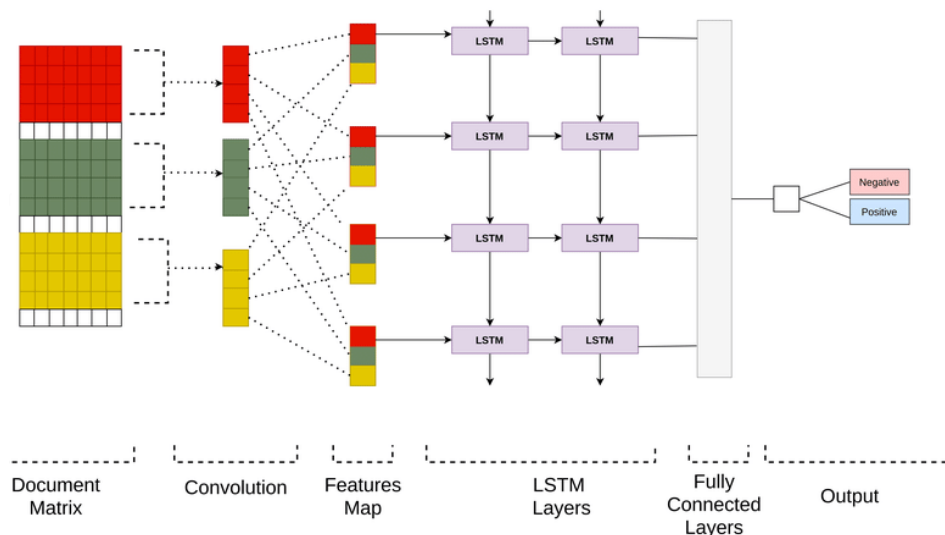
BiLSTM merupakan peningkatan model dari model LSTM yang dapat menangani informasi kontekstual dari token. Dasarnya BiLSTM mendapatkan informasi kontekstual dengan menggabungkan *hidden layer*, *forward* dan *backward*. Dengan menggunakan *forward* and *backward layer*, BiLSTM dapat menangkap informasi masa lalu dan masa depan. (Cahyaningtyas et al., 2021)



Gambar 2. 2 Arsitektur umum BiLSTM (Puteri, Dian., 2023).

CNN-BiLSTM merupakan jenis jaringan saraf tiruan yang menggabungkan model CNN dan BiLSTM. Model yang diterapkan menggabungkan lapisan CNN untuk

ekstraksi fitur dan BiLSTM untuk output dari lapisan max-polling yang digunakan untuk mempertimbangkan ketergantungan jangka panjang. (Cahyaningtyas et al., 2021)



Gambar 2. 3 Arsitektur umum CNN-BiLSTM (Shahriar, Fahim & Bashar, Md. 2021).

2.6 Penelitian Terdahulu

Penelitian terkait aspect based sentiment analysis telah dilakukan oleh cahyaningtyas et al. pada tahun 2021 dengan judul Deep Learning for Aspect Based Sentiment Analysis on Indonesian hotels reviews. Pada penelitian ini proses klasifikasi dilakukan dengan membandingkan delapan metode deep learning yaitu RNN, LSTM, GRU, BiLSTM, Attention BiLSTM, CNN, CNN-LSTM dan CNNBiLSTM dalam klasifikasi aspek harga, hotel, kamar, lokasi, pelayanan dan restoran. Dalam klasifikasi sentimen membandingkan dua scenario yaitu sentimen positif dan sentimen negative. Hasil penelitian menunjukkan bahwa LSTM mencapai model terbaik untuk klasifikasi aspek dengan nilai akurasi sebesar 0,926. Untuk klasifikasi sentimen hasil penelitian menunjukkan bahwa model CNN mendapatkan nilai akurasi rata-rata sebesar 0,904.

Penelitian yang dilakukan oleh Siti Masturoh pada tahun 2020 dengan judul Sentiment Analysis Against the Dana E-Wallet on Google Play Reviews Using the K-Nearest Neighbor Algorithm, pada penelitian ini analisis sentimen e-wallet menggunakan algoritma KNN dikelompokkan menjadi tiga kelas berdasarkan

jumlah bintang pada menu review dan didapatkan nilai akurasi tertinggi dengan pengujian dua kelas pada bintang satu dan bintang lima, dan didapatkan nilai k1 dengan hasil akurasi terbaik sebesar 86,64%. Penelitian selanjutnya dilakukan oleh Surya et al. pada tahun 2021 dengan judul penelitian Analisis Sentimen Terhadap Ulasan Film Menggunakan Word2Vec dan Support Vector Machine. Pada penelitian ini menggunakan data sebanyak 10.000 ulasan dan model klasifikasi SVM linier menghasilkan nilai terbaik sebesar 78,74% f1-score. Lalu penelitian yang dilakukan oleh Mubarak et al pada tahun 2017 dengan judul Aspect-based sentiment analysis to review products using Naïve Bayes, berdasarkan pengujian yang dilakukan Naïve Bayes berkinerja dengan baik untuk aspek berdasarkan analisis sentimen dengan f1-measure sebesar 78,12%. F1-measure untuk klasifikasi aspek adalah 88,13% dan f1-measure untuk klasifikasi sentimen adalah 75%.

Selanjutnya penelitian yang dilakukan oleh Cahyani et al. pada tahun 2019 dengan judul penelitian Analisis Sentimen terhadap Ulasan Hotel menggunakan Boosting Weighted Extreme Learning Machine. Proses klasifikasi menghasilkan sentimen berupa kelas positif atau kelas negative. Berdasarkan pengujian yang dilakukan didapatkan model dengan nilai f-measure tertinggi sebesar 0.953 dengan nilai optimal parameter yang didapatkan nilai $C = 16$, $L = 64$, dan weak learner = 256.

Penelitian yang dilakukan oleh Kristiyanti et al pada tahun 2020 dengan judul penelitian E-Wallet Sentiment Analysis Using Naïve Bayes and Support Vector Machine Algorithm. Penelitian ini melakukan perbandingan antara algoritma Naïve Bayes dan Support Vector Machine dalam klasifikasi ulasan pengguna ewallet OVO dan DANA dengan teks bahasa inggris. OVO mendapatkan akurasi yang cukup tinggi dari kedua algoritma yang dibandingkan. Naïve bayes mencapai 94,90% sedangkan SVM mencapai 91,00% dengan perbedaan presisi yang dekat yaitu 3,90%. Namun jika dibandingkan dengan AUC yang diperoleh, maka nilai AUC untuk SVM mencapai 0,986 (mencapai sempurna) sehingga peneliti menyimpulkan bahwa algoritma SVM mampu bekerja dengan baik dalam menganalisis sentimen.

Tabel 2. 1 Penelitian Terdahulu

No	Peneliti	Metode	Judul	Keterangan
1.	(Siti Masturoh, 2020)	K-Nearest Neighbor	Sentiment Analysis Against the Dana E-Wallet on Google Play Reviews Using the K-Nearest Neighbor Algorithm	Dengan menggunakan KNN pada review e-wallet DANA pada google play di dapatkan akurasi terbaik sebesar 86,67%
2.	(Naquitasia et al., 2022)	CNN dan CNN-BiLSTM	Analisis Sentimen Berbasis Aspek Pada Wisata Halal dengan Metode Deep Learning	Untuk klasifikasi aspek dengan CNN didapat akurasi 98.299% sedangkan untuk klasifikasi sentimen didapat akurasi sebesar 93.96%.
3.	(Mubarok et al., 2017)	Naïve Bayes	Aspect-based sentiment analysis to review products using Naïve Bayes	nilai analisis sentimen dengan f1-measure sebesar 78,12%. F1-measure untuk klasifikasi aspek adalah 88,13% dan f1-measure untuk klasifikasi sentimen adalah 75%.
4.	(Cahyani et al., 2019)	Boosting Weighted ELM	Analisis Sentimen terhadap Ulasan Hotel menggunakan <i>Boosting Weighted Extreme Learning Machine</i>	Berdasarkan pengujian yang dilakukan, didapatkan model dengan nilai f-measure tertinggi sebesar 0,953.

Tabel 2.1 Penelitian Terdahulu (Lanjutan)

5.	(Kristiyanti et al., 2020)	SVM dan Naïve Bayes	E-Wallet Sentiment Analysis Using Naïve Bayes and Support Vector Machine Algorithm	Naïve Bayes dan SVM unggul untuk mengklasifikasikan ulasan pengguna e-wallet OVO dan DANA. Ketepatan Naïve Bayes mencapai 94,90%, sedangkan SVM cukup tinggi yang menghasilkan akurasi 91,00%
6.	(Cahyaningt yas et al., 2021)	RNN, LSTM, GRU, BiLSTM, Attention BiLSTM, CNN, CNN-LSTM, dan CNN-BiLSTM)	Deep learning for aspect-based sentiment analysis on Indonesian hotels reviews	LSTM untuk klasifikasi aspek dengan nilai akurasi sebesar 0,926. Untuk klasifikasi sentimen hasil penelitian menunjukkan bahwa model CNN mendapatkan nilai akurasi rata-rata sebesar 0,904.

2.7 Perbedaan Penelitian

Penelitian yang dilakukan oleh peneliti memiliki perbedaan signifikan dengan penelitian sebelumnya. Sebelumnya, penelitian hanya terbatas pada analisis sentimen terhadap e-wallet Dana, sedangkan penelitian saat ini mengadopsi pendekatan yang lebih canggih dengan menggunakan metode *Convolutional Neural Network – Bidirectional Long Short Term Memory* (CNN-BiLSTM) untuk melakukan analisis sentimen berbasis aspek terhadap ulasan e-wallet DANA pada platform *Google Playstore* dalam bahasa Indonesia.

BAB 3

ANALISIS DAN PERANCANGAN SISTEM

3.1 Dataset

Penelitian ini menggunakan data yang diperoleh dari ulasan pengguna yang mengunggah aplikasi Dana pada *Google Playstore* menggunakan teknik *scraping* dalam format *csv*. Jumlah data yang terkumpul sebanyak 5000 data. Berikut ini beberapa data awal dari hasil *scraping*.

Tabel 3. 1 Hasil *scraping* ulasan pengguna Dana

No	Content
1.	pembayaran lewat QRIS masih belum masuk juga sudah lebih dari 30 jam. Transaksi masih terpending belum masuk juga ke saldo bisnis saya. Ini aplikasi bagus cuman orang2 customer service nya kurang cepat alias lamban menangani setiap masalah daftarnya mudah dan verifikasi premiumnya juga ga sulit.
2.	Sebenarnya aplikasi ini cukup bagus bagi saya dalam hal transaksi, performanya cepat dan tidak memakan waktu lama, pelayanan nya mantap sekali
3.	Turunin bintang, pelayanan gajelas, akun saya tiba tiba dibekukan, banding pembekuan akun udah hampir setahun gak selesai-selesai, padahal data asli dan lengkap udah dilampirkan berkali-kali, tetep aja diminta minta lagi datanya, tapi mklum sih orang ga bisa bedain data asli sama photo copy.
4.	Setelah upgrade dr sebelumnya, 1. Sering minta verifikasi untuk upgrade dana ke premium, padahal (SUDAH PREMIUM) 2. Logout sendiri dan susah login dgn notif jaringan tidak ada! Data/WiFi lancar!! 3. Sering error javascript ngebug parah. Krn ini aplikasi keren,memudahkan transaksi, tapi dengan hadirnya masalah tsb SAYA pun KECEWA!!.
5.	Slogan yg indah, sok"an nerima pengaduan 24x7 jam CS slalu aktif 😊 nyatanya Cs sangat amat SLOW RESPON 🐢 setiap laporan yg di ajukan akan dapat blasan jika sudah 8 hari / lebih (berlaku email) kalo WA mlah perbaikan katanya, 1 bulan msak perbaikan terus 😊. Uang yg ggal pengiriman tdk akan dpt kembali lagi. Padahal udh puluhan juta gw udah pernah transaksi di apk ini, ketika gw tau Cs nya kya bot 😊 1-10 = 2 Oh iya DANA Care nya ganti aja DANA Lelet. registrasinya ribet banget

Dataset yang telah terkumpul sebanyak 5000 data di bagi menjadi dua bagian yaitu data *training* (data latih) sebanyak 80% dan data *testing* (data uji) sebanyak 20% yang selanjutnya data latih akan di bagi kembali menjadi dua bagian yaitu data latih utama sebanyak 80% dan data validasi sebanyak 20%. Pembagian dataset ditampilkan seperti pada table 3.2

Tabel 3. 2 Pembagian Dataset

Training	Validation	Testing
3200	800	1000

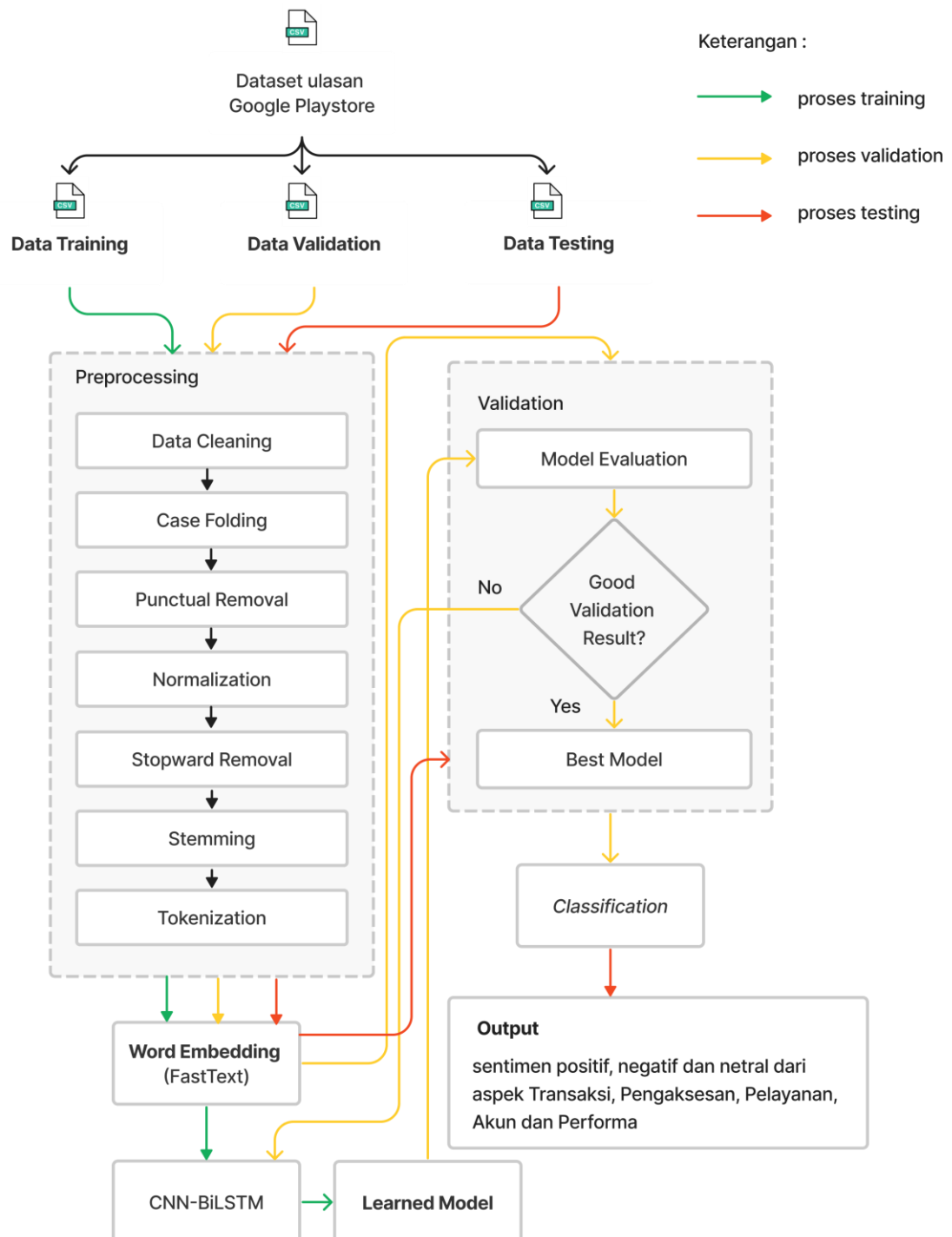
3.2 Arsitektur Umum

Adapun langkah dalam pengerjaan penelitian ini diawali dengan melakukan pengumpulan data berformat *.csv* yang diterapkan pada proses *training* dan *testing*. Selanjutnya dilakukan proses *Preprocessing* yaitu *cleaning data*, *case folding*, *punctual removal*, *normalization*, *stopward removal*, *stemming*, dan *tokenization*. Kemudian data yang sudah melalui proses *Preprocessing* akan melalui pembobotan kata dengan menggunakan *fastText*. Kemudian data akan dilatih dan menghasilkan model yang akan digunakan pada proses uji. Setiap langkah akan dijelaskan secara rinci pada arsitektur umum yang ditampilkan pada Gambar 3.1. Adapun kerja sistem terbagi menjadi beberapa bagian, yaitu:

3.2.1 Input

Pada penelitian ini data input yang digunakan adalah data ulasan berbahasa Indonesia oleh pengguna DANA. Data yang diperoleh di bagi menjadi dua bagian yaitu data *training* (data latih) sebanyak 80% dan data *testing* (data uji) sebanyak 20% yang selanjutnya data latih akan di bagi kembali menjadi dua bagian yaitu data latih utama sebanyak 80% dan data validasi sebanyak 20%. Data *training* dan data *testing* didapat melalui ulasan pengguna melalui situs *Google Play Store* menggunakan teknik *scraping* dalam format *.csv*. Dataset yang digunakan diberi label secara manual pada setiap ulasan dari pengguna. Dalam penentuan topik klasifikasi digunakan metode *Word Cloud* dimana metode ini digunakan untuk melakukan *highlight* pada kata dari frasa populer. Sehingga didapat 5 topik klasifikasi yang akan diidentifikasi yaitu

access, transaction, user account, customer service, dan performance. (Fattahila, n.d.) Setiap aspek memiliki keyword yang dijabarkan pada table 3.3 berikut :



Gambar 3. 1 Arsitektur Umum

Tabel 3. 3 Kata kunci dari setiap aspek

Topik klasifikasi	Kata
Transaction	Transfer, Pembelian, Pembayaran, Saldo, Topup, Beli, Listrik, Transaksi, Biaya, Potongan, Uang
Customer Service	Email, WA, Whatsapp, Telpon, CS, Telfon, Respon, Pelayanan, Laporan, Keluhan, Complain, Call Center
User Account	Upgrade, Verifikasi, KTP, Premium, Akun
Access	Login, Registrasi, OTP, Kode, Pin, Daftar, Akses
Performance	Lambat, Lag, Cepat, Performa, Lelet, Lemot, Bug

Data berupa kumpulan ulasan yang sudah dipisah berdasarkan kategori secara manual ke 3 sentimen yakni positif, negatif, dan netral. Setiap ulasan yang dinyatakan positif (1) adalah ketika menyatakan sentimen berdasarkan pernyataan bahagia dan persetujuan. Review yang dinyatakan negatif (-1) ketika menyatakan pernyataan kecewa, amarah dan penolakan. Dan untuk review yang dinyatakan netral (0) adalah ketika aspek yang bersangkutan tidak terdapat didalam kalimat review. Data tersebut akan menjadi acuan untuk sistem dapat dilatih. Contoh data dapat dilihat pada table 3.4.

Tabel 3. 4 contoh data ulasan pengguna dan label

Ulasan Pengguna	Topik klasifikasi				
	transaksi	pelayanan	akun	akses	performa
Dana nya jelek,, udah premium tiba2 jadi gak premium,, mau upgrade lagi gak bisa gagal terus,, kecewaaaaaaa...	0	0	-1	0	0
Kecewa dengan pelayanan dana sih sebenarnya udah beli pulsanya belum masuk tapi saldo nya udah kurang saat di minta bantuan juga tidak di respon di WA ig juga tda di respon buruk skrg	-1	-1	0	0	0

Tabel 3. 4 contoh data ulasan pengguna dan label (Lanjutan)

Kenapa aplikasi DANA sy g bisa login yah,,, udah di restart hp, udah di update aplikasi tetep g bisa di buka	0	0	0	-1	0
Dana sekarang sangat lambat .. sering banget jaringan sibuk .. buat TF buru2 jadi ngga bisa ... Tolong di perbaiki lagi ya bosku	0	0	0	0	-1
Terima kasih atas responnya. Ternyata kendala karena sudah lama belum terupdate. Sekarang aplikasi sudah bisa digunakan dengan cepat lancar tanpa ada kendala	0	1	0	0	1

3.2.2 Pre-processing

Preprocessing merupakan serangkaian tahapan menyeleksi data yang akan dipakai agar lebih terstruktur sehingga mempermudah penelitian. Adapun tahapan preprocessing adalah sebagai berikut:

3.2.2.1 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.

Function `cleaning(data)`:

```
# Step 1: Replace "\\r" with a space in the
'text' column

data['Tweet_Parsed_1'] =
data['text'].replace("\\\\r", " ")
```

```

# Step 2: Remove non-ASCII characters
(emoticons)

data['Tweet_Parsed_1'] =
data['Tweet_Parsed_1'].astype(str).apply(lambda
x: x.encode('ascii', 'ignore').decode('ascii'))

# Step 3: Remove usernames starting with '@'
data['Tweet_Parsed_1'] =
data['Tweet_Parsed_1'].replace('@[\w]+', '')

# Step 4: Replace "\\n" and "\n" with a space
data['Tweet_Parsed_1'] =
data['Tweet_Parsed_1'].replace("\\\\n", " ")

data['Tweet_Parsed_1'] =
data['Tweet_Parsed_1'].replace("\n", " ")

# Step 5: Replace "\r" with a space
data['Tweet_Parsed_1'] =
data['Tweet_Parsed_1'].replace("\r", " ")

# Step 6: Replace multiple spaces with a
single space
data['Tweet_Parsed_1'] =
data['Tweet_Parsed_1'].replace("    ", " ")

# Step 7: Remove links (URLs)
pattern =
r'''(?i)\b(?:https|http?://|www\d{0,3}[.][a-z0-
9.\-]+\.[a-
z]{2,4})/)(?:[^\s()<>+|\(\([^\s()<>+|\(\([^\s()<>+
\)\)]*\)\))+(\?:\(\([^\s()<>+|\(\([^\s()<>+|\)\)]*\)\)|[
^\s`!()\[\]{};:'".,<>?«»"'"')])'''

data['Tweet_Parsed_1'] =
data['Tweet_Parsed_1'].replace(pattern, " ")

# Step 8: Remove double quotes
data['Tweet_Parsed_1'] =
data['Tweet_Parsed_1'].replace('""', '')

```



```
# Step 9: Remove '\\x' followed by two
alphanumeric characters
data['Tweet_Parsed_1'] =
data['Tweet_Parsed_1'].replace("\\\\x[a-zA-z0-
9][a-zA-z0-9]", "")
# Return the cleaned data
return data
```

Pseudocode Data Cleaning

3.2.2.2 Case Folding

Case folding adalah proses untuk mengkonversi seluruh karakter menjadi huruf kecil (*lowercase*). Hal ini bertujuan untuk mengurangi sensitifitas terhadap kalimat.

Contoh :

Sebelum proses casefolding

Pembayaran kartu kredit sejak tanggal 28 Oktober 2022 sampai sekarang status nya masih order di Proses..pembayaran saya jadi lewat dari tanggal jatuh temponya.. terpaksa bayar lagi.. sudah konfirm via email, WA, dm IG tidak ada hasil..uang saya entah dimana rimbanya :)
buat yg mau atau sudah download DANA silahkan pikir pikir lagi resikonya.. you deserve better :')

Setelah dilakukan proses casefolding

pembayaran kartu kredit sejak tanggal 28 oktober 2022 sampai sekarang status nya masih order di proses..pembayaran saya jadi lewat dari tanggal jatuh temponya.. terpaksa bayar lagi.. sudah konfirm via email, wa, dm ig tidak ada hasil..uang saya entah dimana rimbanya :)
buat yg mau atau sudah download dana silahkan pikir pikir lagi resikonya.. you deserve better :')

Function lowerCase(data):

```
# Step 1: Convert the 'Tweet_Parsed_1' column
to lowercase
```

```

data['Tweet_Parsed_2'] =
data['Tweet_Parsed_1'].toLowerCase()
# Return the dataframe with the lowercase
text in 'Tweet_Parsed_2'
return data

```

Pseudocode Case Folding

3.2.2.3 Punctual Removal

Punctual removal dilakukan untuk menghilangkan semua tanda baca dan angka yang terdapat pada data ulasan sehingga hanya kalimat saja yang dapat di proses.

Contoh :

Sebelum proses punctual removal

pembayaran kartu kredit sejak tanggal 28 oktober 2022 sampai sekarang status nya masih order di proses..pembayaran saya jadi lewat dari tanggal jatuh temponya.. terpaksa bayar lagi.. sudah konfirm via email, wa, dm ig tidak ada hasil..uang saya entah dimana rimbanya :)
 buat yg mau atau sudah download dana silahkan pikir pikir lagi resikonya.. you deserve better :')

Setelah dilakukan proses punctual removal

pembayaran kartu kredit sejak tanggal oktober sampai sekarang status nya masih order di proses pembayaran saya jadi lewat dari tanggal jatuh temponya terpaksa bayar lagi sudah konfirm via email wa dm ig tidak ada hasil uang saya entah dimana rimbanya buat yg mau atau sudah download dana silahkan pikir pikir lagi resikonya you deserve better

Function removePunct(data):

```

# Step 1: Create a new column
'Tweet_Parsed_3' and copy the content from
'Tweet_Parsed_2'
data['Tweet_Parsed_3'] =
data['Tweet_Parsed_2']

```

```

# Step 2: Remove non-alphabetic characters
(punctuation) and replace with a space
data['Tweet_Parsed_3'] =
data['Tweet_Parsed_3'].replace('[^a-zA-Z]+', ' ')
# Return the dataframe with the processed
text in 'Tweet_Parsed_3'
return data

```

Pseudocode Punctual Removal

3.2.2.4 Normalization

Pada tahap ini, normalization bertujuan untuk menormalisasi setiap kata pada dataset. Proses yang akan dilakukan adalah untuk menormalkan kalimat berisi bentuk kata tidak baku berupa singkatan dan typo menjadi kata bentuk baku dan memiliki makna. Peneliti menggunakan kamus yang sudah disediakan.

Contoh :

Sebelum proses punctual removal

pembayaran kartu kredit sejak tanggal oktober sampai sekarang status nya masih order di proses pembayaran saya jadi lewat dari tanggal jatuh temponya terpaksa bayar lagi sudah konfirm via email wa **dm ig** tidak ada hasil uang saya entah dimana rimbanya buat yg mau atau sudah **download** dana silahkan pikir pikir lagi **resikonya** you deserve better

Setelah dilakukan proses punctual removal

pembayaran kartu kredit sejak tanggal oktober sampai sekarang status masih order di proses pembayaran saya jadi lewat dari tanggal jatuh temponya terpaksa bayar lagi sudah konfirm via email wa **direct message instagram** tidak ada hasil duit saya entah dimana rimbanya buat mau atau sudah **unduh** dana silakan pikir pikir lagi **risikonya** you deserve better

```

Function normalize_text(text, stdword_,
nonstdword_):
    # Step 1: Split the input text into a list of
words
    text = text.split(" ")

    # Step 2: Iterate through each word in the
text
    for i in range(len(text)):
        # Step 3: Check if the word is in the
list of non-standard words
        if text[i] in nonstdword_:
            # Step 4: If found, get the index and
replace the word with the corresponding standard
word
            index = nonstdword_.index(text[i])
            text[i] = stdword_[index]

    # Step 5: Join the modified list of words
into a single string
    return ' '.join(map(str, text))

Function normalizeData(data, stdword_,
nonstdword_):
    # Step 1: Create a new column
'Tweet_Parsed_4' and copy the content from
'Tweet_Parsed_3'
    data['Tweet_Parsed_4'] =
data['Tweet_Parsed_3']

    # Step 2: Apply the normalize_text function
to each element in 'Tweet_Parsed_4'
    data['Tweet_Parsed_4'] =
data['Tweet_Parsed_4'].map(lambda com:
normalize_text(com, stdword_, nonstdword_))

```

```
# Return the dataframe with the normalized
text in 'Tweet_Parsed_4'
return data
```

Pseudocode Normalization

3.2.2.5 Stopword Removal

Tahapan selanjutnya yaitu stopwords removal atau penghapusan kata penghubung. Stopwords adalah kata-kata umum yang sering muncul dalam suatu bahasa dan cenderung tidak memberikan kontribusi signifikan terhadap makna kalimat atau dokumen. Meskipun tidak ada standar stopwords yang bersifat universal, seringkali daftar stopwords mencakup kata-kata umum seperti "dan," "atau," "di," dan sebagainya, tergantung pada bahasa yang digunakan dalam analisis.

Contoh :

Sebelum proses stopwords removal

pembayaran kartu kredit sejak tanggal oktober sampai sekarang status masih order di proses pembayaran saya jadi lewat dari tanggal jatuh temponya terpaksa bayar lagi sudah konfirm via email wa direct message instagram tidak ada hasil duit saya entah dimana rimbanya buat mau atau sudah unduh dana silakan pikir pikir lagi risikonya you deserve better

Setelah dilakukan proses stopwords removal

pembayaran kartu kredit tanggal oktober status order proses pembayaran tanggal jatuh temponya terpaksa bayar konfirm via email wa direct message instagram tidak hasil duit dimana rimbanya unduh dana silakan pikir pikir risikonya you deserve better

```
Function removeMeaningless(data, meaningless_):
    # Step 1: Iterate through each stopwords in
the list of meaningless words
    for word in meaningless_:
        # Step 2: Create a regex pattern for the
current stopwords
```

```

        regex_meaningless = r"\b" + word + r"\b"
        # Step 3: Replace occurrences of the
stopword with an empty string in 'Tweet_Parsed_4'
        data['Tweet_Parsed_4'] =
data['Tweet_Parsed_4'].replace(regex_meaningless,
'')
        # Return the dataframe with the meaningless
words removed from 'Tweet_Parsed_4'
        return data

```

Pseudocode Stopword Removal

3.2.2.6 Stemming

Stemming ialah proses mengurangi kata-kata ke bentuk dasarnya dengan menghapus imbuhan. Tujuannya untuk menghasilkan bentuk dasar.

Contoh :

Sebelum proses stemming

pembayaran kartu kredit tanggal oktober status order proses
pembayaran tanggal jatuh temponya terpaksa bayar konfirmasi via
email wa direct message instagram tidak hasil duit dimana rimbanya
unduh dana silakan pikir pikir risikonya you deserve better

Setelah dilakukan proses stemming

bayar kartu kredit tanggal oktober status order proses bayar tanggal
jatuh tempo paksa bayar konfirmasi via email wa direct message
instagram tidak hasil duit mana rimba unduh dana sila pikir pikir
risiko you deserve better

```

Function stemming(text, stemmer_id):
    # Step 1: Split the input text into a list of
words
    text_split = text.split(" ")
    # Step 2: Initialize an empty list to store
the stemmed words

```

```

    stemmed_list = []
    # Step 3: Iterate through each word in the
text
    for i in text_split:
        # Step 4: Stem each word using the
provided stemmer
        stem_text = stemmer_id.stem(i)
        # Step 5: Append the stemmed word to the
list
        stemmed_list.append(stem_text)

    # Step 6: Join the list of stemmed words into
a single string
    stemmed = ' '.join(map(str, stemmed_list))
    # Step 7: Return the stemmed text
    return stemmed
Function stemData(data, stemmer_id):
    # Step 1: Create a new column
'Tweet_Parsed_6' and copy the content from
'Tweet_Parsed_5'
    data['Tweet_Parsed_6'] =
data['Tweet_Parsed_5']
    # Step 2: Apply the stemming function to each
element in 'Tweet_Parsed_6'
    data['Tweet_Parsed_6'] =
data['Tweet_Parsed_6'].map(lambda com:
stemming(com, stemmer_id))
    # Return the dataframe with the stemmed text
in 'Tweet_Parsed_6'
    return data

```

Pseudocode Stemming

3.2.2.7 Tokenization

Tokenization merupakan tahapan membagi kalimat menjadi bagian yang lebih kecil berupa token token.

Contoh :

Sebelum proses tokenization

bayar kartu kredit tanggal oktober status order proses bayar tanggal
jatuh tempo paksa bayar konfirmasi via email wa direct message
instagram tidak hasil duit mana rimba unduh dana sila pikir pikir
risiko you deserve better

Setelah proses tokenization

['bayar', 'kartu', 'kredit', 'tanggal', 'oktober', 'status', 'order', 'proses',
'bayar', 'tanggal', 'jatuh', 'tempo', 'paksa', 'bayar', 'konfirmasi', 'via', 'email',
'wa', 'direct', 'message', 'instagram', 'tidak', 'hasil', 'duit', 'mana', 'rimba',
'unduh', 'dana', 'sila', 'pikir', 'pikir', 'risiko', 'you', 'deserve', 'better']

```
# Step 1: Initialize a Tokenizer with a maximum
number of words (MAX_NB_WORDS)
tokenizer = Tokenizer(num_words=MAX_NB_WORDS,
lower=True, char_level=False)
# Step 2: Fit the Tokenizer on the combined text
data (train_list + test_list + val_list)
tokenizer.fit_on_texts(train_list + test_list +
val_list)
# Step 3: Convert text sequences in the training
set to numerical sequences
train_seq =
tokenizer.texts_to_sequences(train_list)
# Step 4: Convert text sequences in the testing
set to numerical sequences
test_seq =
tokenizer.texts_to_sequences(test_list)
```



```

# Step 5: Convert text sequences in the
validation set to numerical sequences
val_seq = tokenizer.texts_to_sequences(val_list)
# Step 6: Get the word index from the Tokenizer
word_index = tokenizer.word_index
# Step 7: Pad sequences to ensure consistent
length
train_seq_pad = np.array(pad_sequences(train_seq,
maxlen=max_len))
test_seq_pad = np.array(pad_sequences(test_seq,
maxlen=max_len))
val_seq_pad = np.array(pad_sequences(val_seq,
maxlen=max_len))

# Step 8: Save the Tokenizer for later use
pickle.dump(tokenizer, open("tokenizer.pkl",
'wb'))

```

Pseudocode Tokenization

3.2.3 Word Embedding

Pada tahapan ini data yang sudah di bentuk berupa token akan di ubah ke dalam bentuk vektor. Hal ini dilakukan agar komputer mampu mengolah data yang akan digunakan. Ini diperlukan karena pemodelan biasanya hanya mampu memproses data bersifat numerik. *Library* yang digunakan untuk penelitian ini menggunakan *word embedding FastText*. Pada *FastText*, setiap kata direpresentasikan sebagai rata-rata representasi vektor dari karakternya n-gram bersamaan dengan kata itu sendiri.

Setelah terbentuknya korpus, tiap kata akan diubah menjadi representasi vektor dalam ruang berdimensi 300. *Embedding matrix* adalah matriks yang digunakan sebagai representasi vektor untuk setiap kata dalam model NLP. Matrix ini berisi vektor numerik yang mencerminkan makna dan hubungan antar kata dalam suatu konteks. Pada pemodelan bahasa, *embedding matrix*

seringkali digunakan sebagai bobot yang dapat diatur selama proses pelatihan model. Representasi dari *embedding matrix* ditunjukkan pada Gambar 3.2.

(-0.0075	0.0691	0.134	0.062	0.0064
	-0.079	0.1127	-0.0429	-0.1018	-0.1246
	0.1761	-0.0113	0.03	0.0722	0.0078
	0.0207	-0.031	-0.0561	0.0889	0.1045
	-0.1132	-0.0536	0.0102	0.1152	0.01
	-0.0234	0.0462	-0.0238	0.1056	0.0497
	0.1825	0.1084	-0.0444	-0.0551	0.0235
	-0.15549999	-0.0574	-0.0693	0.0654	-0.0615
	0.1064	0.0204	-0.0348	-0.0828	-0.0206
	0.0407	0.0102	0.0496	-0.041	0.0281
	0.0452	-0.0058	-0.0955	-0.1743	0.0664
	0.0982	0.17919999	0.0905	0.0806	-0.0401
	0.0603	0.1468	0.0497	-0.0158	-0.0217
	0.0796	-0.0518	0.0776	0.0589	0.0115
	0.0413	-0.0574	-0.0518	-0.0082	-0.0542
	-0.0396	-0.0397	-0.0903	-0.0883	0.0386
	0.0266	-0.037	-0.0264	0.0635	-0.0521
	-0.0164	-0.0624	-0.2789	-0.0336	-0.017
	0.0529	0.0282	-0.066	0.0635	-0.0978
	-0.1646	0.0332	0.0945	0.1365	-0.0795
	-0.0576	0.0207	-0.088	0.0428	0.1548
	0.0194	0.1016	-0.0986	-0.0364	-0.1162
	0.0281	-0.0735	-0.031	0.0069	-0.0088
	0.0452	-0.0547	-0.0015	0.056	-0.0079
	-0.0254	-0.0603	-0.0019	0.0225	-0.0662
	0.0515	0.0234	0.033	0.0306	0.0295
	0.0064	0.0461	0.0818	0.0098	0.1507
	0.0749	0.0134	0.0152	0.075	-0.0557
	0.0536	-0.22840001	0.0457	0.0038	-0.09
	-0.0744	0.1111	-0.0043	0.0004	0.0627
	0.0816	-0.0368	0.0103	-0.13070001	0.1937
	0.0329	0.0482	0.0462	-0.0432	0.0666
	0.20100001	-0.082	0.0383	0.0339	0.0469
	0.0026	0.0272	-0.0357	0.0099	-0.0136
	-0.0701	-0.0396	0.0226	-0.0557	0.0349
	0.0869	0.0711	-0.0485	-0.0058	-0.0291
	0.0879	-0.0923	0.0252	-0.0128	0.0621
	0.0278	0.094	0.0733	-0.0731	-0.0779
	0.0505	-0.0399	-0.029	0.0505	0.1409
	-0.0423	0.1058	-0.0928	0.0243	0.0021
	-0.0651	-0.0685	-0.1053	0.1023	0.1001
	0.0392	-0.0857	0.0668	-0.1133	0.0409
	0.0266	0.0295	-0.0675	0.0172	0.0174
	-0.1398	-0.0617	-0.0077	-0.0108	0.0429
	0.0219	0.0739	0.0591	0.0407	-0.1109
	0.1635	-0.0313	-0.0027	-0.0601	-0.079
	0.0377	0.0386	-0.0402	0.1426	-0.03
	0.0095	-0.0274	-0.0804	0.061	-0.0606
	0.0284	-0.0325	0.082	-0.1145	-0.1577
	-0.061	-0.1516	0.0343	0.0498	-0.0734
	0.0639	-0.1222	0.0436	-0.1479	0.0162
	-0.0628	0.0964	0.1031	0.0145	-0.0922
	0.0149	-0.0547	0.0195	0.0076	0.0228
	-0.0014	-0.1195	0.0019	-0.0582	0.0555
	-0.0499	0.0135	0.0778	-0.0559	-0.1202
	-0.0193	-0.2236	0.156	-0.0808	-0.0273
	0.0126	-0.0554	0.0008	-0.0051	0.1055
	-0.0032	-0.1285	0.0414	-0.0313	-0.0087
	0.0315	0.1078	-0.1212	0.0248	-0.0116
	0.0623	-0.197	-0.1372	0.0124	0.15970001)

Gambar 3. 2 Embedding Matrix

3.2.4 Model Building

Selanjutnya adalah tahapan pembentukan model menggunakan arsitektur CNN-BiLSTM (*Convolutional Neural Network - Bidirectional Long Short Term Memory*). Model yang dibentuk akan dilengkapi dengan beberapa lapisan yang memiliki tujuan untuk menerima *input* agar dapat menghindari terjadinya *overfitting* dan *underfitting*. Berikut merupakan penjelasan umum tentang beberapa lapisan yang terdapat dalam model tersebut :

1. *Embedding Layer*

Pada lapisan *embedding layer*, terjadi pembentukan matriks input sekuens setelah tahap tokenisasi. Matriks ini disusun dari kata-kata dalam sekuens yang diberikan bobot awal secara acak, dan selanjutnya proses pembelajaran dilakukan agar model dapat mengintegrasikan informasi dari setiap kata ke dalam dataset pelatihan.

2. *Convolutional Layer*

Pada *convolutional layer* akan dilakukan ekstraksi fitur dari representasi embedding teks. Jumlah filter yang digunakan dalam layer ini 64 filter. Kemudian fungsi aktivasi *Rectified Linear Unit* (ReLU) digunakan untuk menambahkan unsur non-linearitas. ReLU memastikan bahwa hasil konvolusi memiliki karakteristik non-linear. *Convolutional layer* ini berfungsi untuk mengekstrak pola atau fitur-fitur lokal dari representasi embedding teks. Setelah dilakukan konvolusi, *pooling layer* bertanggung jawab untuk mereduksi dimensi dari hasil konvolusi. Menggunakan MaxPooling, yang akan mengambil nilai maksimum dari setiap nilai berturut-turut.

3. *Bidirectional LSTM layer*

Setelah melalui tahapan *convolutional layer* selanjutnya *output* yang dihasilkan akan menjadi inputan BiLSTM layer, pada layer ini akan memproses sekuensial dari kedua arah, memungkinkan model memahami konteks dari kedua sisi. Inputan akan dilatih menggunakan model BiLSTM yang terdiri dari 32 unit (*neuron*). Jumlah unit ini dapat dianggap

sebagai dimensi ruang fitur yang dihasilkan oleh lapisan ini. Hasilnya adalah representasi sekuensial yang telah diperkaya dengan informasi kontekstual dan temporal dari kedua arah.

4. *Fully Connected Layer (Dense Layer)*

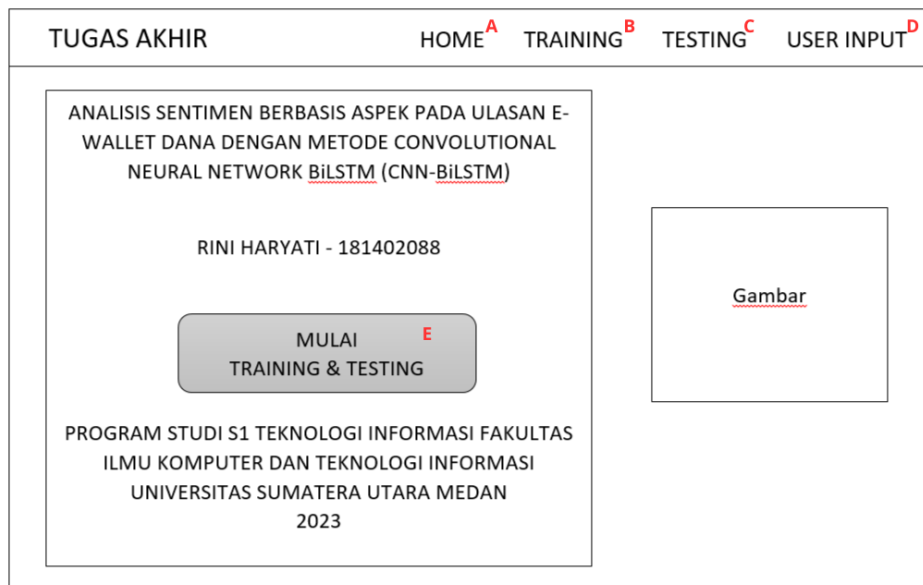
Dense Layer kemudian menerima input dari setiap neuron, layer ini terhubung dengan semua neuron dari layer sebelumnya, membentuk representasi fitur yang lebih abstrak. Setiap lapisan Dense ini adalah output layer untuk kategori atau kelas tertentu. seperti, "transaksi", "pelayanan", "account", "access", dan "performa". Fungsi aktivasi softmax digunakan di output layer untuk menghasilkan distribusi probabilitas untuk setiap kelas. Ini memastikan bahwa probabilitas semua kelas pada setiap output layer jumlahnya 1.

3.3 Perancangan Antarmuka Sistem

Model yang dibangun akan diaplikasikan pada aplikasi berbasis website. User dapat memasukkan file dataset dengan format csv. Perancangan aplikasi ini bertujuan untuk memudahkan pengguna dalam mengakses sistem yang nantinya dibangun. Rancangan ini terdiri dari beberapa tahap yaitu:

3.3.1 Rancangan Halaman Beranda

Halaman beranda menampilkan informasi umum mengenai sistem. Terdapat tiga menu navigasi pada bagian header yaitu menu beranda, menu *training* (pelatihan) dan menu *testing* (pengujian). Pada halaman beranda terdapat satu tombol *shortcut* yang akan memudahkan pengguna untuk menuju halaman training. Halaman beranda diperlihatkan seperti gambar 3.3



Gambar 3. 3 Gambar Rancangan Halaman Beranda

Keterangan:

1. Tombol (A) berguna untuk menampilkan halaman utama dari sistem.
2. Tombol (B) berguna untuk mengarahkan ke halaman training.
3. Tombol (C) berguna untuk mengarahkan ke halaman testing.
4. Tombol (D) berguna untuk mengarahkan ke halaman analisis pengguna (user input)
5. Tombol (E) berguna untuk mengarahkan ke halaman training.

3.3.2 Rancangan Tampilan Halaman Training Data

Halaman training berfungsi sebagai tempat pelatihan model menggunakan dataset yang telah diberi label. Rancangan tampilan halaman training ditunjukkan seperti gambar 3.4

TUGAS AKHIR HOME^A TRAINING^B TESTING^C USER INPUT^D

Training
 pilih file data training berekstensi .csv

Choose file...^E Mulai Training^F

Hasil Data Preprocessing^G

Gambar 3. 4 Rancangan Tampilan Halaman Training Data

Keterangan:

1. Tombol (A) berguna untuk menampilkan halaman utama dari sistem.
2. Tombol (B) berguna untuk mengarahkan ke halaman training.
3. Tombol (C) berguna untuk mengarahkan ke halaman testing.
4. Tombol (D) berguna untuk mengarahkan ke halaman analisis pengguna (user input)
5. Tombol (E) berguna untuk input file data training.
6. Tombol (F) berguna untuk memulai proses training.
7. Tombol (G) berguna untuk menampilkan hasil proses dari data training yang di input.

3.3.3 Rancangan Tampilan Halaman Testing

Halaman testing merupakan halaman yang berfungsi sebagai tempat dilakukan proses pengujian (testing). Setelah suatu model selesai dilatih, diperlukan adanya evaluasi yang berfungsi selain untuk mengidentifikasi suatu model, juga untuk menilai performa sistem saat melakukan analisis sentimen pada ulasan aplikasi Dana yang diujikan berdasarkan aspeknya. Rancangan tampilan halaman testing ditunjukkan pada Gambar 3.5

TUGAS AKHIR **HOME**^A **TRAINING**^B **TESTING**^C **USER INPUT**^D

Testing
pilih file data training berekstensi .csv

Choose file...^E Mulai Testing^F

Hasil Analisis Sentimen Berdasarkan Aspek^G

Gambar 3. 5 Rancangan Tampilan Halaman Testing

Keterangan :

1. Tombol (A) berguna untuk menampilkan halaman utama dari sistem.
2. Tombol (B) berguna untuk mengarahkan ke halaman training.
3. Tombol (C) berguna untuk mengarahkan ke halaman testing.
4. Tombol (D) berguna untuk mengarahkan ke halaman analisis pengguna (user input)
5. Tombol (E) berguna untuk input file data testing.
6. Tombol (F) berguna untuk memulai proses testing.
7. Tombol (G) berguna untuk menampilkan hasil proses data testing yang di input

3.3.4 Rancangan Tampilan Halaman User Input

Halaman user input adalah halaman yang berfungsi sebagai tempat pengguna melakukan analisis terhadap data tunggal. Pada halaman ini pengguna akan memberikan inputan berupa ulasan berbahasa Indonesia terhadap aplikasi Dana. Output yang dihasilkan adalah prediksi sentimen dan aspek berdasarkan ulasan yang diinput oleh pengguna. Rancangan tampilan halaman user input ditunjukkan seperti Gambar 3.6

Gambar 3. 6 Rancangan Tampilan Halaman User Input

Keterangan :

1. Tombol (A) berguna untuk menampilkan halaman utama dari sistem.
2. Tombol (B) berguna untuk mengarahkan ke halaman training.
3. Tombol (C) berguna untuk mengarahkan ke halaman testing.
4. Tombol (D) berguna untuk mengarahkan ke halaman analisis pengguna (user input)
5. Tombol (E) berguna untuk input file data tunggal untuk dilakukan prediksi
6. Tombol (F) berguna untuk memulai proses prediksi
7. Tombol (G) berguna untuk menampilkan hasil proses prediksi yang berupa sentimen dan aspek

3.4 Metode Evaluasi

Perlunya metode evaluasi guna menjadi penilaian kinerja terhadap suatu pemodelan analisis sentimen berbasis aspek pada ulasan pengguna aplikasi Dana pada algoritma yang digunakan. *Confussion Matrix* adalah metode yang digunakan dalam penelitian ini. Evaluasi yang diterapkan adalah mengikuti aturan variable dengan perhitungan seperti *accuracy*, *preciission*, *recall*, dan *F1-Score*.

Tabel 3. 5 Confussion Matrix

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

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

- *True Positive* (TP): Prediksi positif dan hasilnya positif
- *True Negative* (TN): Prediksi negatif dan hasilnya positif
- *False Positive* (FP): Prediksi positif dan hasilnya negatif
- *False Negative* (FN): Prediksi negatif dan hasilnya negative

Confussion matrix juga dapat menghitung berbagai performa matrix, diantaranya:

a. *Accuracy*

Accuracy merupakan persentase tingkat kebenaran dari nilai prediksi dibanding dengan nilai sebenarnya dari sebuah data. Semakin tinggi nilai *accuracy*, maka semakin akurat model yang akan mengklasifikasi data dengan benar. Adapun untuk mendapatkan nilai *accuracy* dapat dilihat pada persamaan 3.1 dibawah ini

$$\text{Akurasi} = \frac{\text{jumlah data diprediksi benar}}{\text{jumlah keseluruhan data}} \quad (3.1)$$

b. *Precision*

Precision dalam konteks ini, adalah ukuran yang menunjukkan rasio antara jumlah prediksi benar positif dengan total hasil yang diprediksi sebagai positif. Adapun 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 persentase keberhasilan model dalam memperoleh informasi yang diminta. Oleh karena itu, *recall* adalah rasio antara jumlah prediksi benar positif dengan total data yang sebenarnya positif. Adapun rumus untuk mendapatkan nilai *recall* dapat dilihat pada persamaan 3.3 dibawah ini

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

d. *F1-Score*

F1-Score sebagai fungsi dari *precision* dan *recall* yang mencerminkan persentase keseimbangan yang diperoleh antara *precision* dan *recall*. *F1-Score* disebut sebagai nilai perbandingan rata-rata dari *precision* dan *recall*. Adapun rumus mendapatkan nilai *F1-Score* dapat dilihat pada persamaan 3.4 dibawah ini

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

BAB 4

IMPLEMENTASI DAN PENGUJIAN

4.1 Implementasi Sistem

Pada tahap ini sistem yang sudah di rancang pada Bab 3 akan dibentuk dengan menggunakan beberapa perangkat keras (*hardware*) dan perangkat lunak (*software*). Adapun perangkat yang digunakan yaitu:

4.1.1 Spesifikasi Perangkat Keras dan Perangkat Lunak

Dalam penelitian ini, spesifikasi perangkat keras yang digunakan antara lain:

1. *Processor* AMD Ryzen 5 3500U @ 2.10 GHz
2. *Memory* (RAM) 8GB

Perangkat lunak yang digunakan dalam perancangan penelitian, yaitu:

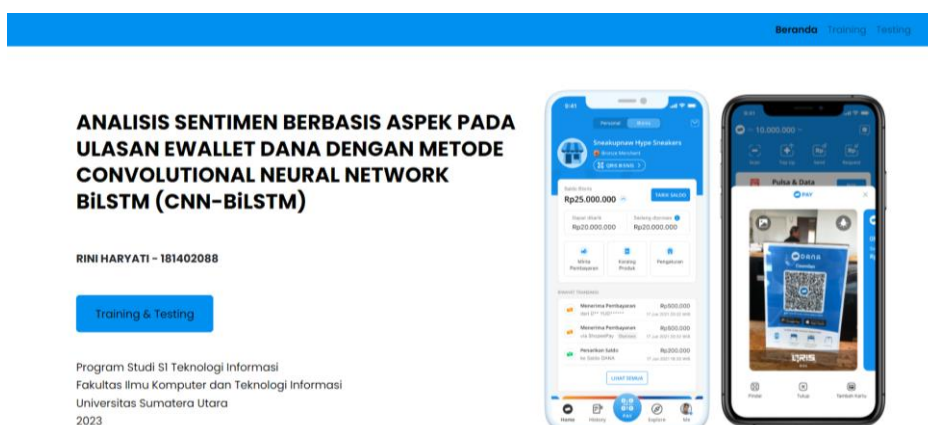
1. Sistem Operasi : Windows 11 Home 64 Bit
2. *Python* versi 3.8
3. *Jupyter Notebook*
4. *Library* yang digunakan pada python :
 - a. Tensorflow dan Keras
 - b. Flask
 - c. Matplotlib
 - d. NLTK
 - e. Pandas
 - f. Numpy
5. *Microsoft Visual Studio Code*

4.1.2 Implementasi Perancangan Tampilan Antarmuka

Adapun Implementasi tampilan antarmuka yang dibangun berdasarkan tampilan yang telah diranacang pada Bab 3 yaitu:

4.1.2.1. Tampilan Halaman Utama

Tampilan halaman utama menampilkan informasi umum berupa judul penelitian dan identitas penulis di sebelah kiri. Terdapat tiga menu navigasi pada bagian header yaitu menu beranda, menu training dan juga menu testing. Pada halaman beranda terdapat satu tombol *shortcut* yang akan memudahkan pengguna untuk menuju halaman *training* dan *testing*. Tampilan Halaman Utama ditampilkan seperti pada Gambar 4.1.



Gambar 4. 1 Tampilan Halaman Utama

4.1.2.2. Tampilan Halaman Training

Pada halaman training berfungsi sebagai tempat pelatihan model menggunakan dataset yang telah diberi label. Pada tampilan ini disediakan tempat untuk mengunggah data latih dengan format *.csv*. Disebelah kiri juga terdapat tempat untuk mengatur epochs dari data yang dilatih. Tampilan halaman training seperti pada Gambar 4.2.

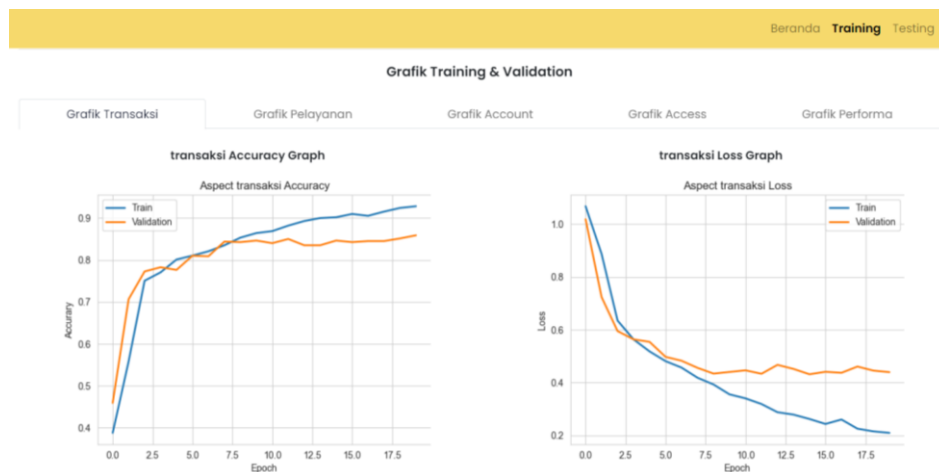
Gambar 4. 2. Tampilan Halaman Training

Setelah proses *training data* selesai, halaman akan menampilkan informasi berupa hasil dari *preprocessing* seluruh data *training*. Hasil *preprocessing* data *training* ditampilkan pada Gambar 4.3.

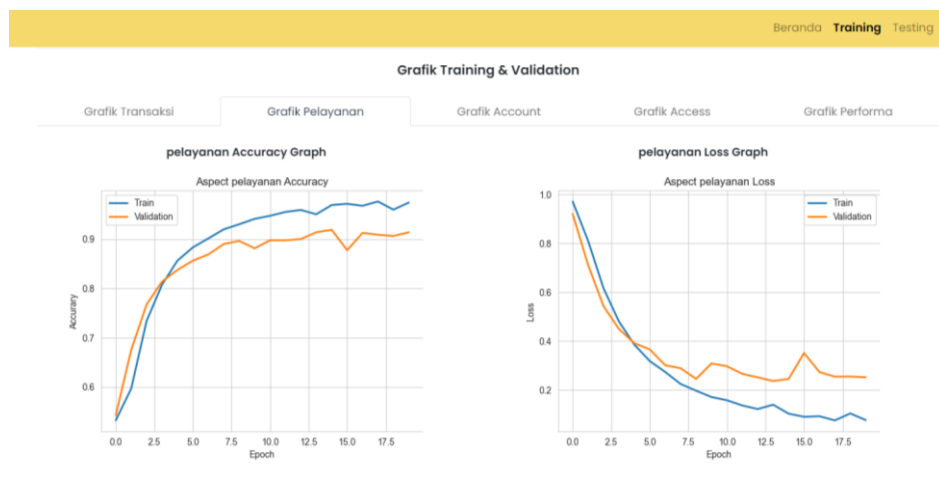
Ulasan	Preprocessing	Transaksi	Pelayanan	Account	Access	Performa
Tapi, sdh berapa kali di coba tidak bisa atau upgrade premium gagal, mohon petunjuk... aplikasinya juga lemot sering banget lelet	kali coba tidak tingkat premium gagal mohon tunjuk aplikasi lot banget lelet	0	0	-1	0	-1
"Aplikasi Telah Berhenti" Bug macam apa ini... Sangat sulit untuk transfer uang semenjak di update, login aplikasi nya aja susah force close trus. Jaringan bagus. Penyimpanan data banyak. Tolong diperbaiki sebelum Bintang 1 dari publik semakin banyak	aplikasi henti bug sulit transfer duit semenjak baru login aplikasi susah force close jaring bagus simpan data tolong baik bintang publik	-1	0	0	-1	-1

Gambar 4. 3. Tampilan Halaman Training (preprocessing)

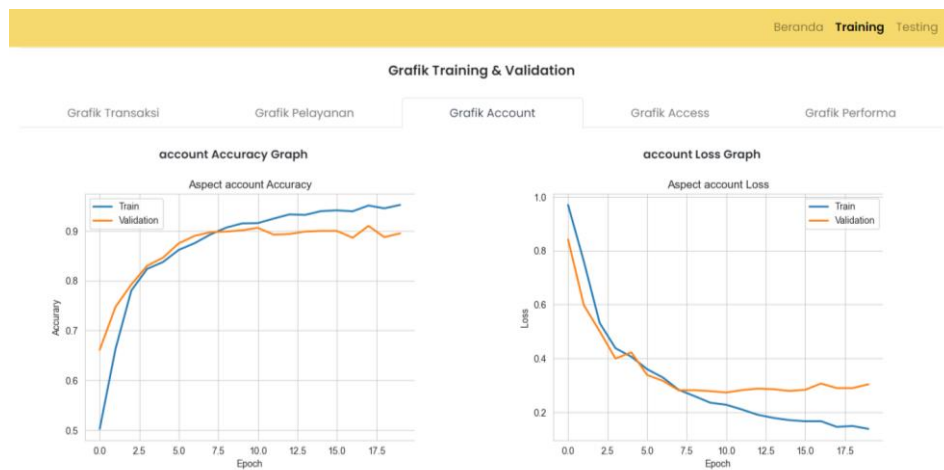
Pada halaman training juga ditampilkan nilai *accuration* dan *loss* yang didapatkan dari tiap aspek. Ditampilkan seperti pada gambar 4.4 sampai gambar 4.8.



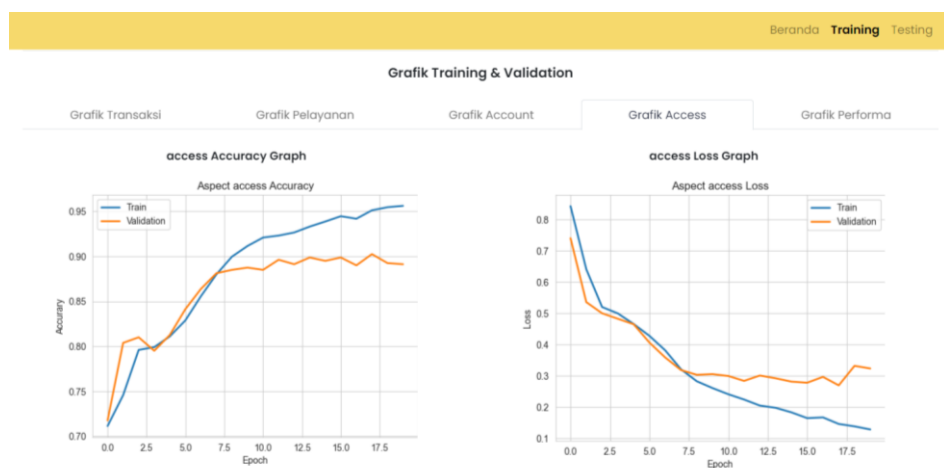
Gambar 4. 4. Tampilan Hasil Training (*accuration* dan *loss* aspek transaksi)



Gambar 4. 5. Tampilan Hasil Training (*accuration* dan *loss* aspek pelayanan)



Gambar 4. 6. Tampilan Hasil Training (*accuration* dan *loss* aspek *account*)



Gambar 4. 7. Tampilan Hasil Training (*accuration* dan *loss* aspek *access*)



Gambar 4. 8. Tampilan Hasil Training (*accuration* dan *loss* aspek performa)

4.1.2.3. Tampilan Halaman Testing

Pada halaman testing disediakan tempat untuk mengunggah data uji dengan format *.csv*. tampilan halaman testing ditampilkan seperti pada gambar 4.9.

The figure shows the 'Testing' interface. At the top, there's a navigation bar with 'Beranda', 'Training', and 'Testing'. The 'Testing' tab is selected. The main heading is 'TESTING'. Below it, the text 'Upload Dataset Untuk Proses Testing' is displayed. There is a file upload area with a 'Choose File' button and the text 'No file chosen'. To the right of the file upload area is a blue button labeled 'Mulai Testing'. Below the upload area, there is a section labeled 'Hasil Prediksi'.

Gambar 4. 9 Tampilan Halaman Testing

Setelah proses pengujian data selesai, selanjutnya akan ditampilkan hasil prediksi dari pengujian yang telah dilakukan. Hasil pengujian ditampilkan dalam format table seperti gambar 4.10.

Beranda

Training

Testing

TESTING

Upload Dataset Untuk Proses Testing

Choose File

data_testing.csv

Reset

Hasil Prediksi

Show

10

entries

Search:

Ulasan	Preprocessing	Prediction				
		Transaksi	Pelayanan	Account	Access	Performa
padahal mudah digunakan untuk topup, upgrade sulit pula banyak kendala.🤔🤔	mudah top up tingkat sulit kendala	0	0	0	0	0
- Pada tanggal 25 Oktober 2022, saya melakukan top up melalui Virtual Account BANK BCA sebesar Rp.192.500 dengan status pembayaran di M-banking BCA BERHASIL. Namun, sampai saat ini saldo pada aplikasi DANA belum	tanggal oktober top up virtual account bank baca rupiah status bayar banking baca hasil saldo aplikasi dana tambah alias tidak masuk tanggal	-1	0	0	0	0

Gambar 4. 10. Tampilan Halaman Testing (hasil prediksi)

4.1.2.4. Tampilan Halaman User Input

Pada halaman user input terdapat *text box* yang memungkinkan pengguna untuk melakukan prediksi terhadap kalimat ulasan pengguna aplikasi Dana. Tampilan user input di tampilan seperti gambar 4.11.

Beranda Training Testing	
Text Input Testing	
Masukkan Kalimat	Mulai Testing
Hasil Prediksi Real Time	

Gambar 4. 11. Tampilan Halaman User Input

Setelah pengguna menuliskan kalimat tunggal akan dihasilkan output dari inputan tersebut yang menghasilkan prediksi berupa sentimen dan aspek dari ulasan pengguna aplikasi Dana. Hasil prediksi pengguna halaman user input ditunjukkan pada gambar 4.12 dan gambar 4.13.

Beranda Training **Testing**

Text Input Testing

kecewa dengan pelayanan dana, udah beli pulsa malah belum masuk tapi saldonya sudah berkurang saat dimintai bantuan juga tidak direspon di wa lg juga tida di respon, buruk sekarang

Reset

Hasil Prediksi Real Time

Input	Transaksi	Pelayanan	Account	Access	Performa
kecewa dengan pelayanan dana, udah beli pulsa malah belum masuk tapi saldonya sudah berkurang saat dimintai bantuan juga tidak direspon di wa lg juga tida di respon, buruk sekarang	Negatif	Negatif	Netral	Netral	Netral

Gambar 4. 12. Tampilan Halaman User Input (sentimen negatif dan netral)

Beranda Training **Testing**

Text Input Testing

aplikasi sangat bagus untuk membantu sekali dalam bertransaksi keuangan.. mulai dari transfer uang ke semua Bank serta transaksi pembayaran online yg lancar.. pelayanannya sangat mantab performanya juga juara banget anti lelet.. Thanks untuk DANA

Reset

Hasil Prediksi Real Time

Input	Transaksi	Pelayanan	Account	Access	Performa
Aplikasi Sangat Bagus dan membantu sekali dalam bertransaksi keuangan.. mulai dari transfer uang ke semua Bank serta transaksi pembayaran online yg lancar... pelayanannya sangat mantab performanya juga juara banget anti lelet.. Thanks untuk DANA	Positif	Positif	Netral	Netral	Positif

Gambar 4. 13. Tampilan Halaman User Input (sentiment positif dan netral)

4.2 Hasil Pengujian Sistem

Tahap selanjutnya adalah melakukan pengujian sistem untuk menevaluasi kinerja model yang telah di buat. Pengujian sistem melibatkan proses pelatihan dan pengujian model untuk memastikan sistem yang telah dibangun berfungsi dengan baik.

4.2.1 Pelatihan Model

Dalam proses pelatihan model, pelatihan data *training* merupakan proses untuk melatih model sehingga menemukan model terbaik sebelum digunakan pada sistem. Berikut adalah contoh data *training*.

Tabel 4. 1 Data Training

Ulasan Pengguna	Topik klasifikasi				
	transaksi	pelayanan	akun	akses	performa
pembelian dengan dana sangat mudah, pelayanannya ramah dan sangat membantu, performa aplikasi sangat baik, transaksi juga lebih aman karna ada akses pin	1	1	0	1	1
udah kapok pake dana. Uang sering kepotong aja. Suka isi pulsa trs dananya kepotong tp pulsanya ga masuk. Laporan ke cs ga ada solusinya. Ga lagi deh pake ini apk verifikasi premium pun gak menjamin bakal aman nyimpan uang disini	-1	-1	-1	0	0
Upgrade nya gagal terus Sudah premium tiba-tiba turun lagi Saldo tidak bisa di apa-apa kan Jadi malas pakai aplikasi ini... Jujur padahal sangat berguna... Tapi kalau begini terus uninstal juga nampak nya	0	0	-1	0	0
Tolong kak ya saya sudah berkali kali coba login ke aplikasi dana nya tapi malah di suruh hubungi cs pokok nya banyak banget tolong benarin aplikasi jika ingin maju trus susah banget login ke aplikasi nya padahal sandi dan no nya udh benar semua tapi gak bisa login ke dana nya	0	0	0	-1	0

Tabel 4. 1 Data Training (Lanjutan)

Minta tolong min danaku suka nge lag sistem error' tolong di perbaiki ya min. Udh bagus bintang nya,, tolong dijaga ya bosskuh.	0	0	0	0	-1
---	---	---	---	---	----

Dataset dibagi menjadi dua dengan ratio 80% untuk pelatihan model klasifikasi dan 20% akan dipilih secara acak untuk pengujian terhadap model. Kemudian data akan melalui proses preprocessing seperti *cleaning data*, *case folding*, *punctual removal*, *normalization*, *stopward removal*, *stemming*, dan *tokenization*. Hasil preprocessing ditunjukkan pada table 4.2

Tabel 4. 2 Hasil Preprocessing

Ulasan	Hasil tokenization	transak si	pelaya nan	Akun	akses	perfor ma
pembelian dengan dana sangat mudah, pelayanannya ramah dan sangat membantu, performa aplikasi sangat baik, transaksi juga lebih aman karna ada akses pin	['beli', 'dana', 'mudah', 'layan', 'ramah', 'bantu', 'performa', 'aplikasi', 'baik', 'transaksi', 'aman', 'akses', 'pin']	1	1	0	1	1
udah kapok pake dana. Uang sering kepotong aja. Suka isi pulsa trs dananya kepotong tp pulsanya ga masuk. Lapor ke cs ga ada solusinya. Ga lagi deh pake ini apk verifikasi premium pun gak menjamin bakal aman nyimpan uang disini	['kapok', 'pakai', 'dana', 'duit', 'potong', 'suka', 'isi', 'pulsa', 'dana', 'potong', 'pulsa', 'tidak', 'masuk', 'lapor', 'rekan', 'tidak', 'solusi', 'tidak', 'pakai', 'verifikasi', 'premium', 'tidak', 'jamin', 'aman', 'simpan', 'duit']	-1	-1	-1	0	0

Tabel 4.2 Hasil Preprocessing (Lanjutan)

Upgrade nya gagal terus Sudah premium tiba-tiba turun lagi Saldo tidak bisa di apa-apa kan Jadi malas pakai aplikasi ini... Jujur padahal sangat berguna... Tapi kalau begini terus uninstal juga nampak nya	['tingkat', 'gagal', 'premium', 'turun', 'saldo', 'tidak', 'malas', 'pakai', 'aplikasi', 'jujur', 'guna', 'uninstal', 'tampak']	0	0	-1	0	0
Tolong kak ya saya sudah berkali kali coba login ke aplikasi dana nya tapi malah di suruh hubungi cs pokok nya banyak banget tolong benarin aplikasi jika ingin maju trus susah bngat login ke aplikasi nya padahal sandi dan no nya udh benar semua tapi gak bisa login ke dana nya	['tolong', 'kakak', 'kali', 'kali', 'coba', 'login', 'aplikasi', 'dana', 'suruh', 'hubung', 'rekan', 'pokok', 'banget', 'tolong', 'benar', 'aplikasi', 'maju', 'susah', 'bngat', 'login', 'aplikasi', 'sandi', 'no', 'benar', 'tidak', 'login', 'dana']	0	0	0	-1	0
Minta tolong min danaku suka nge lag sistem error' tolong di perbaiki ya min. Udh bagus bintang nya,, tolong dijaga ya bosskuh.	['tolong', 'min', 'dana', 'suka', 'lag', 'sistem', 'error', 'tolong', 'baik', 'min', 'bagus', 'bintang', 'tolong', 'jaga', 'bosskuh']	0	0	0	0	-1

Setelah proses *tokenization*, data akan melewati tahap *word embedding*, pada penelitian ini tahapan *word embedding* menggunakan metode *fastText*, data yang sudah di pecah ke dalam bentuk token akan di ubah ke dalam bentuk vektor. Selanjutnya dilakukan pelatihan dengan algoritma CNN-BiLSTM. Algoritma CNN-BiLSTM berupa jaringan saraf tiruan berlapis dengan kemampuan untuk mendeteksi fitur kompleks dalam data. Model yang diterapkan menggabungkan lapisan CNN untuk ekstraksi fitur dan BiLSTM untuk outpun dari lapisan max-

polling yang digunakan untuk mempertimbangkan ketergantungan jangka panjang. Pada tahapan ini algoritma CNN-BiLSTM akan mempelajari pola dari data pelatihan untuk menentukan pengelompokan sesuai dengan label yang telah ditetapkan sebelumnya. Pada tahapan ini akan menghasilkan model pembelajaran yang berisi nilai dari rumus CNN-BiLSTM. Model yang sudah terbentuk akan memahami bahwa label -1 untuk ulasan negatif, label 0 untuk ulasan netral dan label 1 untuk ulasan positif.

Setelah model pembelajaran terbentuk, selanjutnya dilakukan pengujian model terhadap 20% data validasi yang dipilih acak dari data *training*. Hasil dari pengujian model diperoleh akurasi pengujian sebesar 89,5%.

4.2.1 Pengujian Model

Setelah model terbentuk dan dilakukan pelatihan maka dilakukan pengujian terhadap data uji yang telah di pisahkan sebelumnya. Hasil pengujian di tunjukkan pada tabel 4.3.

Tabel 4. 3 Hasil Pengujian

text	Preprocess	tra nsa ksi	pel aya na n	acc ou nt	acc ess	per for ma	tra nsa ksi _pr ed	pel aya na n_ pre d	acc ou nt_ pre d	acc ess _pr ed	per for ma _pr ed
Ngisi saldo via mandiri transaksi sudah berhasil sampai sekarang saldo masih nol. KECEWA BANGET. Ada pemeliharaan sistem bukannya di kasih tau di aplikasi malah diem bae kan asu	isi saldo via mandiri transaksi hasil saldo nol kecewa banget pelihara sistem kasih tau aplikasi diam anjing	-1	0	0	0	0	-1	0	0	0	0

Tabel 4. 3 Hasil Pengujian (Lanjutan)

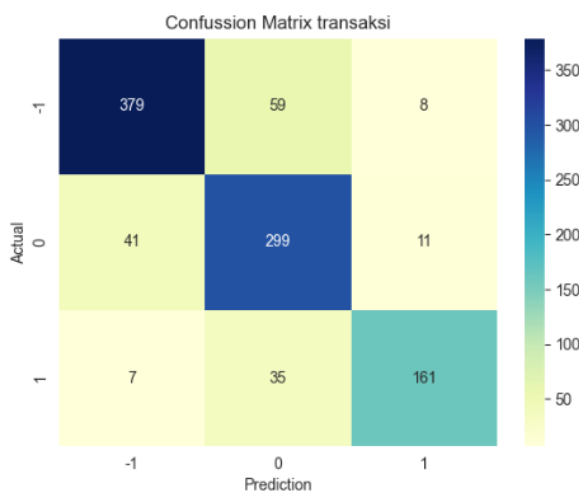
Ko diupdate malah jadi ga bisa tf ini gimana mohon diperbaiki bug nya,mau transfer keluar sistem eror,pengguna lama	diupdate tidak transfer mohon baik bug transfer sistem eror guna lama	0	0	0	0	-1	-1	0	0	0	-1
3 hari ini ada apa dengan DANA susah sekali dalam pengiriman uang. Tidak seperti biasanya Notes nya adalah "sistem sibuk silahkan coba beberapa saat lagi" mhn bantuannya karena sangat mengganggu transaksi kami	dana susah kirim duit tidak notes sistem sibuk sila coba mhn bantu mengganggu transaksi	-1	0	0	0	0	-1	0	0	0	0
Kenapa ya setelah update jadi aneh ini aplikasi bnyak gangguan segala macem transfer uang ga bisa terus aja "untuk sementara tidak bisa melakukan", tolong di perhatikan lah.	baru aneh aplikasi ganggu transfer duit tidak tidak tolong perhati	-1	0	0	0	0	-1	0	0	0	0

Tabel 4. 3 Hasil Pengujian (Lanjutan)

Ini kenapa saya mau upgrade gabisa bisa dibilang ktp sudah pernah digunakan tapi belum pernah digunakan..	0	0	-1	0	0	0	0	-1	0	0
Saya masih ada saldo mau tingkat tidak kirim ke teman bilang kartu aja gabisaa.. tanda duduk Tolong saldo kirim diperbaiki teman gabisaa punya saya. tolong baik										

4.3 Evaluasi

Evaluasi dilakukan untuk mengetahui tingkat performa yang diterapkan ketika proses klasifikasi berdasarkan aspek dengan analisis sentimen yang dapat dilihat berdasarkan perhitungan *confussion matrix*. *Confussion matrix* merupakan alat evaluasi yang umum digunakan dalam pembelajaran mesin untuk mengevaluasi kinerja suatu model klasifikasi. *Confussion matrix* memberikan gambaran lebih rinci mengenai perbandingan antara nilai actual dengan nilai prediksi melalui perhitungan *accuracy*, *preciission*, *recall*, dan *F1-Score* seperti yang ditampilkan pada gambar 4.14 *confussion matrix* dari aspek transaksi.

**Gambar 4. 14** Confussion Matrix Aspek Transaksi

Berdasarkan gambar 4.14 pada aspek transaksi didapatkan nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) dari setiap sentimen negatif, sentimen positif, dan sentimen netral. Nilai tersebut di tampilkan dalam tabel 4.4, table 4.5 dan table 4.6.

Tabel 4.4 Nilai Variabel Sentimen Negatif dari Aspek Transaksi

No	Sentimen Negatif	Total
1	TP (<i>True Positive</i>)	379
2	TN (<i>True Negative</i>)	506
3	FP (<i>False Positive</i>)	48
4	FN (<i>False Negative</i>)	67

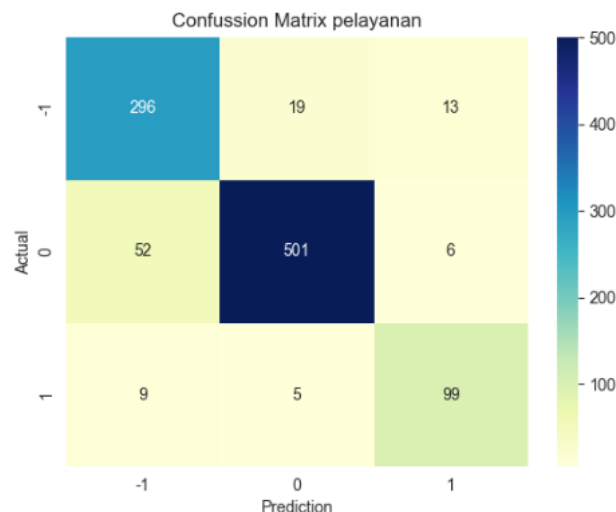
Tabel 4.5 Nilai Variabel Sentimen Positif dari Aspek Transaksi

No	Sentimen Positif	Total
1	TP (<i>True Positive</i>)	161
2	TN (<i>True Negative</i>)	778
3	FP (<i>False Positive</i>)	19
4	FN (<i>False Negative</i>)	42

Tabel 4.6 Nilai Variabel Sentimen Netral dari Aspek Transaksi

No	Sentimen Netral	Total
1	TP (<i>True Positive</i>)	299
2	TN (<i>True Negative</i>)	555
3	FP (<i>False Positive</i>)	94
4	FN (<i>False Negative</i>)	52

Selanjutnya confusion matrix berdasarkan aspek pelayanan. Ditampilkan pada gambar 4.15



Gambar 4. 15 Confussion Matrix Aspek Pelayanan

Berdasarkan gambar 4.15 pada aspek pelayanan didapatkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) dari setiap sentimen negatif, sentimen positif, dan sentimen netral. Nilai tersebut di tampilkan dalam tabel 4.7, table 4.8 dan table 4.9

Tabel 4. 7 Nilai Variabel Sentimen Negatif dari Aspek Pelayanan

No	Sentimen Negatif	Total
1	TP (<i>True Positive</i>)	296
2	TN (<i>True Negative</i>)	611
3	FP (<i>False Positive</i>)	61
4	FN (<i>False Negative</i>)	32

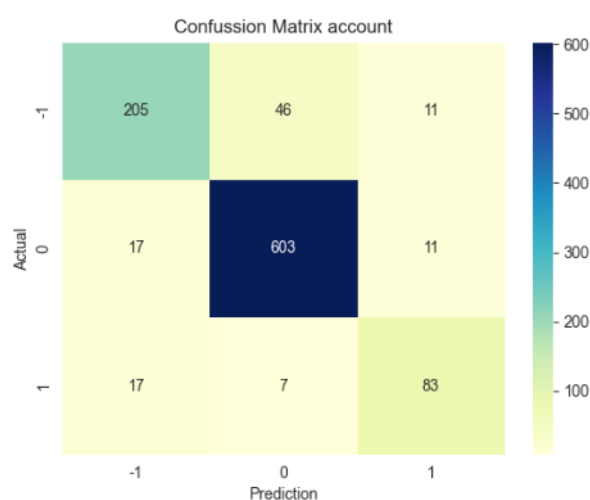
Tabel 4. 8 Nilai Variabel Sentimen Positif dari Aspek Pelayanan

No	Sentimen Positif	Total
1	TP (<i>True Positive</i>)	99
2	TN (<i>True Negative</i>)	868
3	FP (<i>False Positive</i>)	19
4	FN (<i>False Negative</i>)	14

Tabel 4. 9 Nilai Variabel Sentimen Netral dari Aspek Pelayanan

No	Sentimen Netral	Total
1	TP (<i>True Positive</i>)	501
2	TN (<i>True Negative</i>)	417
3	FP (<i>False Positive</i>)	24
4	FN (<i>False Negative</i>)	58

Selanjutnya confusion matrix berdasarkan aspek *account*. Ditampilkan pada gambar 4.16

**Gambar 4. 16** Confussion Matrix Aspek Account

Berdasarkan gambar 4.16 pada aspek Account didapatkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) dari setiap sentimen negatif, sentimen positif, dan sentimen netral. Nilai tersebut di tampilkan dalam tabel 4.10, table 4.11 dan table 4.12.

Tabel 4. 10 Nilai Variabel Sentimen Negatif dari Aspek Account

No	Sentimen Negatif	Total
1	TP (<i>True Positive</i>)	205
2	TN (<i>True Negative</i>)	704
3	FP (<i>False Positive</i>)	34
4	FN (<i>False Negative</i>)	57

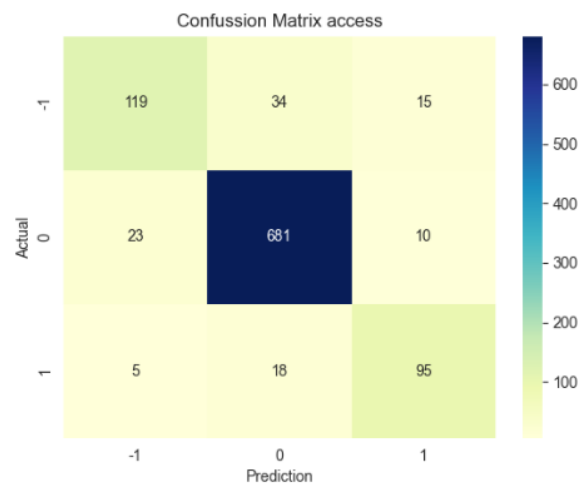
Tabel 4. 11 Nilai Variabel Sentimen Positif dari Aspek Account

No	Sentimen Positif	Total
1	TP (<i>True Positive</i>)	83
2	TN (<i>True Negative</i>)	871
3	FP (<i>False Positive</i>)	22
4	FN (<i>False Negative</i>)	24

Tabel 4. 12 Nilai Variabel Sentimen Netral dari Aspek Account

No	Sentimen Netral	Total
1	TP (<i>True Positive</i>)	603
2	TN (<i>True Negative</i>)	316
3	FP (<i>False Positive</i>)	53
4	FN (<i>False Negative</i>)	28

Selanjutnya confusion matrix berdasarkan aspek *access*. Ditampilkan pada gambar 4.17

**Gambar 4. 17** Confussion Matrix Aspek access

Berdasarkan gambar 4.17 pada aspek access didapatkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) dari setiap sentimen negatif, sentimen positif, dan sentimen netral. Nilai tersebut di tampilkan dalam tabel 4.13, table 4.14 dan table 4.15

Tabel 4.13 Nilai Variable Sentimen Negatif dari Aspek Access

No	Sentimen Negatif	Total
1	TP (<i>True Positive</i>)	119
2	TN (<i>True Negative</i>)	804
3	FP (<i>False Positive</i>)	28
4	FN (<i>False Negative</i>)	49

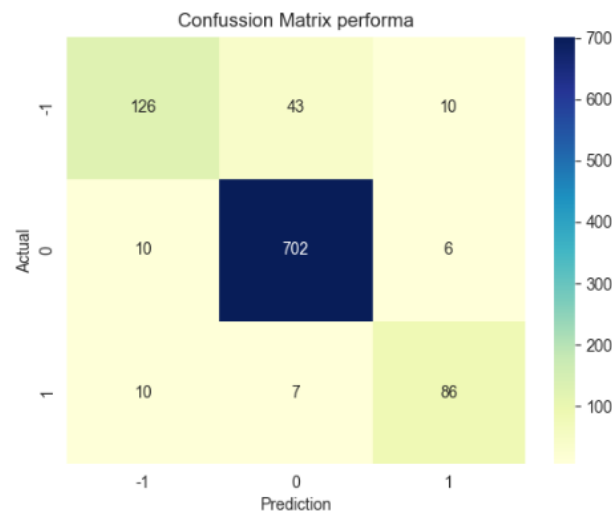
Tabel 4.14 Nilai Variabel Sentimen Positif dari Aspek Access

No	Sentimen Positif	Total
1	TP (<i>True Positive</i>)	95
2	TN (<i>True Negative</i>)	857
3	FP (<i>False Positive</i>)	25
4	FN (<i>False Negative</i>)	23

Tabel 4.15 Nilai Variabel Sentimen Netral dari Aspek Access

No	Sentimen Netral	Total
1	TP (<i>True Positive</i>)	681
2	TN (<i>True Negative</i>)	234
3	FP (<i>False Positive</i>)	52
4	FN (<i>False Negative</i>)	33

Selanjutnya confusion matrix berdasarkan aspek performa. Ditampilkan pada gambar 4.18



Gambar 4. 18 Confussion Matrix Aspek Performa.

Berdasarkan gambar 4.18 pada aspek performa didapatkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) dari setiap sentimen negatif, sentimen positif, dan sentimen netral. Nilai tersebut di tampilkan dalam tabel 4.16, table 4.17 dan table 4.18

Tabel 4. 16 Nilai Variabel Sentimen Negatif dari Aspek Performa

No	Sentimen Negatif	Total
1	TP (<i>True Positive</i>)	126
2	TN (<i>True Negative</i>)	801
3	FP (<i>False Positive</i>)	20
4	FN (<i>False Negative</i>)	53

Tabel 4. 17 Nilai Variabel Sentimen Positif dari Aspek Performa

No	Sentimen Positif	Total
1	TP (<i>True Positive</i>)	86
2	TN (<i>True Negative</i>)	881
3	FP (<i>False Positive</i>)	16
4	FN (<i>False Negative</i>)	17

Tabel 4. 18 Nilai Variabel Sentimen Netral dari Aspek Performa

No	Sentimen Netral	Total
1	TP (<i>True Positive</i>)	702
2	TN (<i>True Negative</i>)	232
3	FP (<i>False Positive</i>)	50
4	FN (<i>False Negative</i>)	16

Berdasarkan nilai variable pada tabel 4.4 sampai 4.18 akan dihitung tingkat akurasi model dengan menggunakan metric akurasi. Perhitungan akurasi dari setiap aspek ditunjukkan sebagai berikut:

$$\text{Akurasi Aspek Transaksi} = \frac{\text{jumlah data diprediksi benar}}{\text{jumlah keseluruhan data}} = \frac{839}{1000} = 0.839$$

$$\text{Akurasi Aspek Pelayanan} = \frac{\text{jumlah data diprediksi benar}}{\text{jumlah keseluruhan data}} = \frac{896}{1000} = 0.896$$

$$\text{Akurasi Aspek Account} = \frac{\text{jumlah data diprediksi benar}}{\text{jumlah keseluruhan data}} = \frac{891}{1000} = 0.891$$

$$\text{Akurasi Aspek Access} = \frac{\text{jumlah data diprediksi benar}}{\text{jumlah keseluruhan data}} = \frac{895}{1000} = 0.895$$

$$\text{Akurasi Aspek Performa} = \frac{\text{jumlah data diprediksi benar}}{\text{jumlah keseluruhan data}} = \frac{914}{1000} = 0.914$$

Dari perhitungan nilai akurasi tiap aspek, maka total nilai akurasi dari kelima aspek adalah:

$$\text{Akurasi} = \frac{0.839+0.896+0.891+0.895+0.914}{5} \times 100\% = 0.887 \times 100\% = 88.7\%$$

Berdasarkan perhitungan akurasi diatas, tingkat akurasi yang diperoleh dari pengujian sistem menggunakan CNN-BiLSTM adalah 88.7%. Dataset yang digunakan pada proses pengujian penelitian ini berjumlah 1000 ulasan dengan 5 aspek dan 3 jenis kategori sentimen. Jumlah data dan jumlah ulasan relevan yang kurang banyak menyebabkan polaritas yang dipelajari model masih terbatas. Kata yang sangat beragam dan kurang berpola juga mempengaruhi ketepatan model dalam melakukan proses pengujian

BAB 5

KESIMPULAN DAN SARAN

Pada Bab ini akan merangkum kesimpulan dari hasil implementasi metode CNN-BiLSTM dalam mengklasifikasi ulasan pengguna dompet digital Dana dan saran yang dapat penulis berikan untuk dijadikan pertimbangan bagi penelitian selanjutnya.

5.1 Kesimpulan

Kesimpulan yang didapat dari keseluruhan penelitian dimulai dari tahap preprocessing dan klasifikasi menggunakan algoritma CNN-BiLSTM adalah sebagai berikut:

1. Hasil pemodelan CNN-BiLSTM dalam mengklasifikasi ulasan pengguna E-Wallet Dana berdasarkan aspek Transaksi, Pelayanan, Account, Access dan performa dengan aspek positif, negative dan netral model CNN-BiLSTM memperoleh akurasi sebesar 88.7%.
2. Dataset yang digunakan sebanyak 5000 data dengan data training sebanyak 3200 data, data testing 1000 data, dan data validation sebanyak 800 data.
3. Berdasarkan akurasi yang telah di peroleh. Algoritma CNN-BiLSTM mampu bekerja dengan baik dalam mengklasifikasi ulasan pengguna berbahasa Indonesia.

5.2 Saran

Adapun saran untuk pengembangan penelitian lanjutan yang berkaitan dengan penelitian ini:

1. Keterbatasan data dan ulasan yang relevan dalam tiap aspek mempengaruhi performa model dalam proses klasifikasi, sehingga dibutuhkan jumlah data pada masing masing label yang seimbang sehingga didapatkan akurasi yang lebih baik.
2. Melakukan perbaikan pada kamus normalisasi untuk mendapatkan data yang lebih bersih.

DAFTAR PUSTAKA

- Anggraeni, L. (2019). *Go-Pay Jadi E-Walletl Paling Populer di Indonesia* - Medcom.id.
- Ari Bangsa, M. T., Priyanta, S., & Suyanto, Y. (2020). Aspect-Based Sentiment Analysis of Online Marketplace Reviews Using Convolutional Neural Network. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 14(2), 123.
- Cahyani, R., Indriati, & Pandu Adikara, P. (2019). Analisis Sentimen terhadap Ulasan Hotel menggunakan Boosting Weighted Extreme Learning Machine. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(8), 7767–7773.
- Cahyaningtyas, S., Hatta Fudholi, D., & Fathan Hidayatullah, A. (2021). Deep Learning for Aspect-Based Sentiment Analysis on Indonesian Hotels Reviews. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 4(3).
- Dr. S. Vijayarani, Ms. J. Ilamathi, M. N. (2015). Preprocessing Techniques for Text Mining Preprocessing Techniques for Text Mining. *International Journal of Computer Science & Communication Networks*, 5(October 2014), 7–16.
- Fattahila, A. A. (n.d.). *Analisis Sentimen Aplikasi Dompot Digital Indonesia Menggunakan Metode CNN Dan LSTM*. 1–9.
- Kristiyanti, D. A., Putri, D. A., Indrayuni, E., Nurhadi, A., & Umam, A. H. (2020). E-Wallet Sentiment Analysis Using Naïve Bayes and Support Vector Machine Algorithm. *Journal of Physics: Conference Series*, 1641(1).
- Liu, B. (2012). Sentiment Analysis: A Fascinating Problem. In *Sentiment Analysis and Opinion Mining*.
- Mubarok, M. S., Adiwijaya, A., & Aldhi, M. D. (2017). Aspect-based sentiment analysis to review products using Naïve Bayes. *AIP Conference Proceedings*, 1867(August).
- Naquitasia, R., Fudholi, D. H., & Iswari, L. (2022). Analisis Sentimen Berbasis Aspek Pada Wisata Halal Dengan Metode Deep Learning. *16*, 156–164.

- 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.
- Puteri, Dian. (2023). Implementasi Long Short Term Memory (LSTM) dan Bidirectional Long Short Term Memory (BiLSTM) Dalam Prediksi Harga Saham Syariah. *Euler : Jurnal Ilmiah Matematika, Sains dan Teknologi*. 11. 35-43. 10.34312/euler.v11i1.19791.
- Saraswita, E. F., Rini, D. P., & Abdiansah, A. (2021). Analisis Sentimen E-Wallet di Twitter Menggunakan Support Vector Machine dan Recursive Feature Elimination. *Jurnal Media Informatika Budidarma*, 5(4), 1195.
- Shahriar, Fahim & Bashar, Md. (2021). Automatic Monitoring Social Dynamics During Big Incidences: A Case Study of COVID-19 in Bangladesh.
- Siti Masturoh, A. B. P. (2020). Sentiment Analysis Against the Dana E-Wallet on Google Play Reviews Using the K-Nearest Neighbor Algorithm. *Ejournal.Nusamandiri.Ac.Id*, 53–58.
- Turjaman, R. M., & Indonesia, U. (2022). Analisis Sentimen Berbasis Aspek Marketing Mix Terhadap Ulasan Aplikasi Dompot Digital (Studi Kasus : Aplikasi Linkaja Pada Twitter). 30(2), 266–275.
- Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), 1–25.
- Zhou, K., & Long, F. (2018). Sentiment analysis of text based on cnn and bi-directional LSTM model. *ICAC 2018 - 2018 24th IEEE International Conference on Automation and Computing: Improving Productivity through Automation and Computing*, September, 1–5.



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 : 24/UN5.2.1.14/SK/SPB/2024

DEKAN FAKULTAS ILMU KOMPUTER
DAN TEKNOLOGI INFORMASI UNIVERSITAS SUMATERA UTARA

- Membaca** : Surat Permohonan Mahasiswa Fasilkom-TI USU tanggal 8 Januari 2024 perihal permohonan ujian skripsi
Nama : RINI HARYATI
NIM : 181402088
Program Studi : Sarjana (S-1) Teknologi Informasi
Judul Skripsi : Analisis Sentimen Berbasis Aspek Pada Ulasan E-Wallet Dana Dengan Algoritma CNN-BILSTM
- 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 : Sarah Purnamawati ST., MSc.
NIP: 198302262010122003
- Sekretaris : Niskarto Zendrato S.Kom., M.Kom
NIP: 198909192018051001
- Anggota Penguji : Dr. Erna Budhiarti Nababan M.IT
NIP: 196210262017042001
- Anggota Penguji : Dedy Arisandi ST., M.Kom.
NIP: 197908312009121002
- Moderator : -
- Panitem : -
- Kedua** : Segala biaya yang diperlukan untuk pelaksanaan kegiatan ini dibebankan pada Dana Penerimaan Buku 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

