PERINGKASAN TEKS UMPAN BALIK PELANGGAN MENGGUNAKAN LSTM DENGAN *WORD EMBEDDING*

SKRIPSI

Oleh :
MOCHAMAD HARIS SYAFIUDDIN
NIM. 210605110082

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025

PERINGKASAN TEKS UMPAN BALIK PELANGGAN MENGGUNAKAN LSTM DENGAN *WORD EMBEDDING*

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh :
MOCHAMAD HARIS SYAFIUDDIN
NIM. 210605110082

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025

HALAMAN PERSETUJUAN

PERINGKASAN TEKS UMPAN BALIK PELANGGAN MENGGUNAKAN LSTM DENGAN *WORD EMBEDDING*

SKRIPSI

Oleh :
MOCHAMAD HARIS SYAFIUDDIN
NIM. 210605110082

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 03 Juni 2025

Pembimbing I,	Pembimbing II,
Supriyono, M.Kom	Dr. Ririen Kusumawati, S.Si., M.Kom
NIP. 19841010 201903 1 012	NIP. 19720309 200501 2 002

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Dr. Ir. Fachrul Kurniawan, M.MT., IPU
NIP. 19771020 200912 1 001

HALAMAN PENGESAHAN

PERINGKASAN TEKS UMPAN BALIK PELANGGAN MENGGUNAKAN LSTM DENGAN *WORD EMBEDDING*

SKRIPSI

Oleh :
MOCHAMAD HARIS SYAFIUDDIN
NIM. 210605110082

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 20 Juni 2025

Susunan Dewan Penguji

Ketua Penguji	: Dr. Ir. Fachrul Kurniawan, M.MT., IPU NIP. 19771020 200912 1 001	( )
Anggota Penguji I	: Nur Fitriyah Ayu Tunjung Sari, M.Cs NIP. 19911226 202012 2 001	( )
Anggota Penguji II	: Supriyono, M.Kom NIP. 19841010 201903 1 012	( )
Anggota Penguji III	: Dr. Ririen Kusumawati, S.Si., M.Kom NIP. 19720309 200501 2 002	( )

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Dr. Ir. Fachrul Kurniawan, M.MT., IPU
NIP. 19771020 200912 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama	: Mochamad Haris Syafiuddin
NIM	: 210605110082
Fakultas / Prodi	: Sains dan Teknologi / Teknik Informatika
Judul Skripsi	: Peringkasan Teks Umpan Balik Pelanggan Menggunakan LSTM dengan *Word Embedding*

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 03 Juni 2025
Yang membuat pernyataan,

Mochamad Haris Syafiuddin
NIM. 210605110082

MOTTO

You don't know what the future holds, that's why its potential is infinite.

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Allah Subhanahu Wa Ta'ala atas segala Rahmat dan Hidayah-Nya, yang telah memungkinkan penulis untuk menyelesaikan studi di Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang, serta berhasil menyelesaikan penelitian Skripsi yang berjudul **Peringkasan Teks Umpan Balik Pelanggan Menggunakan LSTM dengan *Word Embedding*** dengan baik.

Penulis menyadari bahwa dalam penulisan dan penelitian Skripsi ini, banyak pihak yang telah berperan penting dalam memberikan bimbingan, semangat, serta dukungan baik secara moral maupun materiil. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Prof. Dr. H. M. Zainuddin, MA selaku Rektor Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
2. Prof. Dr. Sri Harini, M.Si selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
3. Dr. Ir. Fachrul Kurniawan, ST., M.MT., IPU selaku Ketua Prodi Teknik Informatika Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang sekaligus Ketua Penguji yang telah memberikan banyak saran dan juga dukungan untuk menyelesaikan skripsi ini.
4. Supriyono, M.Kom selaku Dosen Pembimbing I atas segala bimbingan, saran, kritik, waktu, kesabaran yang diberikan selama penulisan skripsi ini.
5. Dr. Ririen Kusumawati, S.Si., M.Kom selaku Dosen Pembimbing II atas bimbingannya selama penulisan skripsi ini hingga selesai dengan baik.

6. Nur Fitriyah Ayu Tunjung Sari, M.Cs selaku Dosen Penguji II yang telah menguji serta memberikan masukan dalam penulisan skripsi ini
7. Dosen dan Staff Program Studi Teknik Informatika UIN Malang atas ilmu dan pengalaman yang telah diberikan selama menempuh perkuliahan.
8. Bapak Andy Syamsudin dan Ibu Siti Choeriyah selaku orangtua penulis, yang dengan tulus memberikan doa, kasih sayang, dan dukungan tanpa henti, serta selalu menjadi sumber inspirasi, kekuatan, dan motivasi dalam setiap langkah hidup, sehingga saya dapat mencapai titik ini.
9. Andraini Puspita Wardhani dan Nabila Nur Rohmah selaku kakak kandung penulis, yang senantiasa memberikan bantuan dalam segala hal serta kepercayaan sehingga penulis mampu menyelesaikan skripsi dengan baik, juga Katon Prasetyo dan Dimas Sutapratama Hariyanto selaku kakak ipar yang telah memberikan dukungan emosional.
10. Adik Rezky Puspita yang mendampingi, memberikan bantuan dan semangat serta motivasi, baik dalam proses perkuliahan maupun di luar kegiatan akademik.
11. Rekan-rekan Teknik Informatika angkatan 21 “ASTER” yang memberikan kontribusi dalam bentuk intelektual dan semangat.

Malang, 06 Juni 2025

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

ABSTRAK

Syafiuddin, Mochamad Haris. 2025. **Peringkasan Teks Umpan Balik Pelanggan Menggunakan LSTM Dengan Word Embedding**. Skripsi. Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Supriyono, M.Kom. (II) Dr. Ririen Kusumawati, S.Si., M.Kom.

Kata kunci: peringkasan teks, umpan balik pelanggan, *Long Short-Term Memory* (LSTM), *word embedding.*

Umpan balik pelanggan sangat penting dalam pengambilan keputusan bisnis, membantu perusahaan memahami kepuasan konsumen, mengidentifikasi kelemahan produk, dan merumuskan strategi pemasaran yang lebih efektif. Pengelolaan umpan balik yang baik dapat meningkatkan pendapatan perusahaan hingga 15-25% melalui loyalitas pelanggan dan perbaikan produk. Umpan balik ini diperoleh dari media sosial, *e-commerce*, *platform review*, dan situs resmi perusahaan. Namun, dengan volume data yang besar, diperlukan solusi otomatis untuk menganalisis dan merangkum informasi penting dengan cepat dan akurat. Penelitian ini bertujuan untuk mengetahui penerapan LSTM dengan *word embedding* dalam peringkasan teks umpan balik pelanggan.. LSTM dipilih karena kemampuannya menangani data sekuensial, memahami konteks percakapan, dan memberikan respons yang relevan serta bermakna. Penelitian ini menggunakan arsitektur dua layer LSTM dengan berbagai skenario pengujian, termasuk percobaan dengan *word embedding* GloVe atau *FastText* dan tanpa *word embedding*. Hasil terbaik diperoleh tanpa *word embedding* melalui F1-score dengan ROUGE-1 score 0.3085, ROUGE-2 score 0.2282, dan ROUGE-L score 0.3064. Dengan demikian, penelitian ini berhasil membuktikan bahwa model LSTM tanpa *word embedding* dapat meringkas umpan balik pelanggan secara efektif.

ABSTRACT

Syafiuddin, Mochamad Haris. 2025. **Customer Feedback Text Summarization Using LSTM with Word Embedding.**. Thesis. Informatics Engineering Study Program, Faculty of Science and Technology, State Islamic University of Maulana Malik Ibrahim Malang. Advisor: (I) Supriyono, M.Kom. (II) Dr. Ririen Kusumawati, S.Si., M.Kom.

Keywords: text summarization, customer feedback, Long Short-Term Memory (LSTM), word embedding.

Customer feedback is crucial in business decision-making, helping companies understand consumer satisfaction, identify product weaknesses, and formulate more effective marketing strategies. Proper management of feedback can increase a company's revenue by 15-25% through customer loyalty and product improvements. This feedback is gathered from social media, e-commerce, review platforms, and company websites. However, with the large volume of data, automated solutions are required to analyze and summarize important information quickly and accurately. This study aims to investigate the application of LSTM with word embedding in summarizing customer feedback text. LSTM was chosen for its ability to handle sequential data, understand the context of conversations, and generate relevant and meaningful responses. The study uses a two-layer LSTM architecture with various testing scenarios, including experiments with GloVe or FastText word embedding and without word embedding. The best results were achieved without word embedding, with an F1-score and ROUGE-1 score of 0.3085, ROUGE-2 score of 0.2282, and ROUGE-L score of 0.3064. Thus, this study successfully proves that the LSTM model word embedding, is effective in summarizing customer feedback.

الملخص

شافي الدين ، محمد حارس. 2025. **تلخيص نص ملاحظات العملاء باستخدام LSTM مع تضمين الكلمات**. الأطروحة. برنامج دراسة هندسة المعلوماتية، كلية العلوم والتكنولوجيا، الجامعة الإسلامية الحكومية، مولانا مالك إبراهيم مالانج. المشرفان : سوبريونو، ماجستير في علوم الحاسوب. (2) الدكتورة ريرين كوسوماواتي، بكالوريوس في العلوم، ماجستير في علوم الحاسوب.

الكلمات المفتاحية: تلخيص النص، وملاحظات العملاء، والذاكرة طويلة المدى القصيرة الأجل (LSTM)، وتضمين الكلما

تُعد ملاحظات العملاء أمرًا بالغ الأهمية في عملية اتخاذ القرارات التجارية، حيث تساعد الشركات على فهم رضا العملاء، وتحديد نقاط ضعف المنتجات، وصياغة استراتيجيات تسويقية أكثر فعالية. يمكن أن تؤدي الإدارة الجيدة للتغذية الراجعة إلى زيادة إيرادات الشركة بنسبة 15-25% من خلال ولاء العملاء وتحسين المنتجات. يتم الحصول على هذه الملاحظات من وسائل التواصل الاجتماعي والتجارة الإلكترونية ومنصات المراجعات والمواقع الإلكترونية للشركات. ومع ذلك، مع وجود كميات كبيرة من البيانات، هناك حاجة إلى حل آلي لتحليل وتلخيص المعلومات المهمة بسرعة ودقة. يهدف هذا البحث إلى معرفة تطبيق LSTM مع تضمين الكلمات في تلخيص نص ملاحظات العملاء. وقد تم اختيار LSTM لقدرته على التعامل مع البيانات المتسلسلة، وفهم سياق المحادثة، وتقديم ردود ذات صلة وذات مغزى. يستخدم هذا البحث بنية LSTM ثنائية الطبقات مع سيناريوهات اختبار مختلفة، بما في ذلك تجارب مع تضمين كلمات GloVe أو FastText وبدون تضمين كلمات. وقد تم الحصول على أفضل النتائج بدون تضمين الكلمات من خلال النتيجة F1 بنتيجة ROUGE-1 بنتيجة 0.3085، و ROUGE-2 بنتيجة 0.2282، و ROUGE-L بنتيجة 0.3064. وبالتالي، أثبتت هذه الدراسة بنجاح أن نموذج LSTM بدون تضمين الكلمات يمكن أن يلخص ملاحظات العملاء بشكل فعال.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam era digital semua hal dapat dilakukan dengan lebih mudah dan efisien, banyak lini pekerjaan yang dapat dilakukan hanya dengan berdiam di rumah atau melalui jarak jauh. Teknologi saat ini yang menggunakan *software* juga berbagai alat digital lain mempermudah dalam berbagai aspek seperti komunikasi dan akses informasi yang dapat dimanfaatkan untuk berbagai kepentingan. Salah satu hal yang dapat dipermudah dengan era digital saat ini adalah umpan balik. Umpan balik (*feedback*) adalah proses dimana output dari suatu sistem, tindakan, atau perilaku dikembalikan sebagai informasi untuk memperbaiki atau mempertahankan performa sistem tersebut. Berbagai jenis umpan balik seperti verbal, non-verbal, konstruktif maupun negatif telah dipermudah dengan berbentuk kuesioner, rating, maupun komentar. Sehingga, pelanggan dapat memberi umpan balik dan perusahaan dapat menerima umpan balik melalui jarak jauh.

Umpan balik pelanggan telah menjadi salah satu elemen kunci dalam pengambilan keputusan bisnis. Perusahaan dari berbagai industri mengandalkan ulasan dan komentar pelanggan untuk memahami kepuasan konsumen, mengidentifikasi kelemahan produk atau layanan, serta menyusun strategi pemasaran yang lebih efektif (Liu, 2012). Studi terbaru menunjukkan bahwa umpan balik pelanggan yang dikelola dengan baik dapat meningkatkan pendapatan perusahaan hingga 15-25% melalui peningkatan loyalitas pelanggan dan strategi

perbaikan produk yang lebih tepat sasaran (Hallikainen *et al*., 2020). Umpan balik ini biasanya dikumpulkan dari berbagai sumber seperti media sosial, *e-commerce*, *platform review*, dan situs web resmi perusahaan (Pang & Lee, 2008). Namun, besarnya volume data yang dihasilkan setiap hari menimbulkan tantangan dalam analisis, karena membaca dan menginterpretasikan ribuan hingga jutaan ulasan pelanggan secara manual bukanlah tugas yang efisien. Oleh karena itu, diperlukan solusi otomatis yang mampu merangkum informasi penting dari umpan balik pelanggan dengan cepat dan akurat (Gambhir & Gupta, 2017).

Salah satu metode yang efektif dalam peringkasan teks otomatis adalah penggunaan model berbasis LSTM yang dikombinasikan dengan *word embedding*. LSTM merupakan jenis *Recurrent Neural Network* (RNN) yang dirancang untuk mengatasi masalah *vanishing gradient*, memungkinkan model untuk memahami konteks dalam data sekuensial seperti teks dalam jangka waktu yang lebih panjang (Hochreiter & Schmidhuber, 1997). Model ini telah terbukti berhasil dalam berbagai tugas pemrosesan bahasa alami, termasuk peringkasan teks, terjemahan otomatis, dan analisis sentimen. Di sisi lain, *word embedding* adalah teknik representasi kata dalam bentuk vektor numerik yang dapat menangkap makna semantik kata-kata dalam suatu teks. Algoritma seperti Word2Vec, GloVe, dan *FastText* telah banyak digunakan untuk meningkatkan pemahaman mesin terhadap teks dengan cara menyusun hubungan antar kata dalam ruang vektor yang lebih bermakna (Mikolov *et al*., 2013). Dengan menggunakan *word embedding*, model dapat lebih memahami hubungan kata-kata dalam umpan balik pelanggan, sehingga menghasilkan ringkasan yang lebih akurat dan informatif.

Sebagai dasar dalam pengembangan teknologi, sains dalam Islam memiliki peran yang sangat penting, sebagaimana tercermin dalam Al-Qur'an yang memberikan banyak petunjuk tentang pencarian ilmu dan pengetahuan. Salah satu ayat yang terkait dengan pentingnya ilmu adalah Surah Al-Alaq (96:1-5), yang berbunyi:

اقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ ١ خَلَقَ الإِنسَانَ مِنْ عَلَقٍ ٢ اقْرَأْ وَرَبُّكَ الْأَكْرَمُ ٣ الَّذِي عَلَّمَ بِالْقَلَمِ ٤ عَلَّمَ الإِنسَانَ مَا لَمْ يَعْلَمْ ٥

"Bacalah dengan menyebut nama Tuhanmu yang menciptakan. Dia telah menciptakan manusia dari segumpal darah. Bacalah, dan Tuhanmulah yang Maha Pemurah, yang mengajarkan dengan pena. Dia mengajarkan manusia apa yang tidak diketahuinya." (QS. Al-Alaq 96:1-5).

Ayat ini menggambarkan pentingnya ilmu dan teknologi dalam kehidupan manusia, yang sesuai dengan semangat penelitian ini dalam mengembangkan sistem berbasis teknologi untuk analisis umpan balik.

Meskipun telah banyak penelitian, yaitu mengenai peringkasan teks, terdapat gap yang masih belum banyak dikaji dalam konteks peringkasan umpan balik pelanggan menggunakan LSTM dengan *word embedding*, sehingga penelitian ini terletak pada penerapan teknik peringkasan teks otomatis untuk menganalisis umpan balik pelanggan. Selain itu, penelitian ini berusaha mengoptimalkan kombinasi LSTM dan *word embedding*, serta mengeksplorasi pengaruh berbagai parameter *word embedding* dalam menghasilkan ringkasan umpan balik pelanggan. Dengan pendekatan ini, diharapkan dapat memberikan kontribusi dalam pengembangan model peringkasan teks yang lebih efektif dan efisien, serta meningkatkan kemampuan perusahaan dalam menganalisis kepuasan pelanggan dan merancang strategi bisnis yang lebih berbasis data.

1.2 Rumusan Masalah

Bagaimana penerapan LSTM dengan *word embedding* dalam peringkasan teks umpan balik pelanggan?

1.3 Tujuan Penelitian

Mengevaluasi penerapan LSTM dengan *word embedding* dalam peringkasan teks umpan balik pelanggan.

1.4 Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan yang ditetapkan untuk menjaga fokus dan validitas hasil yang diperoleh diantaranya adalah sebagai berikut:

a. Penelitian ini dibatasi pada pengembangan model peringkasan teks otomatis untuk umpan balik pelanggan yang berasal dari *platform e-commerce.*

b. Data yang digunakan hanya berupa teks ulasan pelanggan tanpa tambahan informasi seperti metadata atau gambar.

c. Pengembangan model difokuskan pada implementasi LSTM dengan representasi *word embedding* untuk peringkasan teks, tanpa mengintegrasikan teknik-teknik praproses atau post-proses lanjutan di luar cakupan dasar peringkasan.

d. Penelitian ini menggunakan 21.966 ulasan pelanggan yang diambil dari trustpilot.com yang terdapat di Kaggle, sehingga analisis dan hasil yang diperoleh hanya mencerminkan karakteristik serta gaya penulisan ulasan

yang khas dari *platform* tersebut dan mungkin tidak sepenuhnya mewakili ulasan dari *platform e-commerce* lainnya.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat strategis yang signifikan antara lain:

a. Dengan adanya peringkasan teks otomatis, diharapkan perusahaan dapat menganalisis umpan balik pelanggan dengan lebih cepat dan efisien dibandingkan metode manual, sehingga mempercepat pengambilan keputusan.

b. Menjadi referensi dan landasan bagi pengembangan model peringkasan teks otomatis yang dapat diaplikasikan pada berbagai sektor industri.

c. Menambah literatur dan pengetahuan mengenai penerapan teknik LSTM dan *word embedding* dalam konteks peringkasan teks, serta mendorong penelitian lebih lanjut di bidang ini.

BAB II

STUDI PUSTAKA

2.1 Penelitian Terkait

Ghibeche *et al*. (2024) melanjutkan eksplorasi dalam bidang *text summarization* dengan mengembangkan model berbasis LSTM yang dilengkapi mekanisme *attention* dan *word embedding*. Penelitian ini menguji efektivitas pendekatan tersebut pada dataset ulasan makanan, menunjukkan bahwa kombinasi LSTM dan *attention* mampu meningkatkan pemahaman konteks dan relevansi ringkasan yang dihasilkan.

Selain itu, Sahith *et al*. (2024) meneliti penerapan LSTM dan *word embedding* dalam analisis sentimen terhadap umpan balik pelanggan. Studi ini membuktikan bahwa penggunaan *word embedding* dapat memperbaiki pemetaan kata dalam data ulasan, sementara LSTM membantu meningkatkan akurasi analisis sentimen dan peringkasan otomatis. Temuan ini semakin memperkuat peran *deep learning* dalam merangkum teks dengan mempertimbangkan aspek sentimen dan relevansi informasi.

Muniraj *et al*. (2023) mengusulkan model HNTSumm, yaitu pendekatan hibrida yang menggabungkan *word embedding* dan RNN berlapis dua untuk menghasilkan ringkasan teks yang lebih akurat dan informatif. Model ini mampu mengatasi permasalahan transliterasi dalam peringkasan teks berita, yang dapat diadaptasi untuk menangani teks umpan balik pelanggan. Sementara itu, Joshi *et al*. (2023) dalam penelitian DeepSumm mengintegrasikan model topik dengan LSTM

untuk meningkatkan pemahaman terhadap konteks dalam teks. Penelitian ini menunjukkan bahwa kombinasi antara *sequence-to-sequence* (seq2seq) *networks* dengan model topik dapat meningkatkan akurasi peringkasan ekstraktif.

Gupta & Patel (2021) mengeksplorasi penggunaan *Latent Semantic Analysis* (LSA) dan BERT yang dikombinasikan dengan *Bidirectional* LSTM untuk menghasilkan ringkasan yang lebih representatif. Hasil penelitian ini menunjukkan bahwa pendekatan berbasis transformer seperti BERT dapat memberikan hasil yang lebih baik dibandingkan metode tradisional berbasis LSA dalam menangkap makna teks.

Dalam peringkasan teks abstraktif, Tomer & Kumar (2020) mengusulkan pendekatan berbasis seq2seq dengan mekanisme perhatian *(attention mechanism).* Penelitian ini menyoroti keunggulan model LSTM dalam memahami konteks teks yang panjang, serta pentingnya mekanisme perhatian dalam menangkap informasi esensial. Senada dengan itu, Wazery *et al*. (2022) menggunakan LSTM yang dikombinasikan dengan Word2Vec dan *Continuous Bag-of-Words* (CBOW) untuk meningkatkan representasi kata dalam peringkasan teks berbahasa Arab.

Suleiman & Awajan (2020) meneliti pendekatan RNN dan *word embedding* untuk menghasilkan ringkasan teks yang lebih alami dan padat. Sementara itu, Liang *et al*. (2020) memperkenalkan *Selective Reinforced Seq2Seq Attention Model*, yang memungkinkan model untuk secara selektif fokus pada bagian teks yang lebih penting selama proses peringkasan. Metode ini meningkatkan kualitas ringkasan dengan mempertimbangkan relevansi antar kata dalam teks.

Dalam penelitian lain, Bani-Almarjeh & Kurdy (2023) mengkombinasikan RNN dan Transformer dengan *word embedding* untuk menangani peringkasan teks dalam berbagai bahasa, termasuk bahasa Arab. Model berbasis Transformer terbukti lebih unggul dalam memahami hubungan antar kata dalam konteks yang lebih luas. Sementara itu, Gambhir & Gupta (2022) meneliti peringkasan ekstraktif berbasis *Bidirectional* LSTM dengan mekanisme perhatian tingkat kata, yang mampu meningkatkan pemilihan kalimat kunci dalam teks panjang.

Dari berbagai penelitian tersebut, dapat disimpulkan bahwa penggunaan LSTM dan *word embedding* telah memberikan kontribusi signifikan dalam meningkatkan kualitas peringkasan teks. Kombinasi seq2seq models, *attention mechanisms*, dan *word embeddings* semakin populer dalam menghasilkan ringkasan teks yang akurat dan informatif. Oleh karena itu, penelitian ini berfokus pada penerapan LSTM dan *word embedding* dalam peringkasan teks umpan balik pelanggan, dengan harapan dapat meningkatkan pemahaman terhadap kebutuhan dan kepuasan pelanggan melalui analisis teks yang lebih efisien dan otomatis.

Tabel 2. 1 Penelitian Terdahulu

No	Peneliti	Judul	Metode	Hasil
1	Ghibeche et al. (2024)	*Deep Learning for Automatic Text Summarization*	LSTM + *Attention* + *Word Embedding*	Menggunakan mekanisme *attention* dan *word embedding* pada *dataset* ulasan makanan untuk menghasilkan ringkasan otomatis.
2	Sahith et al. (2024)	*Enhanced Sentiment Analysis of Product Feedback: Leveraging Deep Learning Techniques*	LSTM + *Word Embedding*	Menggunakan *word embedding* untuk pemetaan kata dalam umpan balik pelanggan dan meningkatkan akurasi analisis sentimen melalui LSTM.
3	Muniraj et al. (2023)	*HNTSumm: Hybrid Text Summarization of Transliterated News Articles*	*Word Embedding* + *2-layer RNN*	Menghasilkan ringkasan lebih akurat dengan pendekatan hibrid antara ekstraktif dan abstraktif.

No	Peneliti	Judul	Metode	Hasil
4	Joshi et al. (2023)	*DeepSumm: Exploiting Topic Models and Sequence-to-Sequence Networks for Extractive Text Summarization*	*LSTM + Topic Modeling*	Menggabungkan model topik dengan LSTM untuk meningkatkan kualitas ringkasan ekstraktif.
5	Gupta & Patel (2021)	*Method of Text Summarization Using LSA and Sentence-Based Topic Modeling with BERT*	LSA + BERT + *Bidirectional* LSTM	Model menggabungkan BERT dan LSA untuk memahami konteks lebih baik dalam peringkasan teks.
6	Tomer & Kumar (2020)	*Improving Text Summarization Using an Ensembled Approach Based on Fuzzy with LSTM*	*Fuzzy Logic* + LSTM	Menggunakan pendekatan *fuzzy* dan LSTM untuk meningkatkan akurasi peringkasan teks.
7	Wazery et al. (2022)	*Abstractive Arabic Text Summarization Based on Deep Learning*	LSTM + Word2Vec + CBOW	Menggunakan representasi kata untuk meningkatkan pemahaman konteks dalam peringkasan teks.
8	Suleiman & Awajan (2020)	*Deep Learning-Based Abstractive Text Summarization*	*RNN + Word Embedding*	Menggunakan model RNN dan *embedding* kata untuk merangkum teks panjang.
9	Liang et al. (2020)	*Abstractive Social Media Text Summarization Using Selective Reinforced Seq2Seq Attention Model*	Seq2Seq + *Attention Mechanism*	Memperkenalkan mekanisme perhatian selektif untuk meningkatkan akurasi peringkasan.
10	Bani-Almarjeh & Kurdy (2023)	*Arabic Abstractive Text Summarization Using RNN-Based and Transformer-Based Architectures*	RNN + *Transformer + Word Embedding*	Menggunakan kombinasi RNN dan Transformer untuk meningkatkan pemahaman konteks.
11	Gambhir & Gupta (2022)	*Deep Learning-Based Extractive Text Summarization with Word-Level Attention Mechanism*	*Bidirectional* LSTM + *Attention*	Menggunakan mekanisme perhatian tingkat kata untuk meningkatkan pemilihan kalimat penting.

2.2 Umpan Balik Pelanggan

Umpan balik pelanggan merupakan sumber informasi yang sangat penting dalam bisnis, karena dapat membantu perusahaan memahami kebutuhan dan kepuasan pelanggan (Muniraj *et al*., 2023). Informasi yang terkandung dalam umpan balik sering kali berupa teks panjang yang tidak terstruktur, sehingga teknik

pemrosesan bahasa alami sangat diperlukan untuk mengekstrak informasi penting secara efisien (Joshi *et al*., 2023). Analisis terhadap umpan balik ini memungkinkan perusahaan untuk mengidentifikasi aspek produk atau layanan yang perlu ditingkatkan serta mengetahui pola kepuasan pelanggan berdasarkan opini yang diberikan. Dengan menerapkan metode seperti text mining dan sentiment analysis, perusahaan dapat secara otomatis mengklasifikasikan umpan balik pelanggan menjadi kategori positif, negatif, atau netral, yang berguna dalam strategi pemasaran dan pengambilan keputusan bisnis (Gupta & Patel, 2021).

Selain memberikan manfaat bagi perusahaan, umpan balik pelanggan juga bermanfaat bagi calon pelanggan yang ingin mengetahui reputasi suatu perusahaan. Pelanggan cenderung mencari ulasan atau opini dari pelanggan sebelumnya untuk menilai kredibilitas dan kualitas produk atau layanan sebelum melakukan transaksi (Wazery *et al*., 2022). Perusahaan dengan ulasan positif yang konsisten cenderung memiliki daya tarik yang lebih tinggi dibandingkan dengan perusahaan yang banyak menerima umpan balik negatif. Oleh karena itu, memiliki sistem yang mampu merangkum umpan balik pelanggan secara efektif dapat membantu perusahaan dalam membangun citra yang lebih baik serta meningkatkan transparansi kepada konsumen (Elsaid *et al*., 2022).

2.3 Peringkasan Teks

Peringkasan teks bertujuan untuk mereduksi ukuran teks dengan tetap mempertahankan informasi esensial. Berdasarkan metode yang digunakan, peringkasan teks dapat dikategorikan sebagai ekstraktif atau abstraktif (Gupta & Patel, 2021). Peringkasan ekstraktif dilakukan dengan memilih dan menyusun

kembali kalimat-kalimat penting dari teks asli tanpa mengubah struktur bahasanya. Metode ini sering kali menggunakan algoritma berbasis frekuensi kata, TF-IDF, atau pendekatan berbasis grafik seperti TextRank untuk mengidentifikasi bagian teks yang paling representatif (Elsaid *et al*., 2022). Di sisi lain, peringkasan abstraktif menghasilkan ringkasan dengan membangun kembali kalimat baru yang tetap mempertahankan makna asli teks. Model *sequence-to-sequence* (Seq2Seq) dengan arsitektur LSTM dan mekanisme perhatian *(attention mechanism)* telah banyak digunakan dalam peringkasan abstraktif untuk meningkatkan keterbacaan dan koherensi hasil ringkasan (Tomer & Kumar, 2020).

Seiring berkembangnya teknologi kecerdasan buatan dan pemrosesan bahasa alami, model-model peringkasan teks semakin canggih dalam menangkap konteks dan menghasilkan ringkasan yang lebih akurat. Kombinasi teknik *word embedding* seperti Word2Vec, GloVe, dan BERT dengan model berbasis LSTM telah terbukti meningkatkan kualitas peringkasan teks, terutama dalam memahami makna kata dalam berbagai konteks (Wazery *et al*., 2022). Selain itu, penelitian terbaru menunjukkan bahwa pendekatan berbasis Transformer, seperti BERTSUM dan T5, memberikan performa yang lebih baik dibandingkan dengan model RNN tradisional dalam peringkasan teks (Bani-Almarjeh & Kurdy, 2023). Dengan meningkatnya volume data teks di era digital, pengembangan metode peringkasan teks yang efisien dan akurat menjadi semakin penting untuk berbagai aplikasi, termasuk analisis umpan balik pelanggan, pencarian informasi, dan pembuatan laporan otomatis.

2.3.1 Peringkasan Ekstraktif

Peringkasan ekstraktif dilakukan dengan memilih kalimat-kalimat yang paling relevan dari teks asli berdasarkan berbagai metrik seperti *Term Frequency-Inverse Document Frequency* (TF-IDF), *Latent Semantic Analysis* (LSA), atau algoritma berbasis grafik seperti TextRank (Tomer & Kumar, 2020). Pendekatan ini bekerja dengan mengidentifikasi kalimat-kalimat yang mengandung informasi paling penting dan menyusunnya menjadi sebuah ringkasan tanpa mengubah struktur bahasa aslinya. Beberapa teknik populer yang sering digunakan dalam peringkasan ekstraktif melibatkan pendekatan berbasis statistik, seperti frekuensi kata, serta model berbasis jaringan saraf seperti *Bidirectional Long Short-Term Memory* (BiLSTM) yang dapat menentukan bobot kepentingan suatu kalimat dalam konteks teks secara lebih akurat (Elsaid *et al*., 2022). Selain itu, algoritma graph-based seperti TextRank bekerja dengan membangun grafik hubungan antar kalimat dalam dokumen, lalu menghitung tingkat kepentingan setiap kalimat berdasarkan koneksi dengan kalimat lainnya (Elsaid *et al*., 2022).

Meskipun metode ekstraktif memiliki keunggulan dalam kecepatan dan kemudahan implementasi, terdapat beberapa keterbatasan dalam penggunaannya. Ringkasan yang dihasilkan sering kali terasa kurang alami karena terdiri dari potongan-potongan kalimat yang langsung diambil dari teks asli tanpa adanya penyusunan ulang (Bani-Almarjeh & Kurdy, 2023). Selain itu, pendekatan ini tidak mampu menangkap makna tersirat atau melakukan reformulasi kalimat seperti pada metode peringkasan abstraktif. Oleh karena itu, banyak penelitian mencoba meningkatkan performa peringkasan ekstraktif seperti transformer-based models

(BERTSUM, T5), yang dapat lebih memahami hubungan semantik antar kalimat (Wazery *et al*., 2022). Dalam berbagai aplikasi, peringkasan ekstraktif sering digunakan dalam analisis berita, laporan bisnis, serta peringkasan dokumen hukum, di mana informasi harus tetap akurat dan tidak boleh diubah dari teks aslinya.

2.3.2 Peringkasan Abstraktif

Berbeda dengan metode ekstraktif, metode peringkasan abstraktif menghasilkan ringkasan dengan menyusun ulang teks menggunakan model seperti *sequence-to-sequence* (seq2seq) *dengan* LSTM (Suleiman & Awajan, 2020). Metode ini meniru cara manusia merangkum teks dengan memahami inti informasi terlebih dahulu, lalu menyusunnya kembali dalam bentuk yang lebih ringkas dan koheren. Dalam model seq2seq, terdapat dua komponen utama, yaitu encoder yang membaca teks input dan decoder yang menghasilkan ringkasan berdasarkan pemahaman dari encoder. Untuk meningkatkan akurasi, sering digunakan attention mechanism, yang memungkinkan model untuk lebih fokus pada bagian penting dari teks saat menghasilkan ringkasan (Elsaid *et al*., 2022). Dengan kemampuan memahami konteks yang lebih dalam, metode abstraktif dapat menangkap hubungan antar kata dan frasa, sehingga lebih fleksibel dalam menghasilkan ringkasan dibandingkan metode ekstraktif yang hanya memilih kalimat dari teks asli tanpa melakukan reformulasi (Wazery *et al*., 2022).

Pemilihan metode abstraktif dalam peringkasan umpan balik pelanggan lebih disukai dibandingkan metode ekstraktif karena sifatnya yang lebih adaptif terhadap variasi bahasa dalam ulasan pelanggan. Umpan balik pelanggan sering kali mengandung bahasa informal, opini subjektif, atau ekspresi emosional yang tidak

selalu dapat ditangkap dengan metode ekstraktif (Bani-Almarjeh & Kurdy, 2023). Dengan menggunakan pendekatan abstraktif, model menyaring informasi penting, menyusun ulang kalimat, dan menghilangkan bagian yang kurang relevan, sehingga hasil ringkasan lebih mudah dipahami oleh perusahaan dalam menganalisis kepuasan pelanggan. Selain itu, pendekatan ini memungkinkan personalisasi dalam peringkasan, misalnya dengan menyesuaikan tingkat detail yang diinginkan dalam ringkasan berdasarkan kebutuhan bisnis. Meskipun lebih kompleks dalam implementasinya, keunggulan metode abstraktif dalam menghasilkan ringkasan yang lebih ringkas, bermakna, dan alami(Gupta & Patel, 2021).

2.4 *Long Short-Term Memory*

Long Short-Term Memory (LSTM) merupakan jenis jaringan saraf berulang disebut juga RNN yang dirancang untuk mengatasi masalah *vanishing gradient*, yaitu hilangnya informasi selama proses propagasi balik dalam jaringan RNN tradisional (Almuzaini & Azmi, 2020). Model ini bekerja dengan menggunakan tiga jenis gerbang utama, yaitu *input gate, output gate,* dan *forget gate* yang berfungsi untuk mengontrol aliran informasi yang disimpan atau dibuang pada setiap langkah waktu (Elsaid *et al*., 2022). Dengan adanya mekanisme ini, LSTM dapat mempertahankan informasi dalam jangka panjang, sehingga lebih efektif dalam menangani data sekuensial seperti teks. Kemampuan ini menjadikan LSTM sebagai model yang sangat populer dalam berbagai aplikasi pemrosesan bahasa alami, termasuk analisis sentimen, penerjemahan mesin, dan peringkasan teks (Gupta & Patel, 2021).

Pemilihan LSTM dalam peringkasan umpan balik pelanggan didasarkan pada kemampuannya dalam memahami hubungan kontekstual antar kata dalam ulasan pelanggan yang sering kali memiliki struktur bahasa yang tidak teratur. Umpan balik pelanggan sering kali berisi frasa yang panjang, informasi subjektif, serta ekspresi emosional yang memerlukan pemahaman mendalam untuk menghasilkan ringkasan yang akurat (Wazery *et al*., 2022). Dengan kemampuannya dalam menyimpan dan mengingat informasi dalam jangka panjang, LSTM lebih unggul dibandingkan model tradisional seperti RNN yang kesulitan dalam memproses informasi yang berada di posisi awal teks ketika panjang sekuens meningkat (Gupta & Patel, 2021). Selain itu, dengan mengombinasikan *word embedding* seperti *FastText* atau GloVe, model LSTM dapat menangkap hubungan semantik antar kata, memungkinkan model untuk memahami konteks dalam ulasan pelanggan dan menghasilkan ringkasan yang lebih bermakna (Bani-Almarjeh & Kurdy, 2023).

Jika dibandingkan dengan model lain seperti *Gated Recurrent Unit* (GRU) atau *Transformer-based models* (BERTSUM, T5), masing-masing memiliki kelebihan dan kekurangannya sendiri. GRU, sebagai varian dari LSTM, memiliki struktur yang lebih sederhana karena hanya menggunakan dua gerbang utama, sehingga lebih ringan secara komputasi tetapi kurang mampu menangani dependensi jangka panjang seperti LSTM (Elsaid *et al*., 2022). Di sisi lain, model berbasis Transformer, seperti BERTSUM dan T5, memiliki keunggulan dalam memahami hubungan kata secara global dengan mekanisme self-attention, yang memungkinkan pemrosesan lebih cepat dibandingkan LSTM yang bekerja secara

sekuensial (Tomer & Kumar, 2020). Namun, model Transformer membutuhkan *dataset* yang jauh lebih besar dan sumber daya komputasi yang tinggi untuk menghasilkan ringkasan yang berkualitas. Oleh karena itu, dalam konteks peringkasan umpan balik pelanggan, di mana teks memiliki panjang sedang dan memerlukan pemahaman konteks yang baik, LSTM tetap menjadi pilihan yang efektif karena keseimbangan antara performa, akurasi, dan efisiensi komputasi.

2.5 *Word Embedding*

Word embedding adalah teknik representasi kata dalam bentuk vektor berdimensi rendah yang menangkap hubungan semantik antar kata berdasarkan konteks penggunaannya dalam teks (Suleiman & Awajan, 2020). Berbeda dengan pendekatan tradisional seperti *bag-of-words* (BoW) atau *one-hot encoding*, *word embedding* mampu merepresentasikan kata dengan mempertimbangkan makna dan hubungan antar kata dalam suatu kalimat. Teknik ini sangat penting dalam peringkasan teks umpan balik pelanggan, karena ulasan pelanggan sering kali mengandung kata-kata dengan makna yang bergantung pada konteks (Gupta & Patel, 2021). Dengan menggunakan *word embedding*, model dapat memahami hubungan semantik antar kata, memungkinkan sistem untuk menghasilkan ringkasan yang lebih relevan dan bermakna. *Word embedding* juga memungkinkan transfer learning, di mana model dapat memanfaatkan *embedding* yang telah dilatih sebelumnya untuk meningkatkan akurasi peringkasan teks tanpa memerlukan *dataset* pelatihan yang sangat besar (Liang *et al*., 2020)

Dalam penelitian pemrosesan bahasa alami, beberapa metode *word embedding* yang umum digunakan adalah Word2Vec, GloVe, dan *FastText*.

Word2Vec menggunakan arsitektur *Continuous Bag-of-Words* (CBOW) dan Skip-gram untuk menghasilkan vektor kata berdasarkan kemunculannya dalam konteks tertentu (Gupta & Patel, 2021). Sementara itu, *FastText*, yang dikembangkan oleh *Facebook AI Research*, memperluas konsep Word2Vec dengan mempertimbangkan sub-kata (*subword information*), sehingga lebih efektif dalam menangani bahasa dengan morfologi yang kompleks, seperti bahasa aglutinatif atau bahasa yang memiliki banyak variasi bentuk kata (Bani-Almarjeh & Kurdy, 2023). Namun, dalam konteks peringkasan teks umpan balik pelanggan, metode GloVe (*Global Vectors for Word Representation*) sering kali lebih disukai karena kemampuannya dalam menangkap hubungan statistik antar kata dalam skala korpus yang besar dan menghasilkan representasi kata yang lebih stabil (Liang *et al*., 2020)

Pemilihan GloVe dan *FastText* sebagai metode *embedding* dalam peringkasan umpan balik pelanggan digunakan sebagai bagian dari skenario pengujian untuk mengevaluasi performa kedua teknik tersebut dalam memahami teks yang bersifat informal dan dinamis. GloVe memanfaatkan matriks co-occurrence global untuk memahami hubungan antar kata secara lebih luas. Dengan pendekatan ini, GloVe dapat memberikan pemahaman yang lebih baik tentang asosiasi kata dalam dataset besar berdasarkan frekuensi dan konteks kata yang lebih umum (Pennington *et al*., 2014). Sebagai contoh, dalam ulasan pelanggan, GloVe dapat menangkap hubungan yang lebih mendalam antar kata yang sering muncul bersama.

Di sisi lain, *FastText* mengatasi keterbatasan yang dimiliki oleh metode *embedding* lainnya dengan membangun representasi berbasis sub-kata (subword).

FastText sangat efektif untuk menangani kata-kata yang jarang muncul, kata baru, atau kata yang memiliki ejaan yang salah, yang sering kali ditemukan dalam umpan balik pelanggan. *FastText* dapat memperbaiki keterbatasan yang mungkin ada pada GloVe dalam menangani variasi kata yang lebih tidak terstruktur atau tidak dikenal (Bojanowski *et al*., 2017).

Dalam skenario pengujian ini, GloVe digunakan untuk memahami hubungan kata-kata dalam konteks global yang lebih luas, sementara *FastText* diuji untuk kemampuannya dalam mengatasi kata-kata tidak umum atau yang sering muncul dalam bentuk variasi tertentu. Dengan menguji kedua metode ini, diharapkan dapat ditemukan metode yang memberikan keseimbangan terbaik antara pemahaman konteks yang mendalam (seperti yang dilakukan oleh GloVe) dan kemampuan untuk menangani kata-kata yang lebih fleksibel dan tidak standar (seperti yang dilakukan oleh *FastText*).

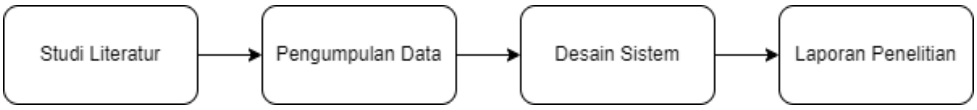
Hasil dari perbandingan kedua metode ini diharapkan dapat memberikan insight lebih dalam mengenai bagaimana keduanya dapat digunakan untuk menghasilkan ringkasan umpan balik yang lebih relevan dan sesuai dengan konteks bisnis yang berubah-ubah (Ma *et al*., 2022).

BAB III

METODOLOGI PENELITIAN

3.1 Prosedur Penelitian

Bab ini menjelaskan langkah-langkah metodologi yang digunakan dalam penelitian peringkasan teks ulasan pelanggan. Penelitian ini terdiri dari beberapa langkah yang dilaksanakan secara sistematis agar tujuan penelitian dapat tercapai dengan optimal. Prosedur penelitian yang disusun bertujuan untuk panduan yang merinci tahap-tahap penelitian serta memberikan penjelasan umum mengenai setiap langkahnya.



Gambar 3. 1 Prosedur Penelitian

Pada Gambar 3.1 menunjukkan rangkaian tahapan penelitian yang diawali dengan studi literatur untuk memperoleh landasan teori dan informasi pendukung, dilanjutkan dengan pengumpulan data ulasan pelanggan yang diperoleh dari *dataset* yang tersedia di *platform* seperti Kaggle. Setelah itu, sistem yang dirancang mencakup perancangan arsitektur dan spesifikasi teknis sesuai kebutuhan penelitian, dan diakhiri dengan analisis hasil untuk menilai keberhasilan penelitian.

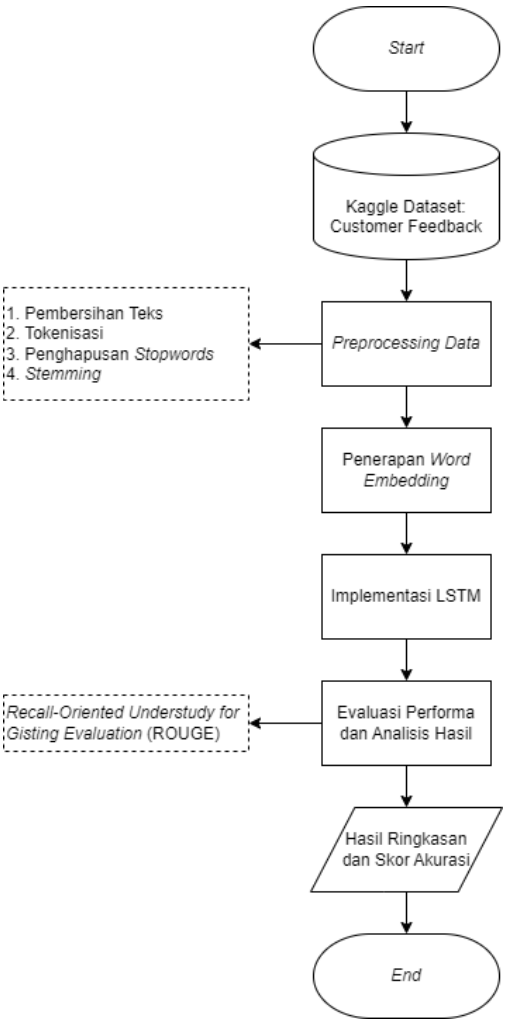
3.1.1 Pengumpulan Data

Data ulasan pelanggan diperoleh dari *dataset* yang tersedia di *platform* Kaggle berisi 21.966 ulasan pelanggan yang diambil dari trustpilot.com. Proses

pengumpulan data dimulai dengan mengunduh *dataset*, lalu dilakukan eksplorasi awal untuk memahami struktur data, termasuk identifikasi kolom teks ulasan, rating produk, serta fitur lain yang berpotensi digunakan dalam pengembangan model (Gupta & Patel, 2021).

3.1.2 Desain Sistem

Pendekatan penelitian ini menggabungkan metode eksperimen dan analisis dalam pengembangan model LSTM dengan teknik *word embedding*.

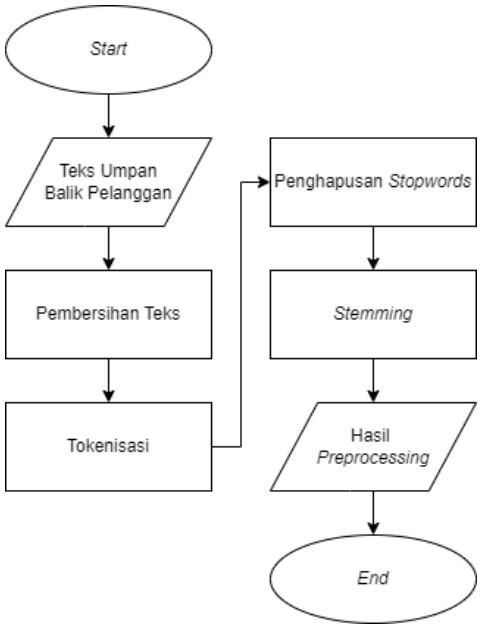


Gambar 3. 2 Desain Sistem

Pada Gambar 3.2 menggambarkan alur proses peringkasan teks umpan balik pelanggan menggunakan model LSTM dengan *word embedding*, dimulai dari pengambilan data ulasan produk secara daring dari Kaggle. Selanjutnya adalah *preprocessing data*, meliputi pembersihan teks, tokenisasi, penghapusan *stopwords*, dan *stemming*, serta penerapan *word embedding* untuk merepresentasikan kata secara vektor. Setelah data siap, model LSTM diimplementasikan untuk menghasilkan ringkasan teks otomatis. Keluaran model tersebut kemudian dievaluasi menggunakan metrik ROUGE guna mengukur kualitas ringkasan. Hasil akhir berupa ringkasan teks dan skor akurasi yang menunjukkan seberapa baik model dalam meringkas umpan balik pelanggan.

3.1.3 *Preprocessing* Data

Tahap *preprocessing* data merupakan langkah penting untuk menyiapkan data mentah menjadi bentuk yang siap untuk penerapan *word embedding*.



Gambar 3. 3 *Flowchart Preprocessing*

Pada Gambar 3.3 memperlihatkan alur *preprocessing* teks umpan balik pelanggan yang dimulai dengan pembersihan teks (menghapus karakter khusus, spasi berlebih, dan tanda baca), dilanjutkan dengan tokenisasi untuk memecah teks menjadi satuan kata, kemudian penghapusan *stopwords*. Hasil akhir dari rangkaian proses ini adalah data teks yang sudah terstruktur dan siap digunakan dalam tahapan pemodelan lebih lanjut, seperti penerapan LSTM untuk peringkasan teks.

a. Pembersihan Teks

Pada tahap pertama, teks yang terdapat dalam kolom *review_head* dibersihkan untuk menghilangkan elemen yang tidak relevan, seperti tanda baca, angka yang tidak penting, simbol khusus, dan *whitespace* berlebih. Misalnya, ulasan seperti:

"Purchased this for my device, it worked as advertised. You can never have too much phone memory, since I download a lot of stuff..." akan dibersihkan dengan menghapus tanda baca dan angka yang tidak memberikan nilai semantik, sehingga menjadi:

"purchased this for my device it worked as advertised you can never have too much phone memory since i download a lot of stuff". Pembersihan ini memastikan bahwa hanya informasi penting yang tersisa untuk tahap analisis selanjutnya.

b. *Tokenisasi*

Setelah proses pembersihan, teks dipecah menjadi unit-unit terkecil yang disebut token, biasanya berupa kata. Proses tokenisasi dilakukan dengan memecah teks berdasarkan spasi atau menggunakan library pemrosesan bahasa alami, seperti

NLTK atau spaCy. Sebagai contoh, teks yang telah dibersihkan pada langkah sebelumnya diubah menjadi daftar token:

["purchased", "this", "for", "my", "device", "it", "worked", "as", "advertised", "you", "can", "never", "have", "too", "much", "phone", "memory", "since", "i", "download", "a", "lot", "of", "stuff"]

c. Penghapusan *Stopwords*

Langkah selanjutnya adalah menghapus stopwords, yaitu kata-kata umum yang tidak memberikan banyak informasi atau konteks, seperti *"this", "for", "my", "it", "as", "a", "of"*, dan sebagainya. Penghapusan stopwords dari daftar token di atas menghasilkan daftar kata seperti [*"purchased", "device", "worked", "advertised", "never", "much", "phone", "memory", "download", "lot", "stuff"*], yang memungkinkan analisis lebih fokus pada kata-kata yang memiliki makna penting.

d. *Stemming*

Pada tahap ini, setiap token diubah ke bentuk dasarnya untuk menyederhanakan variasi kata, seperti mengonversi kata "*worked*" dan "*working*" menjadi "*work*" menggunakan algoritma seperti *Porter Stemmer*. Proses ini membuat daftar token menjadi lebih konsisten, contohnya: [*"purchase", "device", "work", "advertise", "never", "much", "phone", "memory", "download", "lot", "stuff"*], sehingga mengurangi redundansi dan meningkatkan kualitas input bagi model.

3.1.4 Penerapan *Word Embedding*

Langkah terakhir dalam *preprocessing* adalah mengonversi token-token teks menjadi representasi vektor numerik menggunakan teknik *word embedding*. Pada penelitian ini, digunakan metode GloVe atau *FastText* untuk menghasilkan vektor kontinu yang dapat menangkap makna semantik serta hubungan antar kata.

a. GloVe

GloVe (*Global Vectors for Word Representation*) adalah metode berbasis matriks yang mengoptimalkan faktorisasi matriks dari statistik kata-kata yang muncul bersama dalam korpus besar (Pennington *et al*., 2014).

$$J = \sum_{i,j=1}^{v} f(X_{ij}) \left(w_i^T \tilde{w}_j + b_i + \tilde{b}_j - logX_{ij} \right)^2 \quad (3.1)$$

Keterangan :
J : fungsi kerugian yang ingin diminimalkan selama pelatihan.
X_{ij} : fungsi *sigmoid*
$f(X_{ij})$: nilai *forget gate*
w_i^T : *hidden state* pada waktu t
$\tilde{w}_j$: vektor representasi kata i (kata utama).
b_i : bias untuk kata i
$\tilde{b}_j$: bias untuk kata j
$logX_{ij}$: Logaritma frekuensi pasangan kata i dan j

.

Tujuan dari GloVe adalah untuk meminimalkan fungsi kerugian J, yang mengukur kesalahan antara prediksi dan nilai sebenarnya dari probabilitas pasangan kata.

b. *FastText*

FastText adalah sebuah model yang dikembangkan oleh *Facebook* untuk memanfaatkan pembelajaran representasi kata dengan memecah kata menjadi sub-kata (n-gram). Ini sangat berguna untuk menangani kata-kata yang tidak ada dalam

korpus atau untuk kata-kata yang memiliki bentuk yang berbeda (seperti dalam kasus bahasa dengan infleksi tinggi). *FastText* menggunakan representasi kata dengan membaginya menjadi sub-kata atau n-gram, dan setiap kata diwakili oleh jumlah representasi dari n-gram tersebut (Ermawan & Cahyono, 2025). Rumus untuk *FastText* adalah:

$$v_w = \frac{1}{|g|}\sum_{i=1}^{|g|} v_{g_i} \tag{3.2}$$

Keterangan :
v_w : vektor rata-rata yang akan dihitung
$|g|$: sigmoid Jumlah n-gram dari kata *W*
v_{g_i} : vektor representasi untuk n-gram ke-*i*

Dengan menggunakan token yang sudah diproses, seperti *["purchase", "device", "work", "advertise", "never", "much", "phone", "memory", "download", "lot", "stuff"*], *FastText* membagi kata-kata tersebut menjadi n-gram. Misalnya, kata "purchase" akan dibagi menjadi n-gram seperti ["pur", "ur", "rch", "ch", "ha", "as", "se"], dan setiap n-gram ini memiliki vektor representasinya sendiri.

v("pur") = [0.2, -0.1, 0.4, 0.5]

v("rch") = [-0.3, 0.2, 0.1, -0.2]

Untuk mendapatkan representasi kata "*purchase*", kita pertama-tama jumlahkan vektor-vektor n-gram yang mewakili kata tersebut dan kemudian menghitung rata-rata dari hasil jumlah tersebut, sehingga kita memperoleh representasi numerik yang lebih komprehensif dan representatif dari kata "*purchase*".

$$v_{purchase} = \frac{1}{7}(v_{pur} + v_{ur} + v_{rch} + v_{ch} + v_{ha} + v_{as} + v_{se})$$

Berikut contoh perhitungannya:

$v_{pur} = [0.2, -0.1, 0.4, 0.5]$

$v_{ur} = [0.1, 0.1, -0.2, 0.3]$

$v_{rch} = [-0.3, 0.2, 0.1, -0.2]$

$v_{ch} = [0.1, -0.2, 0.3, 0.1]$

$v_{ha} = [0.3, -0.1, -0.1, 0.2]$

$v_{as} = [0.2, 0.1, 0.0, -0.3]$

$v_{se} = [-0.1, 0.2, -0.3, 0.4]$

Kita dapat menghitung vektor untuk "purchase" dengan menjumlahkan semua vektor tersebut dan membaginya dengan jumlah n-gram:

$$v_{purchase} = \frac{1}{7}([0.2, -0.1,0.4,0.5] + [0.1,0.1, -0.2,0.3] + [-0.3,0.2,0.1, -0.2] + [0.1, -0.2,0.3,0.1] + [0.3, -0.1, -0.1,0.2] + [0.2,0.1,0.0, -0.3] + [-0.1,0.2, -0.3,0.4])$$

$$v_{purchase} = [0.1,0.07,0.04,0.13]$$

Setelah menjumlahkan semua vektor dan membaginya dengan 7, kita mendapatkan vektor representasi untuk kata "*purchase*":

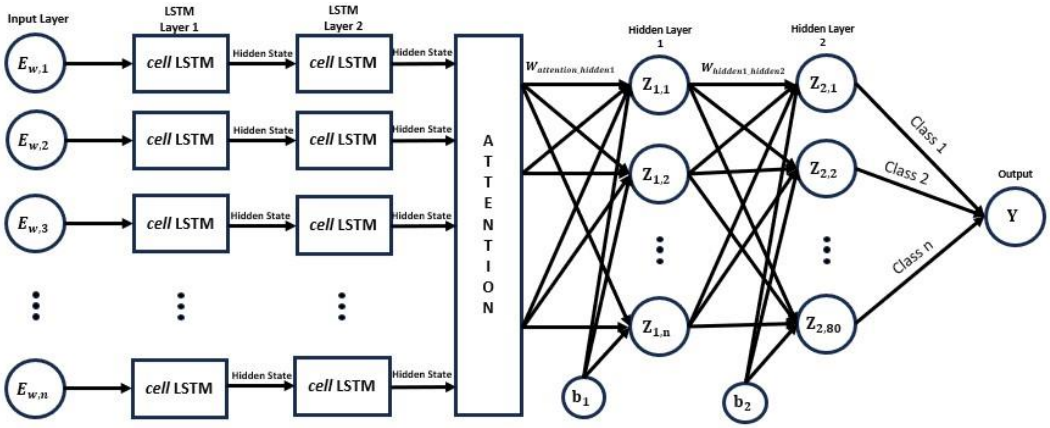
Penerapan GloVe atau *FastText* memungkinkan model LSTM untuk mengenali pola dan konteks dalam teks ulasan secara lebih efektif selama proses pelatihan, karena GloVe atau *FastText* telah dilatih pada korpus teks yang besar dan dapat menangkap konteks global dari setiap kata.

3.1.5 Model *Long Short-Term Memory*

Model LSTM dipilih karena kemampuannya dalam mengolah data sekuensial serta menangani permasalahan *vanishing gradient* yang sering ditemui pada RNN konvensional. Dalam konteks peringkasan umpan balik pelanggan, integrasi teknik *word embedding* memungkinkan setiap kata diubah menjadi representasi vektor numerik yang kaya informasi semantik, sehingga model dapat menangkap konteks dan hubungan antar kata secara lebih mendalam. Pendekatan ini tidak hanya mengandalkan urutan kata, tetapi juga memperhatikan arti dan keterkaitan semantik antar kata yang sangat krusial untuk menghasilkan ringkasan yang akurat dan padat. Dengan mengoptimalkan LSTM melalui representasi *embedding*, model dapat secara efektif menyaring informasi penting dari data mentah, mengurangi kebisingan, serta menangani variasi gaya bahasa dan idiomatik yang kerap muncul pada ulasan pelanggan. Hasilnya, ringkasan yang dihasilkan menjadi lebih representatif dan informatif, memberikan gambaran yang lebih jelas mengenai inti umpan balik pelanggan.

a. Arsitektur Model

Arsitektur model LSTM pada gambar 3.4 menggambarkan dua lapisan LSTM yang digunakan untuk memproses input berbasis urutan data. Model ini juga dilengkapi dengan mekanisme perhatian (attention mechanism) yang membantu fokus pada bagian penting dari input untuk menghasilkan prediksi yang lebih akurat. Setelah melalui beberapa lapisan, hasilnya diproses oleh lapisan tersembunyi dan akhirnya menghasilkan output yang terdiri dari beberapa kelas.



Gambar 3. 4 Arsitektur Model LSTM

Gambar 3.4 menggambarkan arsitektur model jaringan saraf yang memanfaatkan dua lapisan LSTM untuk memproses data urutan, yang diikuti oleh mekanisme perhatian (attention mechanism) dan lapisan tersembunyi untuk klasifikasi. Model dimulai dengan lapisan input yang menerima serangkaian vektor *embedding*, yang mewakili kata atau elemen dalam urutan data. Setiap vektor *embedding* ini kemudian diproses oleh sel LSTM pada dua lapisan berturut-turut. Pada lapisan pertama (LSTM *Layer* 1), data urutan diproses untuk menangkap hubungan temporal dan konteks antara elemen-elemen dalam urutan tersebut. Hasil dari LSTM Layer 1, yang berupa *hidden state*, kemudian diteruskan ke LSTM Layer 2 yang bertugas lebih lanjut mengolah informasi tersebut dan menghasilkan *hidden state* kedua yang lebih terperinci.

Selanjutnya, output dari kedua lapisan LSTM ini diteruskan ke mekanisme perhatian (*attention mechanism*), yang memainkan peran penting dalam memprioritaskan bagian-bagian tertentu dari input yang dianggap lebih relevan untuk tugas klasifikasi. Mekanisme perhatian ini memberikan bobot yang lebih

besar pada bagian informasi yang penting dan lebih kecil pada bagian yang kurang relevan. Setelah diproses melalui *attention*, informasi yang sudah difokuskan tersebut kemudian diteruskan ke dua lapisan tersembunyi (*hidden layers*) yang sepenuhnya terhubung (*fully connected*). Pada lapisan ini, informasi yang diproses diperkuat dan dipersiapkan untuk tahap klasifikasi akhir. Setiap lapisan tersembunyi menghasilkan representasi baru, yang kemudian digunakan untuk menghasilkan prediksi akhir.

Akhirnya, lapisan keluaran (*output layer*) menghubungkan hasil dari lapisan tersembunyi ke prediksi kelas akhir model. Output ini berupa prediksi kelas untuk setiap input yang diberikan, yang dapat berupa beberapa kelas. Secara keseluruhan, model ini menggunakan kombinasi LSTM, mekanisme perhatian, dan lapisan tersembunyi untuk menangkap informasi temporal, memperhatikan bagian relevan dari input, dan menghasilkan prediksi klasifikasi yang lebih akurat.

b. Mekanisme Kerja LSTM

Pada inti LSTM terdapat empat komponen utama, yaitu *input gate*, *forget gate*, *candidate cell state*, dan *output gate*, yang bekerja secara bersamaan untuk memproses informasi dalam urutan data. Berikut adalah contoh perhitungan LSTM yang mengilustrasikan mekanisme kerja pada salah satu token dari kalimat yang telah diproses (*preprocessed*). Misalnya, setelah *preprocessing*, kalimat ulasan "*Purchased this for my device, it worked as advertised. You can never have too much phone memory since I download a lot of stuff*" telah diubah menjadi "*purchase device work advertise never much phone memory download lot stuff*". Proses ini memungkinkan LSTM untuk menangkap informasi dari urutan kata yang relevan

dalam kalimat, yang kemudian digunakan untuk menghasilkan prediksi atau output berdasarkan pemahaman konteks yang telah dipelajari sebelumnya.

Kita akan mengambil token pertama, yaitu "purchase", untuk dilakukan perhitungan. Misalkan representasi vektor dari token "purchase" (hasil dari GloVe) adalah $x_t = [0.5,0.3]$, dan *hidden state* dari waktu sebelumnya adalah $h_{t-1} = [0.2,0.1]$.

Input gate (i_t) menentukan seberapa banyak informasi baru dari input x_t yang akan ditambahkan ke *cell state.* Perhitungannya adalah:

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \quad (3.3)$$

Gabungkan h_{t-1} dan x_t sehingga membentuk vektor input gabungan:

$$h_{t-1}, x_t = [0.2, 0.1, 0.5, 0.3]$$

Misalkan parameter untuk *input gate* adalah:

$$W_i = [0.1, 0.2, 0.3, 0.4]$$
$$b_i = 0.05$$

Maka perhitungan *input gate* adalah:

$$i_t = \sigma(0.1 \times 0.2 + 0.2 \times 0.1 + 0.3 \times 0.5 + 0.4 \times 0.3 + 0.05)$$
$$i_t = \sigma(0.02 + 0.02 + 0.15 + 0.12 + 0.05)$$
$$i_t = \sigma(0.36)$$
$$i_t = 0.59$$

Keterangan :
i_t : nilai *input gate*
σ : fungsi *sigmoid*
W_i : matriks bobot untuk *input*
h_{t-1} : *hidden state* dari waktu sebelumnya
x_t : input pada waktu t
b_i : bias *input*

Persamaan 3.3 menunjukkan bagaimana *Input Gate* menentukan seberapa banyak informasi baru dari input x_t yang harus ditambahkan *ke cell state*; di sini, *hidden state* sebelumnya h_{t-1} dan *input* saat ini x_t digabungkan (*concatenation*), kemudian diolah melalui transformasi linear dengan matriks bobot W_i dan bias b_i, sebelum diterapkan fungsi aktivasi sigmoid σ yang menghasilkan nilai antara 0 dan 1 sebagai pengontrol kontribusi informasi baru.

Forget gate berfungsi untuk menentukan bagian dari *cell state* sebelumnya yang harus dilupakan atau dipertahankan. Persamaannya adalah:

$$f_t = \sigma(W_f[h_{t-1},x_t]+b_f) \quad (3.4)$$

Misalkan parameter untuk *forget gate* adalah:

$W_f = [0.2, 0.1, 0.4, 0.3]$

$b_f = 0.1$

Perhitungan *forget gate* adalah:

$f_t = \sigma(0.2 \times 0.2 + 0.1 \times 0.1 + 0.4 \times 0.5 + 0.3 \times 0.3 + 0.1)$

$f_t = \sigma(0.04 + 0.01 + 0.2 + 0.09 + 0.1)$

$f_t = \sigma(0.44)$

Keterangan :

f_t	: nilai *forget gate*
σ	: fungsi *sigmoid*
W_f	: matriks bobot untuk, *forget*
h_{t-1}	: *hidden state* dari waktu sebelumnya
x_t	: input pada waktu t
b_f	: bias *forget*

Persamaan 3.4 mengatur *Forget Gate* yang memutuskan bagian mana dari *cell state* sebelumnya yang harus dipertahankan atau dilupakan; kombinasi h_{t-1} dan x_t dioperasikan secara linear menggunakan bobot W_f dan bias b_f, lalu fungsi

sigmoid σ memberikan nilai antara 0 dan 1 untuk menentukan tingkat penghapusan informasi lama.

Selanjutnya, informasi baru yang dihasilkan dari kombinasi input x_t dan *hidden state* h_{t-1} diproses untuk membentuk kandidat *cell state*. Proses ini dilakukan dengan menggunakan fungsi aktivasi tanh, yang bertujuan untuk menghasilkan nilai kandidat yang akan dipertimbangkan untuk diperbarui dalam *cell state*. Fungsi tanh memberikan output dalam rentang antara -1 dan 1, yang memastikan bahwa nilai-nilai dalam *cell state* tetap terkendali dan tidak tereskalasi terlalu besar. Fungsi aktivasi tanh ini bertindak sebagai penghalus, mencegah informasi yang tidak relevan atau berlebihan mengganggu proses pembaruan dalam jaringan.

$$\tilde{C}_t = \tanh\left(W_C \times [h_{t-1}, x_t] + b_C\right) \tag{3.5}$$

Misalkan parameter untuk kandidat *cell state* adalah:

$W_C = [0.3, 0.2, 0.1, 0.4]$

$b_C = 0.0$

Maka perhitungannya:

$$\tilde{C}_t = \tanh\left(0.3 \times 0.2 + 0.2 \times 0.1 + 0.1 \times 0.5 + 0.4 \times 0.3\right)$$

$$\tilde{C}_t = \tanh\left(0.06 + 0.02 + 0.05 + 0.12\right)$$

$$\tilde{C}_t = \tanh\left(0.25\right)$$

$$\tilde{C}_t = 0.244$$

Keterangan :

$\tilde{C}_t$: *candidate cell state*
W_c : bobot untuk *candidate cell state*
h_{t-1} : *hidden state* dari waktu sebelumnya
x_t : input pada waktu t
b_C : bias *candidate*

Persamaan 3.5 pada kandidat *cell state* $\tilde{C}_t$ dihasilkan dengan menggabungkan h_{t-1}dan x_t melalui transformasi linear dengan bobot W_C dan bias b_C, kemudian fungsi aktivasi tanh mengubah nilai hasilnya ke dalam rentang −1 hingga 1. Proses ini menghasilkan representasi informasi baru yang lebih bervariasi, yang memungkinkan model untuk menangkap pola yang lebih kompleks dan relevan dalam proses pembelajaran.

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (3.6)$$

Misalkan *cell state* dari waktu sebelumnya adalah

$C_{t-1} = 0.5$

Maka perhitungannya:

$C_t = 0.61 \times 0.5 + 0.59 \times 0.244$

$C_t = 0.449$

Keterangan :
C_t : *candidate cell state*
f_t : nilai *forget gate*
C_{t-1} : *cell state* dari waktu sebelumnya
i_t : nilai *input gate*
$\tilde{C}_t$: *candidate cell state* yang akan ditambahkan

Persamaan 3.6 memperbarui *cell state* C_t dengan menggabungkan dua komponen: bagian dari *cell state* lama C_{t-1} yang dipertahankan melalui *forget gate* f_t dan informasi baru dari kandidat *cell state* C_t yang diatur oleh *input gate* i_t; penjumlahan keduanya menghasilkan *cell state* yang mengintegrasikan memori masa lalu dengan data terbaru.

Langkah terakhir adalah menentukan *output* dari sel LSTM melalui *output gate*, yang mengontrol bagian dari *cell state* yang akan dikirim sebagai *hidden state*, memastikan informasi yang relevan diteruskan ke langkah berikutnya.

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o) \quad (3.7)$$

$$h_t = o_t \times tanh(C_t) \quad (3.8)$$

Misalkan parameter untuk *output gate* adalah:

$W_o = [0.2, 0.3, 0.3, 0.1]$
$b_o = 0.05$

Maka perhitungannya:

$o_t = \sigma(0.2 \times 0.2 + 0.3 \times 0.1 + 0.3 \times 0.5 + 0.1 \times 0.3 + 0.05)$
$o_t = \sigma(0.04 + 0.03 + 0.15 + 0.03 + 0.05)$
$o_t = \sigma(0.30)$
$o_t = 0.574$

Kemudian, *hidden state* dihitung dengan:

$h_t = 0.574 \times tanh(0.449)$

$h_t = 0.574 \times 0.421$

$h_t = 0.241$

Keterangan :
σ : fungsi *sigmoid*
$\tilde{C}_t$: *candidate cell state*
o_t : nilai *output gate*
h_t : *hidden state* pada waktu t
x_t : input pada waktu t
h_{t-1} : *hidden state* dari waktu sebelumnya
b_o : bias *output*
W_o : matriks bobot untuk *output gate*

Persamaan 3.7 mendefinisikan *Output Gate* yang mengontrol bagian dari *cell state* yang akan dijadikan output; dengan menggabungkan h_{t-1} dan x_t melalui

transformasi linear menggunakan bobot W_o dan bias b_o fungsi sigmoid σ menghasilkan nilai antara 0 dan 1 yang menentukan proporsi *cell state* yang dipilih untuk diteruskan ke *hidden state* berikutnya.

Persamaan 3.8 menghasilkan *hidden state* h_t dengan pertama-tama menerapkan fungsi tanh pada *cell state* yang telah diperbarui C_t untuk mengubahnya ke dalam rentang −1 hingga 1, kemudian hasilnya dikalikan dengan *output gate* $o_t$sehingga hanya informasi yang telah dipilih yang diteruskan sebagai output ke langkah waktu selanjutnya.

3.1.6 Pengujian dan Hasil Evaluasi Model

Implementasi model dilakukan dengan menggunakan *framework deep learning* seperti TensorFlow atau PyTorch (Gupta & Patel, 2021). *Dataset* yang telah melalui tahap *preprocessing* dibagi menjadi tiga bagian, yaitu *data training*, validasi, dan *testing*. *Data training* digunakan untuk melatih model, di mana model belajar dengan mengoptimalkan parameter-parameter untuk meminimalkan *loss function*. Setelah model dilatih, data validasi digunakan untuk mengevaluasi kinerja model selama proses pelatihan, tanpa mempengaruhi parameter model. Data validasi membantu memantau apakah model mulai mengalami *overfitting*, yaitu kondisi di mana model terlalu menyesuaikan diri dengan data pelatihan dan kehilangan kemampuan untuk menggeneralisasi pada data baru. Untuk mencegah *overfitting*, teknik *early stopping* diterapkan, yang menghentikan pelatihan lebih awal jika tidak ada peningkatan signifikan pada kinerja model di data validasi. Teknik ini memastikan bahwa model dihentikan pada titik optimal, sebelum terjebak dalam penyesuaian yang berlebihan terhadap data pelatihan. Setelah

pelatihan selesai, model diuji menggunakan data testing untuk mengukur kinerjanya secara objektif dengan data yang belum pernah dilihat sebelumnya. Evaluasi akhir dilakukan menggunakan metrik seperti ROUGE untuk menilai kualitas model. Pembagian dataset ini (*training,* validasi, dan *testing*) penting untuk memastikan bahwa model tidak hanya belajar dengan baik dari data yang ada, tetapi juga dapat menggeneralisasi dengan baik pada data baru, menghasilkan model yang lebih *robust* dan siap diterapkan pada masalah dunia nyata.

$$ROUGE - N = \frac{\sum_{S \in Referensi} \sum_{n-gram \in S} Count_{match}(n-gram)}{\sum_{S \in Referensi} \sum_{n-gram \in S} Count(n-gram)} \qquad (3.9)$$

Keterangan :

$\sum$	: Jumlah n-gram dari semua ringkasan referensi.
$S \in Referensi$	: Setiap ringkasan acuan dalam kumpulan referensi.
$n-gram \in S$	: Sekumpulan n kata berurutan.
$Count_{match}(n-gram)$	: Jumlah n-gram antara ringkasan dan referensi.
$Count(n-gram)$	: Jumlah n-gram dalam ringkasan referensi.

Persamaan 3.9 mengukur kesamaan antara n-gram pada ringkasan yang dihasilkan oleh model dengan n-gram pada ringkasan referensi. Dengan membandingkan elemen tersebut, persamaan ini memberikan gambaran mendetail sejauh mana kualitas ringkasan yang dihasilkan oleh model tersebut mendekati ringkasan referensi, baik dalam hal keakuratan maupun relevansi informasi. Hal ini memungkinkan evaluasi yang objektif terhadap kemampuan model dalam memahami konteks dari data. (Tomer & Kumar, 2020; Weinstein, 1994).

Setelah pelatihan, model menghasilkan ringkasan, contohnya seperti berikut: ringkasan model (candidate) berbunyi "device works well with adequate memory for downloads," sementara ringkasan referensi (ground truth) adalah "device works well with enough memory for downloading files."

Untuk mengevaluasi performa, dihitung metrik ROUGE-1 (unigram):

1. Tokenisasi Ringkasan

   Candidate: ["device", "works", "well", "with", "adequate", "memory", "for", "downloads"]

   Referensi*: ["device", "works", "well", "with", "enough", "memory", "for", "downloading", "files"]*

2. Identifikasi Token Tumpang Tindih

   Selanjutnya, kita cari token-token yang muncul di kedua ringkasan. Dalam kasus ini, token yang sama (dengan pencocokan string secara langsung) adalah:

   "device", "works", "well", "with", "memory", "for"

   Total terdapat 6 token yang tumpeng tindih, lalu total ada 9 token referensi, dan jumlah token pada ringkasan model 8 token. Pada token "*adequate*" tidak cocok dengan "*enough*", dan "*downloads*" tidak sama dengan "*downloading*". Juga, token "*files*" hanya ada pada referensi.

3. Perhitungan ROUGE-1 *Recall* dan *Precision*

   $$Recall = \frac{Jumlah\ token\ yang\ tumpang\ tindih}{Jumlah\ token\ pada\ ringkasan\ referensi} \qquad (3.10)$$

   $$Recall = \frac{6}{9}$$

   $$Recall = 0.667$$

   $$Precision = \frac{Jumlah\ token\ yang\ tumpang\ tindih}{Jumlah\ token\ pada\ ringkasan\ model} \qquad (3.11)$$

   $$Precision = \frac{6}{8}$$

   $$Precision = 0.75$$

4. Perhitungan ROUGE-1 *Recall* dan *Precision*

   F1-score merupakan rata-rata harmonis antara *precision* dan *recall*:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3.12)$$

$$F1 = 2 \times \frac{0.75 \times 0.667}{0.75 + 0.667}$$

$$F1 = 2 \times \frac{0.5}{1.417}$$

$$F1 = 0.707$$

   Dari perhitungan tersebut, didapatkan:

   ROUGE-1 *Recall*: 66.7%

   ROUGE-1 *Precision*: 75%

   ROUGE-1 *F1-score*: Sekitar 70.7%

Evaluasi menggunakan ROUGE bertujuan untuk mengukur kualitas hasil ringkasan dengan membandingkannya dengan ringkasan referensi yang telah dibuat secara manual. Metode ROUGE yang digunakan dalam penelitian ini mencakup ROUGE-N, yang mengukur kesamaan berdasarkan kemunculan n-gram (seperti unigram pada ROUGE-1 dan bigram pada ROUGE-2). Dengan menggunakan metrik ini, dapat dianalisis sejauh mana sistem mampu menghasilkan ringkasan yang informatif dan tetap mempertahankan esensi cerita asli (Tomer & Kumar, 2020).

Metrik ROUGE digunakan secara luas untuk mengevaluasi kualitas ringkasan teks yang dihasilkan oleh model dengan cara mengukur kesamaan antara ringkasan yang dihasilkan dan referensi yang telah ditentukan. Terdapat beberapa variasi dari metrik ROUGE yang digunakan untuk mengukur berbagai aspek

kinerja model dalam *text summarization*, yaitu ROUGE-1, ROUGE-2, dan ROUGE-L. ROUGE-1 mengukur kesamaan unigram (kata tunggal) antara ringkasan model dan referensi, memberikan gambaran tentang seberapa banyak kata penting yang dapat diekstrak oleh model dari teks sumber. Semakin tinggi skor ROUGE-1, semakin baik model dalam menangkap informasi dasar yang terkandung dalam teks asli. Sementara itu, ROUGE-2 mengukur kesamaan bigram (pasangan kata) antara ringkasan model dan referensi, sehingga menilai kemampuan model dalam mempertahankan konteks yang lebih luas dan hubungan antar kata. Skor ROUGE-2 yang lebih tinggi menunjukkan bahwa model berhasil menjaga hubungan antara kata-kata yang relevan dalam teks. Sedangkan ROUGE-L mengukur kesamaan berdasarkan *Longest Common Subsequence* (LCS), yaitu urutan kata yang muncul dengan urutan yang sama di kedua teks, meskipun tidak harus bersebelahan. ROUGE-L memberikan gambaran seberapa baik model mempertahankan urutan alami kalimat, yang berkontribusi pada koherensi dan alur dari ringkasan yang dihasilkan (ShafieiBavani et al., 2018).

Metrik ROUGE sering digunakan untuk mengukur kualitas ringkasan karena memberikan indikasi yang jelas tentang seberapa baik model dalam mengekstraksi informasi relevan dan menyusun teks dengan cara yang koheren. Untuk menilai hasil yang lebih menyeluruh, banyak peneliti menggunakan *F1-score* sebagai ukuran kinerja model dalam metrik ROUGE. *F1-score* adalah rata-rata harmonis dari *precision* dan *recall*, yang memberikan keseimbangan antara keduanya, sehingga menghindari bias yang mungkin timbul jika hanya menggunakan *precision* atau *recall* saja. Precision mengukur seberapa banyak kata

yang dipilih oleh model yang relevan, sedangkan *recall* mengukur seberapa banyak kata relevan yang berhasil dipilih oleh model. *F1-score* sangat berguna dalam konteks ROUGE karena memberikan gambaran yang lebih baik tentang bagaimana model seimbang dalam memilih kata yang relevan tanpa kehilangan konteks penting dari teks asli. Dengan demikian, penggunaan *F1-score* dalam mengukur ROUGE dapat memberikan hasil evaluasi yang lebih akurat dan representatif mengenai kualitas ringkasan yang dihasilkan (Alshawi, 2003).

Skor ROUGE yang lebih tinggi, baik itu ROUGE-1, ROUGE-2, maupun ROUGE-L, menunjukkan bahwa model lebih berhasil dalam mengekstrak informasi yang relevan dan menyajikan ringkasan yang mendekati referensi manual. Nilai ROUGE yang lebih tinggi tidak hanya menunjukkan kemampuan model untuk mengekstrak kalimat yang relevan tetapi juga mempertahankan makna asli dari teks sumber. Misalnya, nilai ROUGE-1 yang lebih tinggi menunjukkan bahwa model berhasil memilih unigram yang relevan, sementara ROUGE-2 menunjukkan kemampuan model untuk mempertahankan konteks yang lebih luas melalui bigram. Meskipun skor ROUGE sering kali bervariasi tergantung pada dataset dan jenis model, nilai yang lebih tinggi secara umum mencerminkan kemampuan model untuk menghasilkan ringkasan yang lebih akurat dan berguna. Namun, dalam beberapa kasus, skor yang lebih rendah, seperti di bawah 0.4, tetap dapat dianggap efektif, terutama untuk aplikasi yang lebih sederhana atau dengan dataset terbatas, seperti analisis umpan balik pelanggan, di mana informasi yang diekstraksi tetap berguna (Nenkova, 2011).

3.2 Skenario Pengujian

Pada bagian LSTM dalam model ini, dua lapisan LSTM digunakan untuk menangkap pola temporal dan representasi fitur dari data sekuensial. LSTM, yang merupakan jenis dari RNN, sangat cocok untuk memproses data sekuensial karena kemampuannya dalam menangani informasi kontekstual yang bergantung pada urutan waktu, seperti yang terdapat pada data teks.

Pada lapisan pertama LSTM, model diuji dengan dua konfigurasi jumlah *hidden state* yang berbeda, yaitu 64 unit pada skenario pertama dan 128 unit pada skenario kedua. Setiap unit dalam *hidden state* berfungsi untuk menyimpan informasi tentang status *internal* model pada setiap langkah waktu, memungkinkan model untuk mengingat pola yang relevan dari data sekuensial. Lapisan pertama LSTM ini bertugas untuk menangkap pola temporal yang lebih sederhana atau dasar, misalnya pola kata atau urutan kata yang muncul secara berulang dalam teks.

Setelah lapisan pertama, model melanjutkan dengan lapisan kedua LSTM, yang bertujuan untuk memperdalam representasi fitur yang telah diperoleh dari lapisan pertama. Pada lapisan kedua ini, model memiliki kapasitas yang lebih besar untuk mempelajari hubungan yang lebih kompleks antara elemen-elemen dalam data, karena jumlah *hidden state* yang digunakan lebih besar, yaitu 128 unit pada skenario pertama dan 256 unit pada skenario kedua. Lapisan kedua ini memungkinkan model untuk menangkap konteks yang lebih mendalam dan pola-pola kompleks dalam data teks yang mungkin tidak terdeteksi pada lapisan pertama.

Selain itu, kedua lapisan LSTM menggunakan *Bidirectional* LSTM, yang berarti bahwa model dapat memproses data sekuensial dalam dua arah, yaitu dari

kiri ke kanan dan dari kanan ke kiri. Hal ini memungkinkan model untuk memanfaatkan informasi kontekstual yang lebih lengkap, di mana kalimat atau urutan kata sebelumnya dan sesudahnya bisa saling melengkapi dalam memahami konteks kalimat secara menyeluruh. Dengan pendekatan ini, model tidak hanya memperhatikan urutan kata yang muncul sebelumnya, tetapi juga kata-kata yang mengikuti, yang dapat sangat penting untuk memahami makna teks secara keseluruhan.

Dalam rangka menguji performa model peringkasan teks umpan balik pelanggan, dilakukan eksperimen dengan pembagian *dataset* ulasan pelanggan ke dalam dua subset utama:

Data Training (90%) digunakan untuk melatih model LSTM dengan *word embedding*, di mana pada tahap ini model belajar mengenali pola, konteks, dan struktur bahasa dari ulasan pelanggan. Proses pelatihan mencakup penerapan *preprocessing* yang kemudian diinput ke dalam arsitektur LSTM untuk menghasilkan ringkasan, dengan *monitoring* terhadap *loss function* dan metrik evaluasi pada *data* validasi untuk mencegah *overfitting*. Sementara itu, *Data Testing* (10%) digunakan untuk menguji kemampuan generalisasi model yang telah dilatih, di mana model menghasilkan ringkasan dari *data testing* yang belum pernah dilihat selama proses pelatihan. Hasil ringkasan ini kemudian dievaluasi menggunakan metrik ROUGE untuk mengukur seberapa baik ringkasan yang dihasilkan mendekati ringkasan referensi yang telah disiapkan secara manual, memberikan gambaran yang jelas tentang seberapa baik model menangkap

informasi dari teks asli dan menghasilkan ringkasan yang sesuai dengan yang diinginkan.

Pembagian *data* secara acak dengan rasio 90:10 memastikan bahwa model memiliki jumlah data yang cukup untuk belajar (90% *data*) dan diuji secara objektif pada *data* yang belum pernah dilihat (10% *data*). Hal ini diharapkan dapat menghasilkan estimasi performa model yang lebih representatif terhadap kemampuan dalam menghasilkan ringkasan yang akurat dan informatif dari ulasan pelanggan.

3.2.1 Model LSTM dengan GloVe

GloVe adalah teknik *word embedding* yang mengubah kata-kata dalam teks menjadi representasi vektor berdimensi rendah yang mampu menangkap hubungan semantik antar kata. Dengan GloVe, model dapat mengonversi kata-kata dalam ulasan pelanggan menjadi vektor yang mewakili makna kata tersebut dalam konteks kalimat. Representasi ini sangat berguna, terutama dalam menangkap hubungan semantik yang lebih kompleks antar kata.

Dataset ulasan pelanggan yang telah melalui tahap *preprocessing* (termasuk penghapusan *stopwords*, dan tokenisasi) dimasukkan ke dalam model LSTM. GloVe memungkinkan model untuk memahami hubungan semantik antar kata yang lebih halus, sehingga model dapat menangkap pola yang lebih kompleks dalam ulasan pelanggan. GloVe memetakan kata-kata yang sering muncul bersama-sama dalam konteks yang sama ke dalam ruang vektor yang berdekatan, yang membantu model memahami bahwa kata-kata yang memiliki arti serupa akan memiliki representasi yang lebih dekat satu sama lain. Dengan GloVe, kata-kata yang

bersinonim atau memiliki makna yang saling terkait (misalnya "*good*" dan "*great*") dapat terhubung dengan cara yang lebih semantik dan relevan.

Skenario ini bertujuan untuk mengukur apakah penggunaan *word embedding* GloVe dapat meningkatkan kualitas ringkasan yang dihasilkan oleh model LSTM, serta bagaimana model dapat menangkap hubungan semantik yang ada dalam ulasan pelanggan. Diharapkan bahwa model yang menggunakan GloVe akan menghasilkan ringkasan yang lebih relevan, koheren, dan tepat sesuai dengan konteks yang diinginkan.

3.2.2 Model LSTM dengan *FastText*

Pada skenario kedua, model LSTM diimplementasikan menggunakan *word embedding FastText*. *FastText*, yang dikembangkan oleh *Facebook*, adalah teknik *word embedding* yang mirip dengan GloVe tetapi dengan perbedaan mendasar. *FastText* tidak hanya mengonversi kata-kata menjadi vektor, tetapi juga membangun representasi kata dari n-gram karakter. Hal ini memungkinkan *FastText* menangkap informasi yang lebih kaya tentang morfologi kata, termasuk kata-kata yang tidak ada dalam *vocabulary* model. Ini sangat berguna untuk menangani kata-kata baru atau kata-kata yang jarang muncul dalam teks.

Pada skenario ini, dataset ulasan pelanggan yang telah diproses dengan langkah-langkah seperti penghapusan stopwords, tokenisasi, dan lemmatization dimasukkan ke dalam model LSTM dengan *FastText*. Dengan *FastText*, model dapat menangkap hubungan semantik antar kata yang lebih dalam, termasuk kemampuan untuk menggeneralisasi kata yang jarang ditemukan dalam dataset. Keunggulan *FastText* ini dapat meningkatkan akurasi dalam menghasilkan

ringkasan yang lebih relevan dan koheren, meskipun ada kata-kata yang tidak dikenal sebelumnya.

Skenario ini bertujuan untuk mengukur apakah penggunaan *FastText* dapat memberikan peningkatan performa dalam menghasilkan ringkasan, terutama ketika model berhadapan dengan kata-kata yang jarang muncul atau tidak terdapat dalam kamus yang sudah ada.

3.2.3 Model LSTM tanpa *word embedding*

Pada skenario ketiga, model LSTM diterapkan tanpa menggunakan *word embedding* apapun. Dalam hal ini, model LSTM bekerja hanya dengan urutan kata dalam teks tanpa representasi semantik yang diberikan oleh GloVe atau *FastText*. Model hanya mengandalkan urutan kata dalam teks yang lebih sederhana, tanpa bantuan vektor representasi kata. Artinya, model harus memproses teks dalam bentuk yang lebih "mentah", tanpa pengetahuan semantik yang ditawarkan oleh teknik *word embedding*.

Dataset ulasan pelanggan yang telah melalui tahap *preprocessing* dimasukkan langsung ke dalam model LSTM tanpa pengolahan lebih lanjut. Tanpa *word embedding*, model akan kesulitan menangkap hubungan semantik antar kata, sehingga mungkin menghasilkan ringkasan yang kurang akurat atau kurang koheren. Sebagai contoh, kata-kata yang sering muncul bersama-sama dalam konteks yang sama (misalnya “*good*” dan “*great*”) akan dianggap berbeda oleh model tanpa adanya representasi semantik yang menghubungkan kedua kata tersebut.

BAB IV

HASIL DAN PEMBAHASAN

Pada bab ini, dibahas secara mendalam mengenai data input yang digunakan, yaitu umpan balik pengguna yang diperoleh dari Kaggle. Selanjutnya, dilakukan *preprocessing* teks yang meliputi pembersihan teks, tokenisasi, penghapusan *stopwords*, dan *stemming*. Setelah itu, diterapkan berbagai metode *word embedding*, yaitu GloVe, *FastText*, atau tanpa menggunakan keduanya. Langkah berikutnya adalah implementasi LSTM untuk peringkasan teks. Terakhir, hasil ringkasan yang diperoleh dievaluasi menggunakan metrik ROUGE *score*.

4.1 Data Input

Langkah pertama sebelum melakukan pengujian adalah mempersiapkan *data* yang akan digunakan. *Data* yang dimanfaatkan pada penelitian ini berasal dari *dataset* yang tersedia di *platform* Kaggle, yang berisi 21.966 ulasan pelanggan yang diambil dari trustpilot.com. Proses pengumpulan data dimulai dengan mengunduh *dataset*, lalu dilakukan identifikasi kolom teks ulasan, rating produk, serta fitur lain yang berpotensi digunakan dalam pengembangan model.

4.2 Hasil *Preprocessing*

Tahap *preprocessing* bertujuan untuk membersihkan dan menyusun data teks agar siap digunakan dalam analisis lebih lanjut. Pada tahap ini, berbagai teknik seperti pembersihan teks, tokenisasi, penghapusan *stopwords*, dan *stemming* diterapkan pada umpan balik pelanggan. Proses ini menghilangkan elemen yang

tidak relevan seperti tanda baca, angka, dan kata-kata umum yang tidak memberikan informasi tambahan. Hasil akhir adalah teks yang lebih padat dan fokus pada kata-kata yang memiliki makna penting.

Tabel 4. 1 Umpan Balik Pelanggan Sebelum *Preprocessing*

NO	***Human Summary***	***Customer Feedback***
1	The best in all that matters	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. The only one in crypto that offers you to actually call a phone number and get to someone with your questions just like that. Highly recommend! I have been a customer for more than a year and I have only good things to say about Celsius network.
2	Celsius Network ROCKS!	If you are looking for the best #HomeForCrypto and where to earn steady yield then there is no better place than Celsius Network. The app is easy to use and understand with a recent update as well. The company is led by a great CEO who is fully engaged with the Celsius community! The mission of Celsius is to do good then do well and they live up to that very much. Rewards compound and pay out weekly!
3	I despise it so much	I despise it so much. Transferring to other wallets is difficult, especially because you cannot swap your coins, and there is a long waiting period when you do send to someone. Verification within 24 hours. When compared to Crypto.com, I find it extremely difficult to gain access to your funds. I'm never going to use it again. My biggest blunder!

4.2.1 Hasil Pembersihan Teks

Tahap awal dalam *preprocessing* adalah pembersihan teks, yang meliputi penghapusan karakter khusus, spasi berlebih, tanda baca yang tidak diperlukan, konversi ke huruf kecil, serta penghilangan simbol khusus.

Tabel 4. 2 Umpan Balik Pelanggan Setelah Pembersihan Teks

NO	***Human Summary***	***Customer Feedback***
1	the best in all that matters	the best in all that matters its a great platform easy and simple to use and beginner friendly the only one in crypto that offers you to actually call a phone number and get to someone with your questions just like that highly recommend i have been a customer for more than a year and i have only good things to say about celsius network
2	celsius network rocks	if you are looking for the best homeforcrypto and where to earn steady yield then there is no better place than celsius network the app is easy to use and understand with a recent update as well the company is led by a great ceo who is fully engaged with the celsius

NO	*Human Summary*	*Customer Feedback*
		community the mission of celsius is to do good then do well and they live up to that very much rewards compound and pay out weekly
3	i despise it so much	i despise it so much transferring to other wallets is difficult especially because you cannot swap your coins and there is a long waiting period when you do send to someone verification within hours when compared to crypto com i find it extremely difficult to gain access to your funds im never going to use it again my biggest blunder

4.2.2 Hasil Tokenisasi

Setelah proses pembersihan teks selesai, langkah berikutnya adalah melakukan tokenisasi, yaitu memecah teks menjadi unit-unit terkecil berupa kata-kata. Proses ini memungkinkan analisis lebih lanjut terhadap masing-masing kata dalam teks, sehingga setiap kata dapat diproses dan dianalisis secara terpisah untuk keperluan pengolahan bahasa alami

Tabel 4. 3 Umpan Balik Pelanggan Setelah Tokenisasi

NO	*Human Summary*	*Customer Feedback*
1	the, best, in, all, that, matters	the, best, in, all, that, matters, its, a, great, platform, easy, and, simple, to, use, and, beginner, friendly, the, only, one, in, crypto, that, offers, you, to, actually, call, a, phone, number, and, get, to, someone, with, your, questions, just, like, that, highly, recommend, i, have, been, a, customer, for, more, than, a, year, and, i, have, only, good, things, to, say, about, celsius, network
2	celsius, network, rocks	if, you, are, looking, for, the, best, homeforcrypto, and, where, to, earn, steady, yield, then, there, is, no, better, place, than, celsius, network, the, app, is, easy, to, use, and, understand, with, a, recent, update, as, well, the, company, is, led, by, a, great, ceo, who, is, fully, engaged, with, the, celsius, community, the, mission, of, celsius, is, to, do, good, then, do, well, and, they, live, up, to, that, very, much, rewards, compound, and, pay, out, weekly
3	i, despise, it, so, much	i, despise, it, so, much, transferring, to, other, wallets, is, difficult, especially, because, you, cannot, swap, your, coins, and, there, is, a, long, waiting, period, when, you, do, send, to, someone, verification, within, hours, when, compared, to, crypto, com, i, find, it, extremely, difficult, to, gain, access, to, your, funds, im, never, going, to, use, it, again, my, biggest, blunder

4.2.3 Hasil Penghapusan *Stopwords*

Setelah proses tokenisasi selesai, langkah selanjutnya adalah penghapusan *stopwords*. *Stopwords* adalah kata-kata yang sering muncul dalam teks tetapi tidak memberikan informasi penting atau signifikan terhadap analisis, seperti "*and*," "*or*," "*with*," dan lainnya. Menghapus *stopwords* membantu mengurangi noise dalam data dan memungkinkan model untuk fokus pada kata-kata yang lebih relevan dan bermakna dalam konteks analisis yang dilakukan.

Tabel 4. 4 Umpan Balik Pelanggan Setelah Penghapusan *Stopwords*

NO	*Human Summary*	*Customer Feedback*
1	best, matters	best, matters, great, platform, easy, simple, use, beginner, friendly, crypto, offers, actually, call, phone, number, get, someone, questions, like, highly, recommend, customer, year, good, things, say, celsius, network
2	celsius, network, rocks	looking, best, homeforcrypto, earn, steady, yield, better, place, celsius, network, app, easy, use, understand, recent, update, company, led, great, ceo, fully, engaged, celsius, community, mission, celsius, good, well, live, rewards, compound, pay, weekly
3	despise, much	despise, much, transferring, wallets, difficult, especially, cannot, swap, coins, long, waiting, period, send, someone, verification, hours, compared, crypto, com, find, extremely, difficult, gain, access, funds, never, use, again, biggest, blunder

4.2.4 Hasil Penghapusan *Stopwords*

Langkah terakhir dalam *preprocessing* teks adalah proses *stemming*. Pada tahap ini, kata-kata yang telah diproses sebelumnya akan diubah menjadi bentuk dasarnya atau akar kata. Misalnya, kata "*running*" akan diubah menjadi "*run*," dan "*programming*" menjadi "*program*." Proses ini bertujuan untuk mengurangi variasi kata yang memiliki makna serupa, sehingga membantu meningkatkan konsistensi data dan mempermudah analisis lebih lanjut.

Tabel 4. 5 Umpan Balik Pelanggan Setelah *Stemming*

NO	*Human Summary*	*Customer Feedback*
1	best, matters	best, matter, great, platform, easy, simpl, use, beginner, friendli, crypto, offer, actual, call, phone, number, get, someon, question, like, high, recommend, custom, year, good, thing, say, celsius, network
2	celsius, network, rocks	look, best, homeforcrypto, earn, steady, yield, better, place, celsius, network, app, easy, use, understand, recent, updat, compani, led, great, ceo, fulli, engag, celsius, communiti, mission, celsius, good, well, live, reward, compound, pay, weekli
3	despise, much	despise, much, transfer, wallet, difficult, especial, cannot, swap, coin, long, wait, period, send, someon, verif, hour, compar, crypto, com, find, extrem, difficult, gain, access, fund, never, use, again, biggest, blunder

4.3 Skenario Uji Coba

Teknik *word embedding* seperti GloVe atau *FastText* akan diterapkan untuk mengonversi kata-kata dalam umpan balik menjadi representasi vektor berdimensi rendah yang lebih mudah diproses oleh model LSTM. Dalam penelitian ini, akan diuji pengaruh penerapan *word embedding* pada kualitas peringkasan teks umpan balik pelanggan, dengan membandingkan hasil dari model LSTM yang menggunakan *embedding* (baik GloVe atau *FastText*) dan model LSTM tanpa *embedding*.

4.3.1 Uji Coba *Word Embedding* GloVe

Setelah tahap *pre-processing* selesai, langkah selanjutnya adalah memproses dan membangun matriks embedding yang digunakan untuk merepresentasikan kata-kata dalam ruang vektor. Proses ini dilakukan dengan menggunakan data dari model *FastText*, yang dikenal karena kemampuannya menangani kata-kata yang tidak ada dalam kosakata model dengan cara memecah kata menjadi subkata. Matriks *embedding* yang dihasilkan dari model *FastText*

kemudian digunakan untuk merepresentasikan kata-kata dalam model pembelajaran mesin, memungkinkan pemahaman yang lebih baik terhadap hubungan semantik antar kata dan meningkatkan kinerja model dalam berbagai tugas pemrosesan bahasa alami.

Pseudocode 4. 1 Tabel Pemrosesan dan Penerapan *Word Embedding* GloVe

```
START

    Inisialisasi embedding_index sebagai dictionary kosong

    BUKA file '/kaggle/input/glove6b/glove.6B.300d.txt' DENGAN
    encoding 'utf-8'

    UNTUK setiap line DALAM file:

        Pisahkan line MENJADI kata-kata

        word = kata-kata[0]

        coefficients = kata-kata[1:]

        embedding = Konversi coefficients MENJADI array numpy
        DENGAN dtype float32

        embedding_index[word] = embedding

    TUTUP file

    Inisialisasi embedding_matrix DENGAN shape (length dari
    tokenizer.word_index + 1, 300)

    UNTUK setiap word, i DALAM tokenizer.word_index:

        JIKA word ADA DI embedding_index:

            embedding_matrix[i] = embedding_index[word]

        ELSE:

            embedding_matrix[i] = Nilai acak dari distribusi
            seragam ANTARA -0.25 dan 0.25

END
```

Pseudocode 4.1 menggambarkan proses pemrosesan *embedding* GloVe dan pembangunannya ke dalam matriks *embedding* yang digunakan untuk model pembelajaran mesin. Dimulai dengan membuka *file* yang berisi *embedding* GloVe dan memuat setiap kata beserta vektornya ke dalam *dictionary embedding_index*.

Kata menjadi *key* dan vektornya menjadi *value* dalam *dictionary* tersebut. Selanjutnya, sebuah matriks *embedding* kosong (*embedding_matrix*) dengan ukuran sesuai jumlah kata dalam tokenizer.word_index diinisialisasi. Untuk setiap kata dalam *word_index*, jika kata tersebut ditemukan dalam *embedding_index*, maka *embedding* dari kata tersebut disalin ke dalam matriks. Jika kata tersebut tidak ditemukan, maka vektor *embedding* diinisialisasi secara acak dengan nilai dari distribusi seragam antara -0.25 dan 0.25 dengan ukuran 300. Dengan demikian, matriks *embedding* siap digunakan untuk representasi kata dalam model.

Tabel 4. 6 Teks Setelah *Preprocessing* dan Nilai Vektor GloVe *Embedding*

NO	***Customer Feedback* Setelah *Preprocessing***	**Vektor GloVe *Embedding***
1	the best in all that matters ...	[0.53582895 0.46177535 0.46074079, ..., 0.479036]
2	if you are looking for the best homeforcrypto...	[0.45473589 0.46880985 0.44877779, ..., 0.55043441]
3	i despise it so much transferring to other wallets...	[0.54054268 0.52815656 0.51153046, ..., 0.49706281]

Tabel 4.6 memiliki dua kolom vektor representasi teks yaitu GloVe *embedding*, yang digunakan untuk menggambarkan teks pada kolom *preprocessed text*. Kolom pertama, *preprocessed text*, berisi paragraf yang telah *preprocessing*.

Kolom Vektor GloVe *Embedding* berisi vektor representasi teks berdasarkan model GloVe yang mengubah setiap kata dalam teks menjadi representasi vektor berdimensi tetap, kemudian menghitung rata-rata vektor-vektor kata tersebut untuk menghasilkan satu vektor representasi untuk seluruh paragraf. Dengan cara ini, GloVe menangkap hubungan statistik antar kata dalam teks dan menghasilkan vektor numerik yang mencerminkan makna keseluruhan teks.

4.3.2 Uji Coba *Word Embedding FastText*

Setelah tahap *preprocessing* selesai, langkah selanjutnya adalah memproses dan membangun matriks *embedding* menggunakan data dari model *FastText* yang kemudian digunakan dalam representasi kata pada model pembelajaran mesin.

Pseudocode 4. 2 Tabel Pemrosesan dan Penerapan *Word Embedding FastText*

```
START

    Inisialisasi embedding_index sebagai dictionary kosong

    BUKA file '/kaggle/input/fasttext/wiki-news-300d-1M-
    subword.vec' DENGAN encoding 'utf-8'

    UNTUK setiap line DALAM file:

        Pisahkan line MENJADI kata-kata

        word = kata-kata[0]

        coefficients = kata-kata[1:]

        embedding = Konversi coefficients MENJADI array numpy
DENGAN dtype float32

        embedding_index[word] = embedding

    TUTUP file

    Inisialisasi embedding_matrix DENGAN shape (length dari
    tokenizer.word_index + 1, 300)

    UNTUK setiap word, i DALAM tokenizer.word_index:

        JIKA word ADA DI embedding_index:

            embedding_matrix[i] = embedding_index[word]

        ELSE:

            embedding_matrix[i] = Nilai acak dari distribusi
            seragam ANTARA -0.25 dan 0.25 DENGAN ukuran (300)

END
```

Pseudocode 4.2 menjelaskan proses untuk memproses *embedding* kata dari *file FastText* dan kemudian membangunnya ke dalam matriks *embedding* yang digunakan untuk model. Dimulai dengan membuka file yang berisi *embedding FastText* dan memuat setiap kata beserta vektornya ke dalam *dictionary*

embedding_index, di mana kata-kata menjadi *key* dan vektornya menjadi *value*. Selanjutnya, sebuah matriks *embedding* kosong *embedding_matrix* diinisialisasi dengan ukuran yang sesuai dengan jumlah kata dalam tokenizer.word_index, dan setiap kata yang ditemukan dalam *dictionary embedding_index* akan diisi dengan *embedding* yang sesuai. Jika kata tidak ditemukan dalam *embedding_index*, maka vektornya akan diinisialisasi secara acak menggunakan distribusi seragam antara -0.25 dan 0.25 dengan dimensi 300. Dengan demikian, matriks *embedding* siap digunakan dalam model untuk representasi kata.

Tabel 4. 7 Teks Setelah *Preprocessing* dan Nilai Vektor *FastText Embedding*

NO	***Customer Feedback* Setelah *Preprocessing***	**Vektor *FastText Embedding***
1	the best in all that matters...	[0.57389807 0.5120527 0.55508093, ..., 0.52477316]
2	if you are looking for the best homeforcrypto...	[0.49095053 0.51555774 0.53726146, ..., 0.47566669]
3	i despise it so much transferring to other wallets...	[0.55041289 0.51807983 0.54433714, ..., 0.4884471]

Tabel 4.7 memiliki dua kolom vektor representasi teks, yaitu *FastText embedding* yang digunakan untuk menggambarkan teks pada kolom preproc*essed text*. Kolom pertama, *preprocessed text*, berisi paragraf yang telah *preprocessing*. Kolom Vektor *FastText Embedding* berisi vektor representasi teks yang dihitung menggunakan model *FastText*, sebuah metode pembelajaran representasi kata yang tidak hanya memperhitungkan kata-kata secara keseluruhan, tetapi juga memecah kata-kata menjadi subword atau bagian-bagian dari kata. Hal ini membuat *FastText* lebih efektif dalam menangani kata-kata yang tidak ada dalam kamus, atau kata yang jarang ditemukan (*out-of-vocabulary*), serta memberikan representasi yang lebih fleksibel dan dapat menangkap informasi morfologis dari kata-kata dalam teks.

4.4 Uji Coba Model LSTM

Setelah menerapkan word embedding menggunakan *GloVe*, *FastText*, atau tanpa *embedding*, langkah selanjutnya adalah membangun dan menerapkan model LSTM. Pada tahap ini, kita akan membahas bagaimana model LSTM dibangun, dikonfigurasi, dan dilatih menggunakan data pelatihan yang telah diproses.

Pseudocode 4.3 menjelaskan langkah-langkah dalam membangun, mengonfigurasi, dan melatih model LSTM. Proses dimulai dengan memeriksa apakah *pre-trained word embeddings* seperti GloVe atau *FastText* sudah tersedia. Jika *embeddings* yang sudah dilatih digunakan, lapisan *Embedding* pada model akan memanfaatkan matriks tersebut dan diset dengan parameter *trainable = False*, sehingga bobot pada lapisan ini tidak diperbarui selama pelatihan.

Jika tidak menggunakan *pre-trained embeddings*, lapisan *Embedding* akan diinisialisasi secara acak dan parameter trainable akan diset ke *True*. Dalam hal ini, bobot *embedding* akan dilatih dari awal selama pelatihan. Selanjutnya, model dilanjutkan dengan lapisan LSTM yang memiliki 128 unit untuk memproses urutan data. LSTM ini akan menangkap pola temporal dalam urutan teks dan mengembalikan seluruh urutan output dengan parameter *return_sequences=True.*

Untuk mencegah *overfitting*, lapisan *Dropout* dengan tingkat 0.5 ditambahkan. *Dropout* akan secara acak menonaktifkan sebagian unit selama pelatihan, membantu model untuk menggeneralisasi lebih baik. Lapisan terakhir adalah lapisan Dense yang memiliki satu unit dan menggunakan fungsi aktivasi sigmoid. Lapisan ini bertanggung jawab menghasilkan output berupa probabilitas untuk klasifikasi biner.

Pseudocode 4. 3 Tabel Membangun dan Mengonfigurasi Model LSTM

```
START
    Inisialisasi model sebagai Sequential()
    Tambahkan layer Embedding KE model DENGAN parameter:
    input_dim = len(tokenizer.word_index) + 1
        output_dim = 300, weights = embedding_matrix
    IF menggunakan_embedding_terlatih:
         trainable = False
    ELSE:
         trainable = True
    Tambahkan layer LSTM KE model DENGAN parameter:
         units = 128, return_sequences = True
    Tambahkan layer Dropout KE model DENGAN parameter:
         rate = 0.5
    Tambahkan layer Dense KE model DENGAN parameter:
         units = 1, activation = 'sigmoid'
    Kompilasi model DENGAN parameter:
         optimizer = Adam DENGAN learning_rate = 1e-4
         loss = binary_crossentropy
         metrics = accuracy
    Inisialisasi earlystop sebagai EarlyStopping() DENGAN
    parameter:
         monitor = 'val_loss'
         patience = 3
         restore_best_weights = True
    Latih model DENGAN parameter:
         X_train, y_train sebagai data pelatihan
         X_val, y_val sebagai data validasi
         epochs = 15, batch_size = 64, callbacks = [earlystop]
END
```

Model kemudian dikompilasi dengan menggunakan *optimizer* Adam pada laju pembelajaran 1e-4, fungsi kerugian *binary_crossentropy*, dan metrik akurasi untuk mengukur kinerja model. Agar tidak *overfitting*, digunakan *callback EarlyStopping* yang akan memantau val_loss dan menghentikan pelatihan jika tidak ada perbaikan setelah beberapa *epoch*. Fitur *restore_best_weights=True* memastikan bahwa bobot yang dipulihkan adalah bobot terbaik yang dicapai selama pelatihan.

Model kemudian dilatih dengan data pelatihan (X_train, y_train) dan validasi (X_val, y_val) selama maksimal 15 *epoch* dengan ukuran *batch* 64. Selama pelatihan, *callback EarlyStopping* membantu menentukan kapan pelatihan sebaiknya dihentikan untuk menghasilkan model yang paling optimal dan menghindari *overfitting*. Dengan cara ini, model dapat menghasilkan prediksi yang lebih akurat pada data yang tidak terlihat sebelumnya dan memiliki kemampuan generalisasi yang lebih baik.

4.4.1 Uji Coba LSTM dengan *Word Embedding* GloVe

Setelah tahap penerapan *word embedding* GloVe yang bertujuan untuk menghasilkan representasi kata yang lebih baik dan akurat, langkah selanjutnya dalam proses ini adalah melakukan model LSTM. Implementasi LSTM ini bertujuan untuk memanfaatkan kemampuan jaringan saraf dalam memproses dan memahami urutan data, yang sangat berguna untuk analisis teks yang memiliki dependensi temporal atau konteks yang saling berhubungan.

```
Epoch 1/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 115s 124ms/step - accuracy: 0.6973 - loss: 0.6099 - val_accuracy: 0.7700 - val_loss: 0.4980
Epoch 2/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 108s 124ms/step - accuracy: 0.7669 - loss: 0.5092 - val_accuracy: 0.7760 - val_loss: 0.4812
Epoch 3/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 107s 123ms/step - accuracy: 0.7780 - loss: 0.4880 - val_accuracy: 0.7797 - val_loss: 0.4752
Epoch 4/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 107s 122ms/step - accuracy: 0.7834 - loss: 0.4772 - val_accuracy: 0.7852 - val_loss: 0.4659
Epoch 5/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 108s 125ms/step - accuracy: 0.7908 - loss: 0.4664 - val_accuracy: 0.7917 - val_loss: 0.4622
Epoch 6/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 107s 123ms/step - accuracy: 0.7898 - loss: 0.4661 - val_accuracy: 0.7865 - val_loss: 0.4613
Epoch 7/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 108s 124ms/step - accuracy: 0.7943 - loss: 0.4582 - val_accuracy: 0.7959 - val_loss: 0.4536
Epoch 8/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 109s 125ms/step - accuracy: 0.7951 - loss: 0.4546 - val_accuracy: 0.7968 - val_loss: 0.4499
Epoch 9/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 107s 123ms/step - accuracy: 0.7990 - loss: 0.4452 - val_accuracy: 0.7997 - val_loss: 0.4506
Epoch 10/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 106s 122ms/step - accuracy: 0.8038 - loss: 0.4387 - val_accuracy: 0.7929 - val_loss: 0.4556
Epoch 11/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 108s 124ms/step - accuracy: 0.8055 - loss: 0.4374 - val_accuracy: 0.7994 - val_loss: 0.4478
Epoch 12/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 106s 122ms/step - accuracy: 0.8048 - loss: 0.4321 - val_accuracy: 0.8004 - val_loss: 0.4465
Epoch 13/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 105s 121ms/step - accuracy: 0.8097 - loss: 0.4270 - val_accuracy: 0.8059 - val_loss: 0.4415
Epoch 14/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 109s 125ms/step - accuracy: 0.8126 - loss: 0.4193 - val_accuracy: 0.8038 - val_loss: 0.4414
Epoch 15/15
870/870 ━━━━━━━━━━━━━━━━━━━━ 107s 123ms/step - accuracy: 0.8181 - loss: 0.4111 - val_accuracy: 0.8036 - val_loss: 0.4448
Training complete. Model and tokenizer saved.
194/194 ━━━━━━━━━━━━━━━━━━━━ 7s 35ms/step - accuracy: 0.8088 - loss: 0.4375
Test Accuracy: 80.38%
```

Gambar 4. 1 Hasil Proses Model LSTM dengan GloVe

Gambar 4.1 menunjukkan hasil pelatihan model LSTM dengan menggunakan *word embedding* GloVe pada tugas peringkasan teks umpan balik pelanggan. Selama 15 epoch pelatihan, model menunjukkan peningkatan kinerja yang stabil. Pada epoch pertama, akurasi pelatihan dimulai dari 69.73%, dan meningkat secara bertahap hingga mencapai 81.81% pada epoch terakhir. Begitu pula dengan akurasi validasi , yang meningkat dari 77.00% pada epoch pertama menjadi 80.36% pada epoch ke-15.

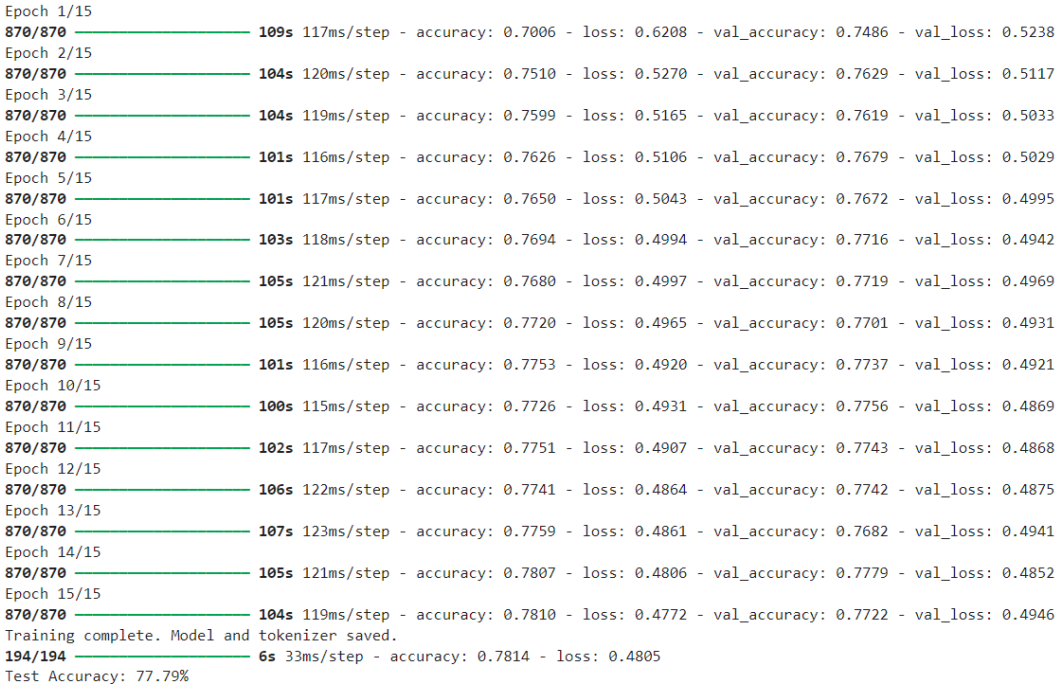
Penurunan *loss* yang terus-menerus menunjukkan bahwa model semakin baik dalam memprediksi teks. *Loss* pada pelatihan berkurang dari 0.6099 pada epoch pertama menjadi 0.4375 pada epoch terakhir, sementara *loss* pada validasi turun dari 0.4980 pada epoch pertama menjadi 0.4448 pada epoch ke-15.

Setelah pelatihan selesai, model berhasil mencapai akurasi pengujian 80.38%, yang mencerminkan kemampuan model dalam memproses dan merangkum data yang belum pernah dilihat sebelumnya. Hasil ini menunjukkan bahwa model LSTM yang menggunakan *embedding* kata dapat diterapkan dengan baik untuk peringkasan teks umpan balik pelanggan, dengan performa yang konsisten dalam meningkatkan akurasi dan mengurangi *error* (*loss*) selama pelatihan.

4.4.2 Uji Coba LSTM dengan *Word Embedding FastText*

Setelah menerapkan *word embedding* menggunakan *FastText* untuk menghasilkan representasi kata yang lebih tepat dan efektif, langkah berikutnya adalah mengimplementasikan LSTM. Penerapan LSTM ini dimaksudkan untuk memanfaatkan kemampuan jaringan saraf dalam menangani dan memahami urutan data, yang sangat penting dalam analisis teks yang memiliki ketergantungan temporal atau konteks yang saling terkait.

Gambar 4.2 menunjukkan hasil pelatihan model LSTM dengan menggunakan *word embedding FastText* pada tugas peringkasan teks umpan balik pelanggan. Selama 15 epoch pelatihan, model menunjukkan peningkatan kinerja yang signifikan. Misalnya, pada epoch pertama, akurasi pelatihan berada di 70.06%, dan pada epoch ke-15, akurasi ini meningkat menjadi 78.10%. Begitu pula dengan akurasi validasi yang menunjukkan tren peningkatan stabil dari 74.86% pada epoch pertama menjadi 77.22% pada epoch terakhir.



Gambar 4. 2 Hasil Proses Model LSTM dengan *FastText*

Penurunan *loss* yang terus terjadi sepanjang pelatihan menunjukkan bahwa model semakin baik dalam mengurangi kesalahan prediksi. *Loss* pada pelatihan turun dari 0.6208 pada epoch pertama menjadi 0.4772 pada epoch ke-15, sementara loss pada validasi juga menunjukkan penurunan dari 0.5238 pada epoch pertama menjadi 0.4946 pada epoch terakhir.

Setelah menyelesaikan pelatihan selama 15 epoch, model berhasil mencapai akurasi pengujian sebesar 77.79%, yang mencerminkan kemampuan model dalam memproses dan merangkum data yang belum pernah dilihat sebelumnya. Hasil ini menunjukkan bahwa model LSTM yang menggunakan *embedding* kata dapat diterapkan dengan baik untuk peringkasan teks umpan balik pelanggan, dengan performa yang baik dalam hal akurasi dan kemampuan untuk mengurangi *error* (*loss*) secara signifikan selama proses pelatihan.

4.4.3 Uji Coba LSTM tanpa *Word Embedding FastText* atau GloVe

Setelah tahap *preprocessing* selesai, langkah selanjutnya adalah memproses dan membangun matriks *embedding* menggunakan data dari model *FastText* yang kemudian digunakan dalam representasi kata pada model pembelajaran mesin.

```
Epoch 1/15
Hidden State (h): [[0.00575664267 0.033601407 0.000154575304 ... 0.010584211 0.0247126017 8.55772814e-05]
 [0.00462205661 0.0357808657 -0.00105777476 ... 0.0096985409 0.0234804302 0.00160393957]
 [0.0292493645 0.0122627793 -0.00536202546 ... 0.00381456059 0.00134059554 0.00963432901]
 ...
 [0.0146128386 -0.00989758875 -0.000592405559 ... 0.00547190104 0.00117450487 -0.00167138816]
 [0.00515952753 0.0336230062 -0.00129431766 ... 0.0101867402 0.0248303451 3.48718022e-05]
 [0.0047527086 0.0353089273 -0.000926487905 ... 0.00962008629 0.0239530019 0.00152827136]]
Cell State (c): [[0.0116568068 0.0666867644 0.000316373946 ... 0.0213809218 0.0501369387 0.000172490109]
 [0.00936566666 0.0709870383 -0.00216596085 ... 0.0196124092 0.0475937799 0.0032310877]
 [0.059242256 0.0242458712 -0.0108597595 ... 0.00763987657 0.00267678313 0.0197785981]
 ...
 [0.0293418802 -0.0199142583 -0.00118157221 ... 0.0109021524 0.00233900128 -0.00333844917]
 [0.0104474891 0.0667322129 -0.00264942646 ... 0.0205802955 0.0503819659 7.02670659e-05]
 [0.0096289767 0.0700587779 -0.00189704308 ... 0.0194500498 0.0485634506 0.00307918945]]
  1/262 ━━━━━━━━━━━━━━━━━━━━ 21:13 5s/step - accuracy: 0.2344 - loss: 0.7053
Hidden State (h): [[0.00651009707 0.0265388153 0.00467630336 ... 0.00801250339 0.0222756192 -0.00513677252]
 [0.00248511531 0.0306722559 0.00692288438 ... 0.00574976485 0.0215962045 -0.00116088556]
 [0.000739094394 0.0323404074 0.00681483 ... 0.00565517321 0.0204608031 2.07323847e-05]
 ...
 [0.0442326143 0.0278267916 -0.027194472 ... -0.0110540576 0.0189355519 0.0059380969]
 [0.0258358065 0.0105905887 -0.0137290582 ... -0.00321377325 0.0105455462 -0.000817034626]
 [0.0367995948 0.0208541341 -0.0124523565 ... -0.0112756714 0.00625625579 0.0029682389]]
Cell State (c): [[0.0131584527 0.0526818708 0.00955629908 ... 0.0161654428 0.0452964604 -0.0103406264]
 [0.00502655236 0.0608763322 0.0141511941 ... 0.0116230519 0.0438464619 -0.00233486923]
 [0.00149583817 0.0641702861 0.0139345676 ... 0.0114408908 0.0415078923 4.1666e-05]
 ...
 [0.0877167583 0.0553765558 -0.0531143844 ... -0.02183038 0.0386485755 0.0117905196]
 [0.0508907177 0.0210982934 -0.0266400371 ... -0.0063294787 0.0215083621 -0.00162252504]
 [0.0743749142 0.0414738357 -0.0245246105 ... -0.0226025395 0.0123989722 0.00595826423]]
  2/262 ━━━━━━━━━━━━━━━━━━━━ 27s 104ms/step - accuracy: 0.3203 - loss: 0.7026
```

Gambar 4. 3 *Output* Pelatihan Model LSTM

Output yang terlihat dalam gambar 4.3 menunjukkan hasil pelatihan model LSTM selama beberapa *epoch* pertama, di mana informasi yang tercetak meliputi status tersembunyi (*hidden state*) dan status sel (*cell state*) dari unit LSTM serta metrik evaluasi seperti akurasi dan *loss*.

Pada setiap *epoch* pelatihan, dalam hal ini *epoch* pertama, informasi tentang *hidden state* (h) dan *cell state* (c) dari unit LSTM dicetak. *Hidden state* adalah vektor yang menggambarkan keadaan internal LSTM yang berfungsi untuk

memproses dan menyimpan informasi tentang input yang telah diproses sejauh ini. Ini berperan dalam membawa informasi yang relevan ke langkah-langkah berikutnya dalam urutan. Sementara itu, *cell state* berfungsi untuk menyimpan informasi jangka panjang dan memelihara dependensi temporal atau hubungan yang jauh dalam urutan data. Kedua status ini memberikan pemahaman yang lebih dalam tentang bagaimana model mengelola informasi dari data urutan selama pelatihan.

Setiap baris yang mencetak status tersembunyi dan status sel menunjukkan langkah pelatihan tertentu, seperti yang terlihat pada 1/262, yang berarti model sedang berada pada langkah pertama dari 262 langkah dalam satu *epoch*. *Output* ini berisi vektor yang sangat panjang, yang mencakup nilai-nilai untuk setiap elemen dalam *hidden state* dan *cell state*, di mana setiap nilai dalam vektor ini mewakili informasi yang diambil atau disimpan oleh LSTM setelah memproses data urutan.

Di bagian bawah output, terdapat metrik evaluasi yang sangat penting untuk mengukur kinerja model, yaitu *accuracy* dan *loss*. *Accuracy* mengukur seberapa sering prediksi model sesuai dengan label yang benar dari data pelatihan atau validasi. Pada *epoch* pertama, *accuracy* tercatat sekitar 0.2344, yang berarti bahwa model hanya benar dalam memprediksi kelas sekitar 23% dari waktu. Ini menunjukkan bahwa model belum belajar dengan baik pada awal pelatihan. Sementara itu, *loss* yang tercatat sebesar 0.7053 menunjukkan besar kesalahan model dalam memprediksi hasil yang benar. *Loss* merupakan ukuran seberapa jauh prediksi model dari nilai yang diharapkan, dan nilai yang lebih rendah biasanya menunjukkan performa model yang lebih baik.

Kecepatan pelatihan juga dicatat dalam *step per second*, yang menggambarkan berapa lama waktu yang diperlukan untuk memproses setiap langkah pelatihan. Misalnya, pada 21:13 5s/*step*, model memerlukan waktu sekitar 5 detik untuk memproses setiap langkah pelatihan. Hal ini membantu kita memahami seberapa cepat model berjalan dan apakah pelatihan dapat dilakukan dalam waktu yang wajar, mengingat ukuran data dan kompleksitas model.

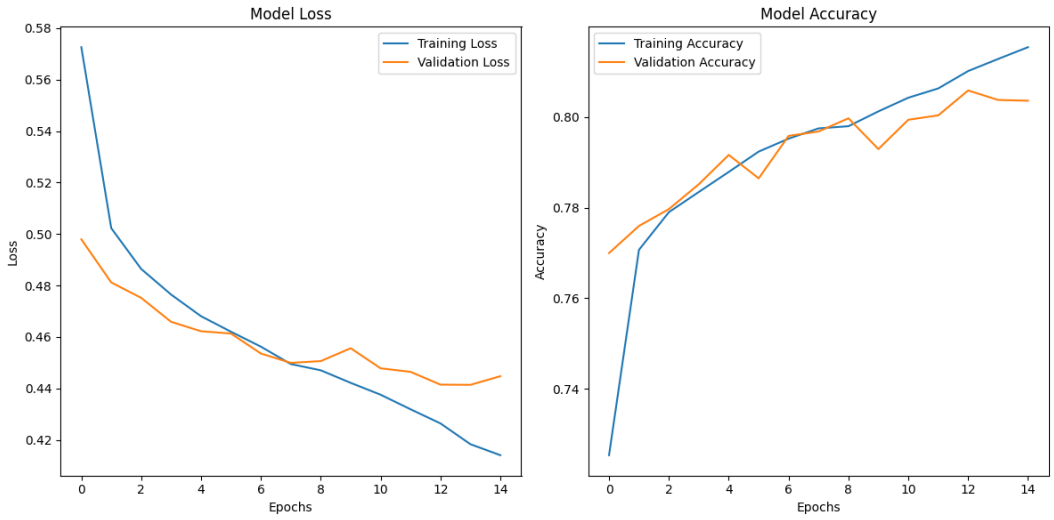
Secara keseluruhan, output ini memberikan wawasan penting mengenai bagaimana unit LSTM mengelola dan memproses informasi sekuensial, serta seberapa baik model belajar pada data pelatihan. Melalui *hidden state* dan *cell state*, kita dapat melihat bagaimana model menyimpan informasi jangka pendek dan panjang. Selain itu, metrik *accuracy* dan *loss* memberi gambaran tentang kualitas model saat belajar, dengan tujuan untuk memperbaiki akurasi dan mengurangi loss seiring berjalannya pelatihan.

4.5 Hasil Loss dan Akurasi Uji Coba

Pada bagian ini, hasil ringkasan yang dihasilkan dari model prediksi dibandingkan dengan ringkasan yang dibuat oleh manusia. Untuk mengukur kualitas dari ringkasan yang dihasilkan, digunakan metrik ROUGE, yang mencakup skor ROUGE-1, ROUGE-2, dan ROUGE-L. Skor ROUGE-1 mengukur kesamaan antara unigram (kata-kata tunggal) dalam ringkasan prediksi dan ringkasan manusia, ROUGE-2 mengukur kesamaan bigram (pasangan kata), dan ROUGE-L mengukur kesamaan berdasarkan urutan kata yang lebih panjang.

4.5.1 Hasil Loss dan Akurasi LSTM dengan *Word Embedding* GloVe

Hasil evaluasi akurasi model menunjukkan kinerja yang sangat baik dalam mempelajari data pelatihan dan mengaplikasikan pengetahuannya pada data baru. *Test accuracy* sebesar 80,38% menunjukkan bahwa model memiliki kemampuan generalisasi yang baik, mampu menerapkan pengetahuannya pada data yang belum pernah dilihat sebelumnya. Perbedaan kecil antara akurasi pelatihan dan pengujian menunjukkan bahwa model tidak mengalami *overfitting* dan dapat bekerja dengan baik pada data yang baru.



Gambar 4. 4 Grafik Akurasi Model LSTM dengan GloVe

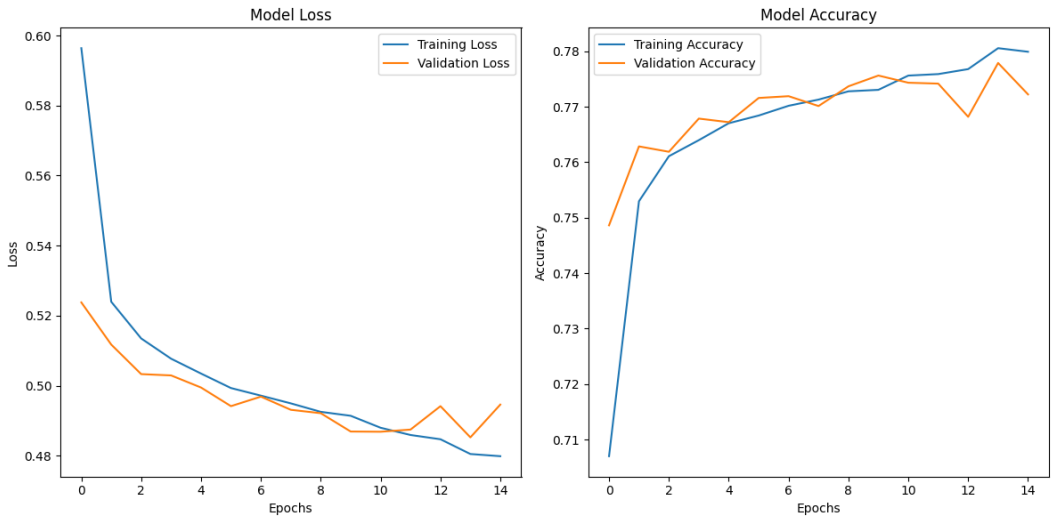
Grafik 4.4 menggambarkan perkembangan *accuracy* dan *loss* pada tiga kategori berbeda: *Training, Test*, dan *Validation* selama proses pelatihan model. Grafik pertama (*model accuracy*) menunjukkan bahwa akurasi pelatihan terus meningkat seiring dengan bertambahnya epoch, sementara akurasi pengujian dan validasi menunjukkan peningkatan yang lebih stabil namun sedikit lebih rendah dibandingkan dengan pelatihan. Hal ini menunjukkan bahwa meskipun model lebih

baik dalam mempelajari data pelatihan, ia masih dapat menggeneralisasi dengan baik pada data yang tidak terlihat sebelumnya.

Grafik kedua (*model loss*) menunjukkan bahwa *training loss* terus menurun dengan stabil sepanjang pelatihan, yang menunjukkan bahwa model semakin efektif dalam meminimalkan kesalahan pada data pelatihan. *Validation loss* dan *test loss* juga mengalami penurunan, meskipun *validation loss* sedikit lebih tinggi daripada *training loss*, yang normal karena data validasi tidak digunakan dalam pelatihan langsung. Penurunan *loss* yang konsisten pada *test* dan *validation loss* menunjukkan bahwa model tidak hanya berhasil belajar dari data pelatihan, tetapi juga dapat menggeneralisasi dengan baik pada data baru. Secara keseluruhan, grafik ini mencerminkan bahwa model memiliki kinerja yang baik dalam mempelajari pola dari data pelatihan dan mengaplikasikannya pada data pengujian serta data validasi tanpa menunjukkan tanda-tanda *overfitting*.

4.5.2 Hasil Loss dan Akurasi LSTM dengan *Word Embedding FastText*

Hasil evaluasi akurasi model menunjukkan kinerja yang sangat baik dalam mempelajari data pelatihan dan mengaplikasikan pengetahuannya pada data baru. *Test accuracy* sebesar 77,79% menunjukkan bahwa model memiliki kemampuan generalisasi yang baik, mampu menerapkan pengetahuannya pada data yang belum pernah dilihat sebelumnya. Perbedaan kecil antara akurasi pelatihan dan pengujian menunjukkan bahwa model tidak mengalami *overfitting* dan dapat bekerja dengan baik pada data yang baru.



Gambar 4. 5 Grafik Akurasi Model LSTM dengan *FastText*

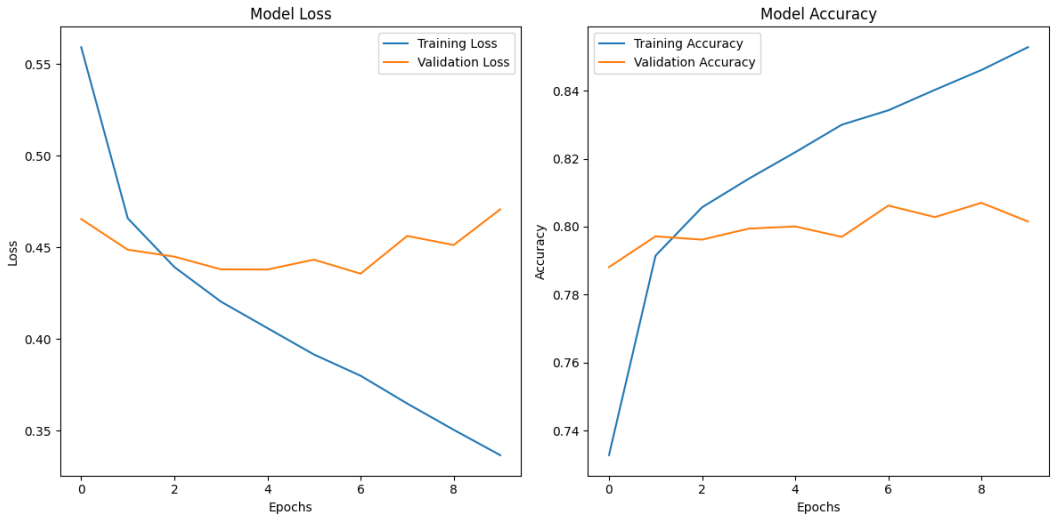
Grafik 4.5 menggambarkan perkembangan accuracy dan loss pada tiga kategori berbeda: *Training, Test,* dan *Validation* selama proses pelatihan model. Grafik pertama (*model accuracy*) menunjukkan bahwa akurasi pelatihan terus meningkat seiring dengan bertambahnya epoch, sementara akurasi pengujian dan validasi menunjukkan peningkatan yang lebih stabil namun sedikit lebih rendah dibandingkan dengan pelatihan. Hal ini menunjukkan bahwa meskipun model lebih baik dalam mempelajari data pelatihan, ia masih dapat menggeneralisasi dengan baik pada data yang tidak terlihat sebelumnya.

Grafik kedua (*model loss*) menunjukkan bahwa *training loss* terus menurun dengan stabil sepanjang pelatihan, yang menunjukkan bahwa model semakin efektif dalam meminimalkan kesalahan pada data pelatihan. *Validation loss* dan *test loss* juga mengalami penurunan, meskipun *validation loss* sedikit lebih tinggi daripada training loss, yang normal karena data validasi tidak digunakan dalam pelatihan langsung. Penurunan *loss* yang konsisten pada *test* dan *validation loss*

menunjukkan bahwa model tidak hanya berhasil belajar dari data pelatihan, tetapi juga dapat menggeneralisasi dengan baik pada data baru. Secara keseluruhan, grafik ini mencerminkan bahwa model memiliki kinerja yang baik dalam mempelajari pola dari data pelatihan dan mengaplikasikannya pada data pengujian serta data validasi tanpa menunjukkan tanda-tanda *overfitting*.

4.5.3 Hasil Loss dan Akurasi LSTM tanpa *Word Embedding*

Hasil evaluasi akurasi model menunjukkan kinerja yang sangat baik dalam mempelajari data pelatihan dan mengaplikasikan pengetahuannya pada data baru. *Test accuracy* sebesar 80,62% menunjukkan bahwa model memiliki kemampuan generalisasi yang baik, mampu menerapkan pengetahuannya pada data yang belum pernah dilihat sebelumnya. Perbedaan kecil antara akurasi pelatihan dan pengujian menunjukkan bahwa model tidak mengalami *overfitting* dan dapat bekerja dengan baik pada data yang baru.



Gambar 4. 6 Grafik Akurasi Model LSTM tanpa *Word Embedding*

Grafik 4.6 menggambarkan perkembangan accuracy dan loss pada tiga kategori berbeda: *Training, Test, dan Validation* selama proses pelatihan model. Grafik pertama (*model accuracy*) menunjukkan bahwa akurasi pelatihan terus meningkat seiring dengan bertambahnya *epoch*, sementara akurasi pengujian dan validasi menunjukkan peningkatan yang lebih stabil namun sedikit lebih rendah dibandingkan dengan pelatihan. Hal ini menunjukkan bahwa meskipun model lebih baik dalam mempelajari data pelatihan, ia masih dapat menggeneralisasi dengan baik pada data yang tidak terlihat sebelumnya.

Grafik kedua (*model loss)* menunjukkan bahwa *training loss* terus menurun dengan stabil sepanjang pelatihan, yang menunjukkan bahwa model semakin efektif dalam meminimalkan kesalahan pada data pelatihan. *Validation loss* dan *test loss* juga mengalami penurunan, meskipun validation loss sedikit lebih tinggi daripada *training loss*, yang normal karena data validasi tidak digunakan dalam pelatihan langsung. Penurunan *loss* yang konsisten pada *test* dan *validation loss* menunjukkan bahwa model tidak hanya berhasil belajar dari data pelatihan, tetapi juga dapat menggeneralisasi dengan baik pada data baru. Secara keseluruhan, grafik ini mencerminkan bahwa model memiliki kinerja yang baik dalam mempelajari pola dari data pelatihan dan mengaplikasikannya pada data pengujian serta data validasi tanpa menunjukkan tanda-tanda *overfitting*.

4.6 Hasil Ringkasan dan Skor ROUGE Uji Coba

Pada bagian ini, hasil ringkasan yang dihasilkan dari model prediksi dibandingkan dengan ringkasan yang dibuat oleh manusia. Untuk mengukur kualitas dari ringkasan yang dihasilkan, digunakan metrik ROUGE, yang mencakup

skor ROUGE-1, ROUGE-2, dan ROUGE-L. Skor ROUGE-1 mengukur kesamaan antara unigram (kata-kata tunggal) dalam ringkasan prediksi dan ringkasan manusia, ROUGE-2 mengukur kesamaan bigram (pasangan kata), dan ROUGE-L mengukur kesamaan berdasarkan urutan kata yang lebih panjang.

4.6.1 Hasil Ringkasan dan Skor ROUGE *Word Embedding* GloVe

Setelah menerapkan LSTM dengan *word embedding* menggunakan GloVe, langkah berikutnya adalah menampilkan hasil evaluasi model menggunakan skor ROUGE. Skor ROUGE ini digunakan untuk mengukur kinerja model dalam menghasilkan teks yang relevan dan sesuai dengan referensi, memberikan gambaran mengenai kualitas prediksi yang dihasilkan.

Tabel 4. 8 Hasil Ringkasan dan Skor ROUGE LSTM dengan GloVe

NO	*Human Summary*	Umpan Balik Pelanggan	*Predicted Summary*	Rouge-1	Rouge-2	Rouge-L
1	The best in all that matters	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. The only one in crypto that offers you to actually call a phone number and get to someone with your questions just like that. Highly recommend! I have been a customer for more than a year and I have only good things to say about Celsius network.	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. I have been a customer for more than a year and I have only good things to say about Celsius network.	0.27027	0.195122	0.27027
2	Celsius Network ROCKS!	If you are looking for the best #HomeForCrypto and where to earn steady yield then there is no better place than Celsius Network. The app is easy to use and understand with a	If you are looking for the best #HomeForCrypto and where to earn steady yield then there is no better place than Celsius Network. The app is easy	0.086957	0.037037	0.086957

NO	*Human Summary*	Umpan Balik Pelanggan	*Predicted Summary*	Rouge-1	Rouge-2	Rouge-L
		recent update as well. The company is led by a great CEO who is fully engaged with the Celsius community! The mission of Celsius is to do good then do well and they live up to that very much. Rewards compound and pay out weekly!	to use and understand with a recent update as well. The company is led by a great CEO who is fully engaged with the Celsius community!			
3	I despise it so much	I despise it so much. Transferring to other wallets is difficult, especially because you cannot swap your coins, and there is a long waiting period when you do send to someone. Verification within 24 hours. When compared to Crypto.com, I find it extremely difficult to gain access to your funds. I'm never going to use it again. My biggest blunder!	Transferring to other wallets is difficult, especially because you cannot swap your coins, and there is a long waiting period when you do send to someone. When compared to Crypto.com, I find it extremely difficult to gain access to your funds. I'm never going to use it again.	0.088889	0	0.088889

Tabel 4.8 memperlihatkan hasil evaluasi ringkasan model prediksi dibandingkan dengan ringkasan manusia melalui skor ROUGE. Skor yang lebih tinggi, seperti pada contoh nomor 1 (ROUGE-1, ROUGE-2, dan ROUGE-L masing-masing sebesar 0.270, 0.195, dan 0.270), menunjukkan tingkat kesamaan yang lebih baik. Sementara itu, skor yang lebih rendah, seperti pada contoh nomor 2 (ROUGE-1 sebesar 0.087, ROUGE-2 sebesar 0.037, dan ROUGE-L sebesar 0.087), menunjukkan adanya perbedaan yang lebih signifikan antara ringkasan

model dan ringkasan manusia. Hal ini menyoroti tantangan dalam menghasilkan ringkasan yang akurat dan relevan, terutama dalam konteks umpan balik pelanggan.

4.6.2 Hasil Ringkasan dan Skor ROUGE *Word Embedding FastText*

Setelah menerapkan LSTM dengan *word embedding* menggunakan *FastText* untuk menghasilkan representasi kata yang lebih efektif, langkah berikutnya adalah menampilkan hasil evaluasi model melalui skor ROUGE. Skor ROUGE ini digunakan untuk menilai kinerja model dalam menghasilkan teks yang relevan dan sesuai dengan referensi, memberikan gambaran mengenai kualitas prediksi yang dihasilkan

Tabel 4. 9 Hasil Ringkasan dan Skor ROUGE LSTM dengan *FastText*

NO	*Human Summary*	Umpan Balik Pelanggan	*Predicted Summary*	Rouge-1	Rouge-2	Rouge-L
1	The best in all that matters	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. The only one in crypto that offers you to actually call a phone number and get to someone with your questions just like that. Highly recommend! I have been a customer for more than a year and I have only good things to say about Celsius network.	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. Highly recommend! I have been a customer for more than a year and I have only good things to say about Celsius network.	0.25641	0.186047	0.25641
2	Celsius Network ROCKS!	If you are looking for the best #HomeForCrypto and where to earn steady yield then there is no better place than Celsius Network. The app is easy to use and	The app is easy to use and understand with a recent update as well. The company is led by a great CEO who is fully engaged with the Celsius	0.051282	0	0.051282

NO	*Human Summary*	Umpan Balik Pelanggan	*Predicted Summary*	Rouge-1	Rouge-2	Rouge-L
		understand with a recent update as well. The company is led by a great CEO who is fully engaged with the Celsius community! The mission of Celsius is to do good then do well and they live up to that very much. Rewards compound and pay out weekly!	community! The mission of Celsius is to do good then do well and they live up to that very much.			
3	I despise it so much	I despise it so much. Transferring to other wallets is difficult, especially because you cannot swap your coins, and there is a long waiting period when you do send to someone. Verification within 24 hours. When compared to Crypto.com, I find it extremely difficult to gain access to your funds. I'm never going to use it again. My biggest blunder!	Transferring to other wallets is difficult, especially because you cannot swap your coins, and there is a long waiting period when you do send to someone. When compared to Crypto.com, I find it extremely difficult to gain access to your funds. I'm never going to use it again. My biggest blunder!	0.083333	0	0.083333

Tabel 4.9 memperlihatkan hasil evaluasi ringkasan model prediksi dibandingkan dengan ringkasan manusia melalui skor ROUGE. Skor yang lebih tinggi, seperti pada contoh nomor 1 (ROUGE-1, ROUGE-2, dan ROUGE-L masing-masing sebesar 0.256, 0.186, dan 0.256), menunjukkan tingkat kesamaan yang lebih baik. Sebaliknya, skor yang lebih rendah, seperti pada contoh nomor 2 (ROUGE-1 sebesar 0.051, ROUGE-2 sebesar 0, dan ROUGE-L sebesar 0.051),

menunjukkan adanya perbedaan yang signifikan antara ringkasan model dan ringkasan manusia. Hal ini menggambarkan tantangan dalam menghasilkan ringkasan yang akurat dan relevan dalam konteks umpan balik pelanggan.

4.6.3 Hasil Ringkasan dan Skor ROUGE tanpa *Word Embedding*

Setelah menerapkan model LSTM dengan representasi kata tanpa *word embedding* GloVe ataupun *FastText*, langkah selanjutnya adalah menampilkan hasil evaluasi model menggunakan skor ROUGE. Skor ROUGE ini digunakan untuk mengukur kinerja model dalam menghasilkan teks yang relevan dan sesuai dengan referensi, memberikan gambaran mengenai kualitas hasil prediksi yang diperoleh

Tabel 4. 10 Hasil Ringkasan dan Skor ROUGE LSTM Tanpa *Word Embedding*

NO	***Human Summary***	***Customer Feedback***	***Predicted Summary***	**Rouge-1**	**Rouge-2**	**Rouge-L**
1	The best in all that matters	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. The only one in crypto that offers you to actually call a phone number and get to someone with your questions just like that. Highly recommend! I have been a customer for more than a year and I have only good things to say about Celsius network.	The best in all that matters	0.27027	0.195122	0.27027
2	Celsius Network ROCKS!	If you are looking for the best #HomeForCrypto and where to earn steady yield then there is no better place than Celsius Network. The app is easy to use and understand with a recent update as well. The company is led by a great CEO who is fully engaged with the Celsius community! The mission of Celsius is to do good then do well and they live up to that very much.	Celsius Network ROCKS!	0.108108	0.052632	0.108108

NO	*Human Summary*	*Customer Feedback*	*Predicted Summary*	Rouge-1	Rouge-2	Rouge-L
		Rewards compound and pay out weekly!				
3	I despise it so much	I despise it so much. Transferring to other wallets is difficult, especially because you cannot swap your coins, and there is a long waiting period when you do send to someone. Verification within 24 hours. When compared to Crypto.com, I find it extremely difficult to gain access to your funds. I'm never going to use it again. My biggest blunder!	I despise it so much	1	1	1

Tabel 4.10 memperlihatkan hasil evaluasi ringkasan model prediksi dibandingkan dengan ringkasan manusia melalui skor ROUGE. Skor tinggi, seperti pada contoh nomor 3 (ROUGE-1, ROUGE-2, dan ROUGE-L mencapai 1), menunjukkan kesamaan sempurna. Sebaliknya, skor rendah, seperti pada nomor 2 (ROUGE-1 sebesar 0.108108), mencerminkan perbedaan signifikan. Hal ini menyoroti tantangan menghasilkan ringkasan yang akurat dan relevan dalam konteks umpan balik pelanggan.

4.7 Pembahasan Hasil

Pada bagian ini, hasil dari implementasi model LSTM yang diintegrasikan dengan teknik *word embedding* seperti GloVe dan *FastText*, serta tanpa menggunakan *word embedding*, akan dianalisis dan dibahas secara mendalam. Pembahasan ini akan menyoroti kelebihan, kekurangan, serta aspek-aspek yang perlu diperbaiki dari setiap pendekatan yang digunakan dalam penelitian ini.

Tabel 4.11 merupakan tabel perbedaan hasil evaluasi model LSTM dengan GloVe, *FastText*, dan tanpa *word embedding* berdasarkan metrik ROUGE (ROUGE-1, ROUGE-2, dan ROUGE-L).

Tabel 4. 11 Hasil Rata-Rata ROUGE dan Akurasi Model LSTM

Metrik	LSTM dengan GloVe	LSTM dengan *FastText*	LSTM tanpa *Word Embedding*
ROUGE-1 score	0.2620	0.2503	0.3085
ROUGE-2 score	0.1907	0.1796	0.2282
ROUGE-L score	0.2604	0.2486	0.3064
Test Accuracy	80.38%	77.79%	80.62%

4.7.1 Pembahasan Model LSTM dengan GloVe

Keunggulan utama dari penggunaan GloVe sebagai teknik *word embedding* adalah kemampuannya untuk menangkap makna semantik kata secara lebih mendalam. GloVe menghasilkan representasi kata dalam bentuk vektor yang mempertimbangkan konteks kata-kata di sekitarnya dalam korpus besar, sehingga model dapat mempelajari hubungan antara kata-kata dalam kalimat dengan lebih baik. Hasil akurasi yang mencapai 80,38% pada data uji menunjukkan bahwa model LSTM dengan GloVe memiliki kemampuan yang baik dalam menggeneralisasi pengetahuannya dari data pelatihan ke data yang belum pernah dilihat sebelumnya. Meskipun model ini lebih baik dalam memahami konteks kalimat secara keseluruhan, tetap ada keterbatasan dalam hal menangkap hubungan antar kata yang lebih kompleks, seperti bigram, yang tercermin dalam skor ROUGE-2 yang rendah (0.1907). Ini menunjukkan bahwa meskipun GloVe efektif dalam menangkap hubungan semantik kata, model ini masih kesulitan dalam memahami konteks yang lebih rumit yang melibatkan urutan kata dalam kalimat yang lebih panjang. Selain itu, meskipun model berhasil menangkap banyak informasi penting

dalam teks, skor ROUGE-1 (0.2620) dan ROUGE-L (0.2604) mengindikasikan bahwa ada potensi untuk meningkatkan koherensi dan cakupan elemen-elemen penting dalam ringkasan. Hal ini mungkin dapat dicapai dengan penerapan teknik yang lebih maju dalam pemrosesan urutan kalimat atau penggunaan model berbasis transformer yang lebih kuat dalam menangkap hubungan konteks antar kalimat.

4.7.2 Pembahasan Model LSTM dengan *FastText*

Penggunaan *FastText* dalam model LSTM memberikan keuntungan dalam hal kemampuan model untuk menangani kata-kata yang tidak ada dalam kosakata pelatihan. *FastText* menggunakan sub-kata untuk menghasilkan representasi kata yang lebih fleksibel, yang memungkinkan model untuk memahami kata-kata baru yang belum pernah ditemukan sebelumnya. Ini sangat berguna dalam aplikasi seperti peringkasan teks ulasan pelanggan, di mana kata-kata yang tidak baku atau baru sering kali muncul. Namun, meskipun *FastText* memberikan kemampuan lebih dalam menangkap makna kata yang tidak dikenal, hasil evaluasi dengan *FastText* menunjukkan skor ROUGE yang lebih rendah dibandingkan dengan GloVe. Skor ROUGE-1 (0.2503), ROUGE-2 (0.1796), dan ROUGE-L (0.2486) menunjukkan bahwa model ini masih kesulitan dalam menangkap hubungan antar kata dalam konteks yang lebih besar, seperti bigram, serta dalam menjaga koherensi antar kalimat. Hal ini mungkin disebabkan oleh cara *FastText* menggabungkan informasi sub-kata, yang meskipun lebih fleksibel, terkadang tidak cukup kuat untuk menangkap hubungan konteks yang lebih dalam dalam teks. Oleh karena itu, meskipun *FastText* membantu model dalam memahami kata-kata baru dan variasi

kata, perbaikan dalam pengolahan hubungan antar kata dan kalimat diperlukan agar model dapat menghasilkan ringkasan yang lebih koheren dan akurat.

4.7.3 Pembahasan Model LSTM Tanpa *Word Embedding*

Penerapan model LSTM tanpa menggunakan teknik *word embedding* seperti GloVe atau *FastText* lebih sederhana namun tetap efektif dalam menghasilkan ringkasan teks. Model ini bekerja dengan representasi kata yang lebih sederhana, seperti one-hot encoding atau indeks kata, yang tidak mempertimbangkan hubungan semantik antar kata secara mendalam. Meskipun demikian, hasil evaluasi menunjukkan bahwa model ini masih dapat menghasilkan ringkasan yang cukup baik, dengan akurasi uji mencapai 80,62%. Keunggulan utama dari pendekatan ini adalah kesederhanaannya, yang memungkinkan model untuk bekerja dengan lebih efisien tanpa memerlukan tahapan tambahan dalam representasi kata. Namun, kelemahan dari pendekatan ini terletak pada keterbatasan dalam menangkap makna semantik dan hubungan antar kata yang lebih kompleks. Model ini kesulitan dalam memahami konteks lebih dalam, yang tercermin dalam skor ROUGE-2 (0.2282) dan ROUGE-L (0.3064) yang lebih rendah dibandingkan dengan model berbasis *word embedding*. Meskipun model ini dapat menghasilkan ringkasan yang informatif, penggunaan representasi kata yang lebih sederhana mengurangi kemampuannya untuk memahami makna yang lebih halus dalam teks dan menghubungkan kalimat dengan cara yang lebih koheren.

4.7.4 Pembahasan Tabel Perbedaan Hasil

Pada tabel 4.13 yang menunjukkan hasil evaluasi menggunakan metrik ROUGE untuk ketiga pendekatan model LSTM, yaitu dengan GloVe, *FastText*, dan

tanpa menggunakan *word embedding*, terdapat perbedaan yang signifikan dalam skor yang diperoleh. Tabel ini menyajikan nilai rata-rata untuk tiga metrik utama dalam ROUGE, yaitu ROUGE-1, ROUGE-2, dan ROUGE-L, yang mengukur kesamaan antara ringkasan yang dihasilkan oleh model dengan ringkasan referensi yang telah disiapkan secara manual.

Pertama, pada ROUGE-1, yang mengukur kesamaan unigram, model LSTM dengan GloVe memperoleh skor 0.2620, yang menunjukkan bahwa model ini mampu menangkap sekitar 26% dari informasi kunci dalam bentuk kata tunggal. Meskipun hasil ini cukup baik, ada ruang untuk peningkatan dalam memperluas cakupan elemen penting dalam ringkasan. Di sisi lain, model LSTM dengan *FastText* memperoleh skor 0.2503, sedikit lebih rendah dibandingkan dengan GloVe. Hal ini bisa terjadi karena meskipun *FastText* efektif dalam menangani kata-kata yang tidak ada dalam kosakata pelatihan, model ini masih memiliki kesulitan dalam menangkap kata-kata yang lebih umum yang telah terlihat dalam data pelatihan. Sementara itu, model LSTM tanpa *word embedding* memperoleh skor tertinggi untuk ROUGE-1, yaitu 0.3085. Ini menunjukkan bahwa meskipun tidak menggunakan representasi semantik kata yang mendalam, model ini lebih baik dalam menangkap informasi penting dalam teks yang lebih panjang. Salah satu alasan skor ini lebih tinggi bisa jadi karena model ini bekerja dengan struktur kalimat yang lebih sederhana, sehingga bisa lebih mudah menangkap elemen-elemen kunci tanpa terganggu oleh kompleksitas hubungan semantik antar kata.

Pada ROUGE-2, yang mengukur kesamaan bigram, model LSTM dengan GloVe memperoleh skor 0.1907, yang sedikit lebih rendah dibandingkan dengan

ROUGE-1. Ini menunjukkan bahwa meskipun model ini mampu menangkap banyak informasi penting dalam bentuk unigram, ia masih kesulitan dalam menangkap hubungan antar kata berturut-turut (bigram), yang penting untuk memahami konteks dalam kalimat yang lebih kompleks. LSTM dengan *FastText* memperoleh skor 0.1796 pada ROUGE-2, yang lebih rendah dibandingkan dengan GloVe. Hal ini menunjukkan bahwa meskipun *FastText* sangat berguna untuk kata-kata baru, model ini tetap memerlukan perbaikan dalam menangkap konteks hubungan antar kata yang lebih terstruktur dan dinamis. LSTM tanpa *word embedding*, di sisi lain, memperoleh skor 0.2282 untuk ROUGE-2, yang sedikit lebih tinggi dari kedua model berbasis *word embedding*. Meskipun model ini tidak menggunakan representasi semantik yang mendalam, model ini mampu menangkap hubungan antar kata dalam bigram dengan lebih baik, mungkin karena struktur kalimat yang lebih sederhana dan pendekatan langsung dalam memproses urutan kata.

Untuk ROUGE-L, yang mengukur kesamaan berdasarkan urutan kata yang lebih panjang, model LSTM dengan GloVe memperoleh skor 0.2604, yang menunjukkan bahwa meskipun model ini cukup baik dalam menangkap urutan kata yang lebih panjang dan hubungan antar kalimat, masih ada ruang untuk meningkatkan koherensi dan alur dalam ringkasan yang dihasilkan. LSTM dengan *FastText* memperoleh skor 0.2486, sedikit lebih rendah dibandingkan GloVe, yang mengindikasikan bahwa meskipun *FastText* dapat menangani kata-kata baru dengan lebih baik, model ini masih kesulitan dalam menjaga koherensi dan urutan kata yang lebih panjang dalam ringkasan. Hal ini mungkin disebabkan oleh

ketidakmampuan *FastText* dalam menangkap hubungan antara kalimat yang lebih kompleks, meskipun representasi kata yang dihasilkan lebih fleksibel dan dapat menangani kata-kata yang tidak ada dalam kosakata pelatihan. Sementara itu, LSTM tanpa *word embedding* memperoleh skor 0.3064, yang merupakan skor tertinggi untuk ROUGE-L di antara ketiga pendekatan. Hal ini menunjukkan bahwa meskipun model tanpa *word embedding* tidak memiliki pemahaman semantik kata yang mendalam, model ini lebih baik dalam menjaga urutan kata dan hubungan antar kalimat, menghasilkan ringkasan yang lebih koheren. Meskipun demikian, model ini masih memiliki keterbatasan dalam memahami makna kata secara lebih mendalam, yang memengaruhi kualitas keseluruhan dari ringkasan yang dihasilkan.

Secara keseluruhan, hasil dari tabel perbandingan ini menunjukkan bahwa LSTM tanpa *word embedding* memiliki keunggulan dalam menjaga koherensi dan menangkap elemen-elemen kunci dari teks, namun memiliki keterbatasan dalam memahami hubungan semantik yang lebih kompleks. Sementara itu, LSTM dengan GloVe menunjukkan kemampuan yang lebih baik dalam memahami konteks semantik dan menangkap informasi penting dalam teks, meskipun masih ada ruang untuk perbaikan dalam menangkap hubungan antar kata dan kalimat. LSTM dengan *FastText*, meskipun lebih unggul dalam menangani kata-kata yang tidak ada dalam kosakata, masih perlu peningkatan dalam hal memahami hubungan antar kata dan kalimat yang lebih rumit. Perbedaan skor ini memberikan wawasan tentang bagaimana masing-masing teknik berkontribusi pada kinerja model dalam menghasilkan ringkasan teks, serta area-area yang perlu diperbaiki untuk meningkatkan kualitas keseluruhan dari ringkasan yang dihasilkan.

4.7.5 Kesimpulan Hasil

Dari hasil evaluasi yang disajikan dalam tabel, kita dapat menarik beberapa kesimpulan. LSTM tanpa *word embedding* menunjukkan hasil yang lebih baik dalam hal ROUGE-1 dan ROUGE-L, yang mengindikasikan bahwa model ini lebih baik dalam menangkap elemen-elemen kunci dalam teks dan menjaga koherensi ringkasan. Namun, model ini masih memiliki keterbatasan dalam menangkap hubungan antar kata yang lebih kompleks (seperti bigram), yang tercermin dalam skor ROUGE-2 yang lebih rendah.

Sementara itu, LSTM dengan GloVe menunjukkan kinerja yang cukup baik di hampir semua metrik ROUGE, terutama dalam hal pemahaman semantik kata, meskipun masih ada ruang untuk perbaikan dalam menangkap hubungan bigram dan menjaga koherensi antar kalimat. LSTM dengan *FastText*, meskipun memiliki keunggulan dalam menangani kata yang tidak ada dalam kosakata, tidak seefektif GloVe dalam memahami hubungan antar kata dan kalimat yang lebih kompleks, yang tercermin dalam skor ROUGE-2 dan ROUGE-L yang lebih rendah.

Secara keseluruhan, meskipun *word embedding* meningkatkan pemahaman semantik kata, model tanpa *word embedding* lebih unggul dalam menjaga koherensi ringkasan dan menangkap elemen kunci teks.

Penelitian ini mengintegrasikan nilai-nilai Islam yang mendorong menuntut ilmu, berkomunikasi bijaksana, dan menghindari pemborosan. Prinsip-prinsip ini relevan dengan tujuan penelitian, yakni mengolah umpan balik pelanggan secara efisien menggunakan teknologi modern seperti LSTM dengan *word embedding* untuk merangkum teks tanpa mengurangi makna. Penelitian ini tidak hanya

berfokus pada aspek teknis, tetapi juga pada penerapan nilai Islam dalam ilmu pengetahuan dan komunikasi.

Surat An-Nahl (16:125), Allah mengajarkan kita untuk menyampaikan pesan dengan cara yang bijaksana dan sesuai dengan konteks. Ayat ini berbunyi:

ادْعُ إِلَىٰ سَبِيلِ رَبِّكَ بِالْحِكْمَةِ وَالْمَوْعِظَةِ الْحَسَنَةِ وَجَادِلْهُم بِالَّتِي هِيَ أَحْسَنُ

"Serulah (manusia) kepada jalan Tuhanmu dengan hikmah dan pelajaran yang baik, dan bantahlah mereka dengan cara yang lebih baik." (QS. An-Nahl 16:125).

Prinsip ini relevan dalam penelitian ini karena peringkasan teks umpan balik pelanggan harus dilakukan dengan cara yang tidak hanya tepat, tetapi juga bermanfaat. Teknik seperti *Word Embedding* yang digunakan dalam peringkasan teks harus memastikan bahwa informasi yang disampaikan tetap relevan dan mudah dipahami, menjaga esensi pesan dengan cara yang bijaksana dan efektif.

Selain itu, Surat Al-Qamar (54:40) mengingatkan kita tentang pentingnya mengolah informasi dengan tepat dan mengingatnya dengan bijaksana. Ayat ini berbunyi:

وَلَقَدْ يَسَّرْنَا الْقُرْآنَ لِلذِّكْرِ فَهَلْ مِن مُّدَّكِرٍ

"Dan Kami telah jadikan Al-Qur'an itu mudah untuk dipahami dan diingat, maka adakah orang yang mengambil pelajaran?" (QS. Al-Qamar 54:40).

Penerapan prinsip ini dalam penelitian ini mengarah pada penggunaan teknologi seperti LSTM yang mempermudah pengolahan dan pemahaman informasi yang kompleks, menjadikannya lebih mudah diingat dan diolah tanpa kehilangan esensinya. Dalam peringkasan teks, hal ini menunjukkan bagaimana kita dapat menyaring informasi yang ada, memudahkan proses pembelajaran, dan menyampaikan pesan dengan cara yang jelas dan tepat.

Terakhir, dalam Surat Al-A'raf (7:31), Allah mengingatkan kita untuk menghindari segala bentuk pemborosan atau berlebihan dalam hidup, termasuk dalam komunikasi dan penyampaian informasi. Ayat ini berbunyi:

يَا بَنِي آدَمَ خُذُوا زِينَتَكُمْ عِندَ كُلِّ مَسْجِدٍ وَكُلُوا وَاشْرَبُوا وَلَا تُسْرِفُوا إِنَّهُ لَا يُحِبُّ الْمُسْرِفِينَ

"Hai anak-anak Adam, pakailah pakaianmu yang indah di setiap mesjid, dan makan serta minumlah, tetapi jangan berlebihan. Sesungguhnya Allah tidak menyukai orang yang berlebihan." (QS. Al-A'raf 7:31)

Prinsip ini mengajarkan kita untuk tidak berlebihan dalam menyampaikan informasi. Dalam konteks peringkasan teks, ini berarti kita harus menjaga keseimbangan, tidak menambahkan kata-kata yang tidak perlu. Peringkasan harus dilakukan dengan bijak agar informasi yang disajikan tetap padat, jelas, dan tidak mengurangi makna yang penting.

Secara keseluruhan, prinsip-prinsip dalam Islam mengarahkan kita untuk mengolah dan menyampaikan informasi dengan cara yang tepat, efisien, dan bermanfaat, yang juga sejalan dengan tujuan dari penelitian ini dalam menggunakan teknologi untuk merangkum umpan balik pelanggan secara akurat dan efisien.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa model LSTM yang diterapkan pada penelitian peringkasan teks menunjukkan kinerja yang bervariasi bergantung pada teknik representasi kata yang digunakan. Hasil evaluasi menggunakan metrik ROUGE menunjukkan bahwa model LSTM tanpa *word embedding* memiliki kinerja terbaik dengan skor ROUGE-1 (0.3085) dan ROUGE-L (0.3064), yang menunjukkan bahwa model ini lebih baik dalam menangkap elemen-elemen kunci dalam teks dan menjaga koherensi ringkasan. Namun, model ini kesulitan dalam menangkap hubungan antar kata yang lebih kompleks, seperti bigram, yang tercermin dalam skor ROUGE-2 (0.2282) yang lebih rendah. Di sisi lain, model LSTM dengan GloVe menunjukkan kinerja yang baik dalam memahami konteks semantik kata, dengan skor ROUGE-1 (0.2620) dan ROUGE-L (0.2604), tetapi masih mengalami kesulitan dalam menangkap hubungan antar kata yang lebih kompleks (ROUGE-2: 0.1907). Model LSTM dengan *FastText* meskipun lebih unggul dalam menangani kata-kata baru yang tidak ada dalam kosakata, ternyata memiliki skor yang lebih rendah pada metrik ROUGE, terutama pada ROUGE-2 (0.1796) dan ROUGE-L (0.2486), yang menunjukkan bahwa model ini kesulitan dalam menjaga koherensi dan menangkap hubungan antar kalimat yang lebih rumit.

Secara keseluruhan, meskipun LSTM dengan teknik *word embedding* seperti GloVe dan *FastText* menunjukkan peningkatan dalam hal pemahaman

konteks semantik kata, model tanpa *word embedding* lebih unggul dalam menjaga koherensi dan kualitas ringkasan secara keseluruhan. Ini menunjukkan bahwa tantangan utama dalam peringkasan teks masih terletak pada pemahaman hubungan antar kata dan kalimat yang lebih kompleks, yang perlu diperbaiki untuk menghasilkan ringkasan yang lebih akurat dan informatif.

5.2 Saran

Berdasarkan temuan ini, beberapa saran untuk penelitian selanjutnya adalah sebagai berikut. Pertama, meskipun LSTM tanpa *word embedding* menunjukkan kinerja yang lebih baik dalam menjaga koherensi dan mengidentifikasi elemen-elemen penting dalam teks, model berbasis transformer seperti BERT atau GPT dapat dipertimbangkan untuk penelitian lebih lanjut, karena model-model ini lebih kuat dalam memahami konteks kalimat yang lebih kompleks dan menangkap hubungan antar kalimat. Pendekatan berbasis transformer mungkin dapat memperbaiki kelemahan yang ditemukan pada model LSTM, terutama dalam menangkap hubungan antar kata dan kalimat yang lebih rumit.

Kedua, eksperimen dengan kombinasi model *hybrid*, yang menggabungkan LSTM dengan teknik *embedding* dan model transformer, bisa menjadi langkah menarik. Pendekatan ini dapat mengoptimalkan kemampuan model dalam memahami baik konteks semantik yang lebih dalam maupun hubungan antar kalimat yang lebih kompleks. Ketiga, penelitian lebih lanjut dapat mempertimbangkan peningkatan teknik pemrosesan bigram dan trigram untuk memperbaiki kinerja model dalam memahami hubungan antar kata yang lebih terstruktur.

DAFTAR PUSTAKA

Almuzaini, H. A., & Azmi, A. M. (2020). Impact of Stemming and Word Embedding on Deep Learning-Based Arabic Text Categorization. *IEEE Access*, *8*, 127913–127928. https://doi.org/10.1109/ACCESS.2020.3009217

Alshawi, H. (2003). Effective utterance classification with unsupervised phonotactic models. *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - NAACL '03*, *1*, 1–7. https://doi.org/10.3115/1073445.1073446

Bani-Almarjeh, M., & Kurdy, M.-B. (2023). Arabic abstractive text summarization using RNN-based and transformer-based architectures. *Information Processing & Management*, *60*(2), 103227. https://doi.org/10.1016/j.ipm.2022.103227

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, *5*, 135–146. https://doi.org/10.1162/tacl_a_00051

Elsaid, A., Mohammed, A., Ibrahim, L. F., & Sakre, M. M. (2022). A Comprehensive Review of Arabic Text Summarization. *IEEE Access*, *10*, 38012–38030. https://doi.org/10.1109/ACCESS.2022.3163292

Ermawan, B. R., & Cahyono, N. (2025). OPTIMASI METODE KLASIFIKASI MENGGUNAKAN FASTTEXT DAN GRID SEARCH PADA ANALISIS SENTIMEN ULASAN APLIKASI SEABANK. *JIKO (Jurnal Informatika Dan Komputer)*, *9*(1), 226. https://doi.org/10.26798/jiko.v9i1.1523

Gambhir, M., & Gupta, V. (2017). Recent automatic text summarization techniques: A survey. *Artificial Intelligence Review*, *47*(1), 1–66. https://doi.org/10.1007/s10462-016-9475-9

Gambhir, M., & Gupta, V. (2022). Deep learning-based extractive text summarization with word-level attention mechanism. *Multimedia Tools and Applications*, *81*(15), 20829–20852. https://doi.org/10.1007/s11042-022-12729-y

Ghibeche, Y., Sellam, A., Nouri, N. A., Laggoun, S., & Khadroun, N. O. I. (2024). Deep Learning for Automatic Text Summarization. *2024 International Conference on Telecommunications and Intelligent Systems (ICTIS)*, 1–5. https://doi.org/10.1109/ICTIS62692.2024.10894686

Gupta, H., & Patel, M. (2021). Method Of Text Summarization Using Lsa And Sentence Based Topic Modelling With Bert. *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, 511–517. https://doi.org/10.1109/ICAIS50930.2021.9395976

Hallikainen, H., Savimäki, E., & Laukkanen, T. (2020). Fostering B2B sales with customer big data analytics. *Industrial Marketing Management*, *86*, 90–98. https://doi.org/10.1016/j.indmarman.2019.12.005

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Joshi, A., Fidalgo, E., Alegre, E., & Fernández-Robles, L. (2023). DeepSumm: Exploiting topic models and sequence to sequence networks for extractive text summarization. *Expert Systems with Applications*, *211*, 118442. https://doi.org/10.1016/j.eswa.2022.118442

Liang, Z., Du, J., & Li, C. (2020). Abstractive social media text summarization using selective reinforced Seq2Seq attention model. *Neurocomputing*, *410*, 432–440. https://doi.org/10.1016/j.neucom.2020.04.137

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Springer International Publishing. https://doi.org/10.1007/978-3-031-02145-9

Ma, T., Pan, Q., Rong, H., Qian, Y., Tian, Y., & Al-Nabhan, N. (2022). T-BERTSum: Topic-Aware Text Summarization Based on BERT. *IEEE Transactions on Computational Social Systems*, *9*(3), 879–890. https://doi.org/10.1109/TCSS.2021.3088506

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space* (arXiv:1301.3781). arXiv. https://doi.org/10.48550/arXiv.1301.3781

Muniraj, P., Sabarmathi, K. R., Leelavathi, R., & B, S. B. (2023). HNTSumm: Hybrid text summarization of transliterated news articles. *International Journal of Intelligent Networks*, *4*, 53–61. https://doi.org/10.1016/j.ijin.2023.03.001

Nenkova, A. (2011). Automatic Summarization. *Foundations and Trends® in Information Retrieval*, *5*(2), 103–233. https://doi.org/10.1561/1500000015

Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1–135. https://doi.org/10.1561/1500000011

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543. https://doi.org/10.3115/v1/D14-1162

Sahith, B., Ganesh Reddy, T., Abhinay, Y., Biradar, V., Regulwar, G. B., & Begum, F. (2024). Enhanced Sentiment Analysis of Product Feedback: Leveraging Deep Learning Techniques. *2024 International Conference on Data Science and Network Security (ICDSNS)*, 1–6. https://doi.org/10.1109/ICDSNS62112.2024.10691192

ShafieiBavani, E., Ebrahimi, M., Wong, R., & Chen, F. (2018). A Graph-theoretic Summary Evaluation for ROUGE. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 762–767. https://doi.org/10.18653/v1/D18-1085

Suleiman, D., & Awajan, A. (2020). Deep Learning Based Abstractive Text Summarization: Approaches, Datasets, Evaluation Measures, and Challenges. *Mathematical Problems in Engineering*, *2020*, 1–29. https://doi.org/10.1155/2020/9365340

Tomer, M., & Kumar, M. (2020). Improving Text Summarization using Ensembled Approach based on Fuzzy with LSTM. *Arabian Journal for Science and Engineering*, *45*(12), 10743–10754. https://doi.org/10.1007/s13369-020-04827-6

Wazery, Y. M., Saleh, M. E., Alharbi, A., & Ali, A. A. (2022). Abstractive Arabic Text Summarization Based on Deep Learning. *Computational Intelligence and Neuroscience*, *2022*, 1–14. https://doi.org/10.1155/2022/1566890

Weinstein, C. J. (1994). Overview of the 1994 ARPA human language technology workshop. *Proceedings of the Workshop on Human Language Technology - HLT '94*, 3–5. https://doi.org/10.3115/1075812.1075813

LAMPIRAN

Lampiran Umpan Balik dan Ringkasan LSTM dengan GloVe

Humman Summary	***Customer Feedback***	***Predicted Summary***
The best in all that matters	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. The only one in crypto that offers you to actually call a phone number and get to someone with your questions just like that. Highly recommend! I have been a customer for more than a year and I have only good things to say about Celsius network.	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. I have been a customer for more than a year and I have only good things to say about Celsius network.
Celsius Network ROCKS!	If you are looking for the best #HomeForCrypto and where to earn steady yield then there is no better place than Celsius Network. The app is easy to use and understand with a recent update as well. The company is led by a great CEO who is fully engaged with the Celsius community! The mission of Celsius is to do good then do well and they live up to that very much. Rewards compound and pay out weekly!	If you are looking for the best #HomeForCrypto and where to earn steady yield then there is no better place than Celsius Network. The app is easy to use and understand with a recent update as well. The company is led by a great CEO who is fully engaged with the Celsius community!
I despise it so much	I despise it so much. Transferring to other wallets is difficult, especially because you cannot swap your coins, and there is a long waiting period when you do send to someone. Verification within 24 hours. When compared to Crypto.com, I find it extremely difficult to gain access to your funds. I'm never going to use it again. My biggest blunder!	Transferring to other wallets is difficult, especially because you cannot swap your coins, and there is a long waiting period when you do send to someone. When compared to Crypto.com, I find it extremely difficult to gain access to your funds. I'm never going to use it again.
Worst customer service and worst…	Worst customer service and worst company to deal with don't waste your time depositing your coins in there. You will get a run around. I am trying to withdraw coins from their platforms to another wallet and this been going on for over a month. Every time I call everyone is clueless and also they said that another country is in charge of the withdraw is all BS BE AWARE DONT WASTE YOUR TIME	Worst customer service and worst company to deal with don't waste your time depositing your coins in there.

Humman Summary	***Customer Feedback***	***Predicted Summary***
Celsius is the BEST in Crypto	Celsius is the most transparent and responsive company I've seen in the crypto space. I have never experienced any issues using their products, and they have consistently paid higher rates on crypto than other yield platforms. It is my first recommendation to crypto newbies and veterans alike. You will try it once and never want to leave!	Celsius is the most transparent and responsive company I've seen in the crypto space.
Best place to store crypto	Best wallet to manage, store and get rewards with my crypto. Using the app since 2019, Celsius never miss weekly rewards deposited to my wallet. Best place to buy and swap between different crypto asses all this services and more with zero fees. Just go and try it.Leo	Best wallet to manage, store and get rewards with my crypto. Using the app since 2019, Celsius never miss weekly rewards deposited to my wallet. Best place to buy and swap between different crypto asses all this services and more with zero fees.
Loan Collateral Held Hostage	I took out a loan but now I cant pay the principle back. I thought I was going to be liquidated which is fine because I can take back the collateral left over and reinvest. Now I am being told that my funds will be held indefinitely until I pay the principle (which I cant)	I took out a loan but now I cant pay the principle back.
Celsius Network is one of the best…	Celsius Network is one of the best Crypto Apps and Websites I have ever used. I earn a passive income with ease and not have to look at my phone as to watch the market, if my BTC/ETH is up or down every 5 mins. I earn an APY % on my Crypto and know that Celsius has my best interest at heart! I love Celsius Netowrk...HODL CELSIUS!!!	Celsius Network is one of the best Crypto Apps and Websites I have ever used. I earn an APY % on my Crypto and know that Celsius has my best interest at heart! I love Celsius Netowrk...HODL CELSIUS!!
Underrated!	Celsius is one of the most underrated crypto related web sites. Their referral codes are great, they have reward programs which are better than most of their competitors. The community is awesome. The founders are trustworthy. I have been using them	Celsius is one of the most underrated crypto related web sites. The community is awesome. The founders are trustworthy. I have been using them nearly for a year now, I had zero problems.

Humman Summary	***Customer Feedback***	***Predicted Summary***
	nearly for a year now, I had zero problems.	
Poor customer service! Atencion al cliente muy mala.	They suspended my account without reason. They didn't give me any opportunity to explain or at least answer their concerns, just shut down and I loose access to my coins!!! I've been waiting an answer to my ticket for more than a month! :(EDIT: The unlocked my account almost 3 moths later.Estoy super decepcionado, despues de casi 2 años como cliente de la nada me suspenden la cuenta. Pido explicaciones y nada, desastre.Tengo un ticket abierto hace más de 1 mes y no responden :(EDIT: desbloquearon mi cuent casi 3 meses despues.	They suspended my account without reason. I've been waiting an answer to my ticket for more than a month! :(EDIT: The unlocked my account almost 3 moths later.Estoy super decepcionado, despues de casi 2 años como cliente de la nada me suspenden la cuenta. Pido explicaciones y nada, desastre.Tengo un ticket abierto hace más de 1 mes y no responden :(EDIT: desbloquearon mi cuent casi 3 meses despues.
Ive being a celsiasn for over 2 years…	Ive being a celsiasn for over 2 years so excited to be part of the celsius comunity. I don’t miss to participate in the live ask our CEO @Alexmachinski every Tuesday in open spaces Twitter and fridays AMA you tube being hodling and even uses the loan application for .75% love to share with others and do good to them while we do good strongly recomend to Hodl you’re coins and gain yield helping other to reach 100 million users to a better financial freedom	Ive being a celsiasn for over 2 years so excited to be part of the celsius comunity. I don’t miss to participate in the live ask our CEO @Alexmachinski every Tuesday in open spaces Twitter and fridays AMA you tube being hodling and even uses the loan application for .75% love to share with others and do good to them while we do good strongly recomend to Hodl you’re coins and gain yield helping other to reach 100 million users to a better financial freedom
Best home for your assets and passive income!	Been on the platform for almost three years and couldn't be happier with my experience. While my state is stricter then most and I can only use the yield part of the app I am fine earning passive income on my assets. Secure platform and a great, welcoming community. What more could you ask for?	Been on the platform for almost three years and couldn't be happier with my experience. Secure platform and a great, welcoming community.

Humman Summary	***Customer Feedback***	***Predicted Summary***
Simply the Best!	Simply the Best!No fees ever for anything! Rewards paid weekly! Friendly customer service that you can actually call! Weekly updates from CEO on Celsius Network YouTube! 1% loans or less available! Great community supporting the company! Available in many nations! Great app design and top notch security!	Simply the Best!No fees ever for anything! Rewards paid weekly! Friendly customer service that you can actually call! Great community supporting the company! Available in many nations! Great app design and top notch security!
Great app, team and community!	Great app, team and community. Have been using Celsius since 2018 and never had a bad experience. I know the recent news about pausing withdrawals is concerning but they had to do restore liquidity and I am hopeful that they will resume withdrawals soon	Great app, team and community. Have been using Celsius since 2018 and never had a bad experience.
Celsius Network Promo Code	Took advantage of one of the Celsius Network promo codes. Waited 90 days for it to take effect. On the 90th day received a notification saying the funds had been credited to my account. Opened the app and didn't see the transaction until I scrolled down to when the transaction was initiated - 90 days ago. There it was. Very cool!	Took advantage of one of the Celsius Network promo codes. Opened the app and didn't see the transaction until I scrolled down to when the transaction was initiated - 90 days ago. Very cool!
Best yield and lending platform in…	Best yield and lending platform in crypto. Very consistent yield rates and the safest custodian for your assets. Fantastic customer service and they emphasise full transparency of how they distribute weekly rewards. Only company in crypto to have paid their community over $1Billion in yield. Oh and no fees!!!!!!	Best yield and lending platform in crypto. Very consistent yield rates and the safest custodian for your assets. Fantastic customer service and they emphasise full transparency of how they distribute weekly rewards.
be aware of these scammers	be aware of these scammers. they get your money and dont let you withdraw it. for every withdrawal you should go through backgroubd checks!!! they dont release your fund and ask you to upload a video a picture with your id paper. another id verification background. hell of a nightmare. you put your money there then you should beg to get it	for every withdrawal you should go through backgroubd checks!!! hell of a nightmare. be careful guys. i think they are running some scam.

Humman Summary	***Customer Feedback***	***Predicted Summary***
	back. be careful guys. i think they are running some scam.	
CeFi Project with Mediocre Rates	I have been with Celsius for several years. Never had problems… until recently. Seems like Celsius mismanaged their liquidity. They are currently unable to pay out their users.It might have been necessary and actually be in the interest of the community to disable withdrawals for now. However, if it wasn't for the mismanagement of funds, Celsius would not have been in that situation in the first place.Nice bonus: Celsius pays the network fees for outgoing transactions… should they ever resume.	I have been with Celsius for several years. Seems like Celsius mismanaged their liquidity.
Great app and community opening up opportunities	Great app, with a great community. I've been through the highs and have experienced the lows. This app and community gives you the opportunity to participate in financial tools not widely available. It is still up to you to know how to use these tools.	Great app, with a great community. I've been through the highs and have experienced the lows. This app and community gives you the opportunity to participate in financial tools not widely available.
This is the worst wallet ever	This is the worst wallet ever. To send funds out, you have to wait 24 hours to get the approval on the address. I found out today that when you send a large amount out, it can take up to ANOTHER 24 hours for them to send. They are no better than a bank holding your funds. The Unbank Yourself shirt the CEO Mashinsky wears is BS and hypocritical. Their customer service is basically non-existent except for snail mail.	This is the worst wallet ever. To send funds out, you have to wait 24 hours to get the approval on the address. Their customer service is basically non-existent except for snail mail.
Very bad customer service	Very bad customer service, doesn't respond at all, forget of receiving any response from them (e.g. if you are a new user and something is not clear to you). They give you codes to attract you telling we give you USD 50 and you only have to buy or transfer 400 USD worth crypto and keep it for 30 days, then when you actually wanna do it, the minimum	Very bad customer service, doesn't respond at all, forget of receiving any response from them (e.g. They give you codes to attract you telling we give you USD 50 and you only have to buy or transfer 400 USD worth crypto and keep it for 30 days,

Humman Summary	***Customer Feedback***	***Predicted Summary***
	transaction is EUR 850, or you pay a high transaction fee, not practical and not user-friendly.	then when you actually wanna do it, the minimum transaction is EUR 850, or you pay a high transaction fee, not practical and not user-friendly.
Deserved top-notch reputation	Celsius has a the reputation to want to please its customers and I did not really believe it.But after 2 years as a customer I must recognize that it's true ! I've only had great experiences using their product and the customer service is so quick, efficient and good-willed ! I'm a customer for life now.	Celsius has a the reputation to want to please its customers and I did not really believe it.But after 2 years as a customer I must recognize that it's true ! I've only had great experiences using their product and the customer service is so quick, efficient and good-willed !
I have been with Celsius for years	I have been with Celsius for years. Celsius is a great platform with no fees even to withdraw you coins. Also Celsius pays you interest on your tokens you hold on the Celsius platform. What is better then that?Great Company, Great customer service!!	I have been with Celsius for years. Celsius is a great platform with no fees even to withdraw you coins. What is better then that?Great Company, Great customer service!
Worst customer service I've seen so…	Worst customer service I've seen so far. The verification process on other similar sites takes about 5 minutes. Here, you cannot complete it in hours. They do not accept your photos for hours and when you try to contact customer support, you get instantly spammed by their bot. Look for alternatives.	Worst customer service I've seen so far. The verification process on other similar sites takes about 5 minutes.
I accidentally sent funds to Celsius…	I accidentally sent funds to Celsius using the BEP20 chain instead of ERC20 (it's easily done, these chains weren't an option when I joined Celsius). They offered to help me recover the funds and would send them back to me but then didn't bother communicating in about 6 months. How can they be trusted when they're holding funds like this? It's showing on the ledger in my wallet but no one will help.	I accidentally sent funds to Celsius using the BEP20 chain instead of ERC20 (it's easily done, these chains weren't an option when I joined Celsius). It's showing on the ledger in my wallet but no one will help.
awful customer service	awful customer service, sent an inquiry about missing matic on	awful customer service, sent an inquiry about

Humman Summary	***Customer Feedback***	***Predicted Summary***
	november 1st and still have nothing to show for it. been trading crypto for a while now, so i know its not an error on my end. You know something shady is going on because when you apply for a loan they reply to you within seconds/minutes, but for a response about an error on their end for missing coins it takes weeks/months. It's sad to see that a man like AleX maschinsky lets this mediocrity slide in his company.	missing matic on november 1st and still have nothing to show for it. been trading crypto for a while now, so i know its not an error on my end. You know something shady is going on because when you apply for a loan they reply to you within seconds/minutes, but for a response about an error on their end for missing coins it takes weeks/months.
Very bad company and communication. SKAM please avoid	Very bad company and communication. I have not received referral reward more than a month. They have told me that my referral have made successfully deposited but rewards are not here. More the a month and 15 emails I still receive answers like,, your request is at our tech department,,. Come on guys you can do better than this. At the end this is big SKAM please avoid.	Very bad company and communication. I have not received referral reward more than a month. Come on guys you can do better than this.
Their support to unfreeze my account…	Their support to unfreeze my account after suspicious activity has been NONexistent! I like the interest rates, but the support makes that all worthless if you can not get into your account. I had considered doubling my the amount I have in my account, but after the problems continuing for a week now, I don't think I will be doing that any time soon. So much for their promised "48" hour policy .	Their support to unfreeze my account after suspicious activity has been NONexistent! I had considered doubling my the amount I have in my account, but after the problems continuing for a week now, I don't think I will be doing that any time soon.
Horrible experience	Horrible experience, feels like a scam. I created account and sent little amount of crypto, everything seemed ok. I deposited some more. And... app logged out, I was trying to log in but it didn't allow to paste 2FA. I tried several times and they locked my account. Support doesn't answer. Now I don't know if I will get my crypto back	Horrible experience, feels like a scam. Now I don't know if I will get my crypto back

Humman Summary	*Customer Feedback*	*Predicted Summary*
3 reasons why I don't trust anyone else	1. Do you know any other company who is as transparent as Celsius?this means regulatory as well as CEO Alex Mashinsky's own availability for everyone to question2. $CEL token is deflationary3. The app and services offered are specifically built in the best interest for long term holders	Do you know any other company who is as transparent as Celsius?this means regulatory as well as CEO Alex Mashinsky's own availability for everyone to question2. The app and services offered are specifically built in the best interest for long term holders

Lampiran Umpan Balik dan Ringkasan LSTM dengan *FastText*

Humman Summary	*Customer Feedback*	*Predicted Summary*
The best in all that matters	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. The only one in crypto that offers you to actually call a phone number and get to someone with your questions just like that. Highly recommend! I have been a customer for more than a year and I have only good things to say about Celsius network.	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. Highly recommend! I have been a customer for more than a year and I have only good things to say about Celsius network.
Celsius Network ROCKS!	If you are looking for the best #HomeForCrypto and where to earn steady yield then there is no better place than Celsius Network. The app is easy to use and understand with a recent update as well. The company is led by a great CEO who is fully engaged with the Celsius community! The mission of Celsius is to do good then do well and they live up to that very much. Rewards compound and pay out weekly!	The app is easy to use and understand with a recent update as well. The company is led by a great CEO who is fully engaged with the Celsius community! The mission of Celsius is to do good then do well and they live up to that very much.
I despise it so much	I despise it so much. Transferring to other wallets is difficult, especially because you cannot swap your coins, and there is a long waiting period when you do	Transferring to other wallets is difficult, especially because you cannot swap your coins, and there is a long

Humman Summary	***Customer Feedback***	***Predicted Summary***
	send to someone. Verification within 24 hours. When compared to Crypto.com, I find it extremely difficult to gain access to your funds. I'm never going to use it again. My biggest blunder!	waiting period when you do send to someone. When compared to Crypto.com, I find it extremely difficult to gain access to your funds. I'm never going to use it again. My biggest blunder!
Worst customer service and worst…	Worst customer service and worst company to deal with don't waste your time depositing your coins in there. You will get a run around. I am trying to withdraw coins from their platforms to another wallet and this been going on for over a month. Every time I call everyone is clueless and also they said that another country is in charge of the withdraw is all BS BE AWARE DONT WASTE YOUR TIME	Worst customer service and worst company to deal with don't waste your time depositing your coins in there.
Celsius is the BEST in Crypto	Celsius is the most transparent and responsive company I've seen in the crypto space. I have never experienced any issues using their products, and they have consistently paid higher rates on crypto than other yield platforms. It is my first recommendation to crypto newbies and veterans alike. You will try it once and never want to leave!	Celsius is the most transparent and responsive company I've seen in the crypto space. I have never experienced any issues using their products, and they have consistently paid higher rates on crypto than other yield platforms. It is my first recommendation to crypto newbies and veterans alike.
Best place to store crypto	Best wallet to manage, store and get rewards with my crypto. Using the app since 2019, Celsius never miss weekly rewards deposited to my wallet. Best place to buy and swap between different crypto asses all	Best wallet to manage, store and get rewards with my crypto. Using the app since 2019, Celsius never miss weekly rewards deposited to my wallet. Best place to buy and

Humman Summary	*Customer Feedback*	*Predicted Summary*
	this services and more with zero fees. Just go and try it.Leo	swap between different crypto asses all this services and more with zero fees.
Loan Collateral Held Hostage	I took out a loan but now I cant pay the principle back. I thought I was going to be liquidated which is fine because I can take back the collateral left over and reinvest. Now I am being told that my funds will be held indefinitely until I pay the principle (which I cant)	I took out a loan but now I cant pay the principle back.
Celsius Network is one of the best…	Celsius Network is one of the best Crypto Apps and Websites I have ever used. I earn a passive income with ease and not have to look at my phone as to watch the market, if my BTC/ETH is up or down every 5 mins. I earn an APY % on my Crypto and know that Celsius has my best interest at heart! I love Celsius Netowrk...HODL CELSIUS!!!	Celsius Network is one of the best Crypto Apps and Websites I have ever used. I love Celsius Netowrk...HODL CELSIUS!!
Underrated!	Celsius is one of the most underrated crypto related web sites. Their referral codes are great, they have reward programs which are better than most of their competitors. The community is awesome. The founders are trustworthy. I have been using them nearly for a year now, I had zero problems.	Celsius is one of the most underrated crypto related web sites. The community is awesome. The founders are trustworthy. I have been using them nearly for a year now, I had zero problems.
Poor customer service! Atencion al cliente muy mala.	They suspended my account without reason. They didn't give me any opportunity to explain or at least answer their concerns, just shut down and I loose access to my coins!!! I've been waiting an answer to my ticket for more than a month! :(EDIT: The unlocked my account almost 3 moths later.Estoy super decepcionado, despues de casi 2	I've been waiting an answer to my ticket for more than a month! :(EDIT: The unlocked my account almost 3 moths later.Estoy super decepcionado, despues de casi 2 años como cliente de la nada me suspenden la cuenta.

Humman Summary	*Customer Feedback*	*Predicted Summary*
	años como cliente de la nada me suspenden la cuenta. Pido explicaciones y nada, desastre.Tengo un ticket abierto hace más de 1 mes y no responden :(EDIT: desbloquearon mi cuent casi 3 meses despues.	
Ive being a celsiasn for over 2 years…	Ive being a celsiasn for over 2 years so excited to be part of the celsius comunity. I don't miss to participate in the live ask our CEO @Alexmachinski every Tuesday in open spaces Twitter and fridays AMA you tube being hodling and even uses the loan application for .75% love to share with others and do good to them while we do good strongly recomend to Hodl you're coins and gain yield helping other to reach 100 million users to a better financial freedom	Ive being a celsiasn for over 2 years so excited to be part of the celsius comunity. I don't miss to participate in the live ask our CEO @Alexmachinski every Tuesday in open spaces Twitter and fridays AMA you tube being hodling and even uses the loan application for .75% love to share with others and do good to them while we do good strongly recomend to Hodl you're coins and gain yield helping other to reach 100 million users to a better financial freedom
Best home for your assets and passive income!	Been on the platform for almost three years and couldn't be happier with my experience. While my state is stricter then most and I can only use the yield part of the app I am fine earning passive income on my assets. Secure platform and a great, welcoming community. What more could you ask for?	Been on the platform for almost three years and couldn't be happier with my experience. Secure platform and a great, welcoming community.
Simply the Best!	Simply the Best!No fees ever for anything! Rewards paid weekly! Friendly customer service that you can actually call! Weekly updates from CEO on Celsius Network YouTube! 1% loans or	Simply the Best!No fees ever for anything! Friendly customer service that you can actually call! Great community supporting

Humman Summary	***Customer Feedback***	***Predicted Summary***
	less available! Great community supporting the company! Available in many nations! Great app design and top notch security!	the company! Great app design and top notch security!
Great app, team and community!	Great app, team and community. Have been using Celsius since 2018 and never had a bad experience. I know the recent news about pausing withdrawals is concerning but they had to do restore liquidity and I am hopeful that they will resume withdrawals soon	Great app, team and community. Have been using Celsius since 2018 and never had a bad experience.
Celsius Network Promo Code	Took advantage of one of the Celsius Network promo codes. Waited 90 days for it to take effect. On the 90th day received a notification saying the funds had been credited to my account. Opened the app and didn't see the transaction until I scrolled down to when the transaction was initiated - 90 days ago. There it was. Very cool!	Took advantage of one of the Celsius Network promo codes. Opened the app and didn't see the transaction until I scrolled down to when the transaction was initiated - 90 days ago. Very cool!
Best yield and lending platform in…	Best yield and lending platform in crypto. Very consistent yield rates and the safest custodian for your assets. Fantastic customer service and they emphasise full transparency of how they distribute weekly rewards. Only company in crypto to have paid their community over $1Billion in yield. Oh and no fees!!!!!!	Best yield and lending platform in crypto. Very consistent yield rates and the safest custodian for your assets. Fantastic customer service and they emphasise full transparency of how they distribute weekly rewards. Only company in crypto to have paid their community over $1Billion in yield.
be aware of these scammers	be aware of these scammers. they get your money and dont let you withdraw it. for every withdrawal you should go through backgroubd checks!!! they dont release your fund and ask you to	for every withdrawal you should go through backgroubd checks!!! hell of a nightmare.

Humman Summary	***Customer Feedback***	***Predicted Summary***
	upload a video a picture with your id paper. another id verification background. hell of a nightmare. you put your money there then you should beg to get it back. be careful guys. i think they are running some scam.	
CeFi Project with Mediocre Rates	I have been with Celsius for several years. Never had problems… until recently. Seems like Celsius mismanaged their liquidity. They are currently unable to pay out their users.It might have been necessary and actually be in the interest of the community to disable withdrawals for now. However, if it wasn't for the mismanagement of funds, Celsius would not have been in that situation in the first place.Nice bonus: Celsius pays the network fees for outgoing transactions… should they ever resume.	I have been with Celsius for several years. Seems like Celsius mismanaged their liquidity.
Great app and community opening up opportunities	Great app, with a great community. I've been through the highs and have experienced the lows. This app and community gives you the opportunity to participate in financial tools not widely available. It is still up to you to know how to use these tools.	Great app, with a great community. I've been through the highs and have experienced the lows. This app and community gives you the opportunity to participate in financial tools not widely available.
This is the worst wallet ever	This is the worst wallet ever. To send funds out, you have to wait 24 hours to get the approval on the address. I found out today that when you send a large amount out, it can take up to ANOTHER 24 hours for them to send. They are no better than a bank holding your funds. The Unbank Yourself shirt the CEO	This is the worst wallet ever. They are no better than a bank holding your funds. Their customer service is basically non-existent except for snail mail.

Humman Summary	*Customer Feedback*	*Predicted Summary*
	Mashinsky wears is BS and hypocritical. Their customer service is basically non-existent except for snail mail.	
Very bad customer service	Very bad customer service, doesn't respond at all, forget of receiving any response from them (e.g. if you are a new user and something is not clear to you). They give you codes to attract you telling we give you USD 50 and you only have to buy or transfer 400 USD worth crypto and keep it for 30 days, then when you actually wanna do it, the minimum transaction is EUR 850, or you pay a high transaction fee, not practical and not user-friendly.	Very bad customer service, doesn't respond at all, forget of receiving any response from them (e.g. They give you codes to attract you telling we give you USD 50 and you only have to buy or transfer 400 USD worth crypto and keep it for 30 days, then when you actually wanna do it, the minimum transaction is EUR 850, or you pay a high transaction fee, not practical and not user-friendly.
Deserved top-notch reputation	Celsius has a the reputation to want to please its customers and I did not really believe it.But after 2 years as a customer I must recognize that it's true ! I've only had great experiences using their product and the customer service is so quick, efficient and good-willed ! I'm a customer for life now.	Celsius has a the reputation to want to please its customers and I did not really believe it.But after 2 years as a customer I must recognize that it's true ! I've only had great experiences using their product and the customer service is so quick, efficient and good-willed !
I have been with Celsius for years	I have been with Celsius for years. Celsius is a great platform with no fees even to withdraw you coins. Also Celsius pays you interest on your tokens you hold on the Celsius platform. What is better then that?Great Company, Great customer service!!	I have been with Celsius for years. Celsius is a great platform with no fees even to withdraw you coins. Also Celsius pays you interest on your tokens you hold on the Celsius platform. What is better then

Humman Summary	*Customer Feedback*	*Predicted Summary*
		that?Great Company, Great customer service!
Worst customer service I've seen so…	Worst customer service I've seen so far. The verification process on other similar sites takes about 5 minutes. Here, you cannot complete it in hours. They do not accept your photos for hours and when you try to contact customer support, you get instantly spammed by their bot. Look for alternatives.	Worst customer service I've seen so far. They do not accept your photos for hours and when you try to contact customer support, you get instantly spammed by their bot.
I accidentally sent funds to Celsius…	I accidentally sent funds to Celsius using the BEP20 chain instead of ERC20 (it's easily done, these chains weren't an option when I joined Celsius). They offered to help me recover the funds and would send them back to me but then didn't bother communicating in about 6 months. How can they be trusted when they're holding funds like this? It's showing on the ledger in my wallet but no one will help.	I accidentally sent funds to Celsius using the BEP20 chain instead of ERC20 (it's easily done, these chains weren't an option when I joined Celsius).
awful customer service	awful customer service, sent an inquiry about missing matic on november 1st and still have nothing to show for it. been trading crypto for a while now, so i know its not an error on my end. You know something shady is going on because when you apply for a loan they reply to you within seconds/minutes, but for a response about an error on their end for missing coins it takes weeks/months. It's sad to see that a man like AleX maschinsky lets this mediocrity slide in his company.	awful customer service, sent an inquiry about missing matic on november 1st and still have nothing to show for it. been trading crypto for a while now, so i know its not an error on my end. You know something shady is going on because when you apply for a loan they reply to you within seconds/minutes, but for a response about an error on their end for missing coins it takes weeks/months.
Very bad company and	Very bad company and communication. I have not	Very bad company and communication. I have

Humman Summary	***Customer Feedback***	***Predicted Summary***
communication. SKAM please avoid	received referral reward more than a month. They have told me that my referral have made successfully deposited but rewards are not here. More the a month and 15 emails I still receive answers like,, your request is at our tech department,,. Come on guys you can do better than this. At the end this is big SKAM please avoid.	not received referral reward more than a month. At the end this is big SKAM please avoid.
Their support to unfreeze my account…	Their support to unfreeze my account after suspicious activity has been NONexistent! I like the interest rates, but the support makes that all worthless if you can not get into your account. I had considered doubling my the amount I have in my account, but after the problems continuing for a week now, I don't think I will be doing that any time soon. So much for their promised "48" hour policy .	Their support to unfreeze my account after suspicious activity has been NONexistent! I like the interest rates, but the support makes that all worthless if you can not get into your account. I had considered doubling my the amount I have in my account, but after the problems continuing for a week now, I don't think I will be doing that any time soon.
Horrible experience	Horrible experience, feels like a scam. I created account and sent little amount of crypto, everything seemed ok. I deposited some more. And... app logged out, I was trying to log in but it didn't allow to paste 2FA. I tried several times and they locked my account. Support doesn't answer. Now I don't know if I will get my crypto back	Horrible experience, feels like a scam. I created account and sent little amount of crypto, everything seemed ok.
3 reasons why I don't trust anyone else	1. Do you know any other company who is as transparent as Celsius?this means regulatory as well as CEO Alex Mashinsky's own availability for everyone to question2. $CEL token is deflationary3. The app and	$CEL token is deflationary3. The app and services offered are specifically built in the best interest for long term holders

Humman Summary	*Customer Feedback*	*Predicted Summary*
	services offered are specifically built in the best interest for long term holders	

Lampiran Umpan Balik dan Ringkasan LSTM tanpa *Word Embedding*

Humman Summary	*Customer Feedback*	*Predicted Summary*
The best in all that matters	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. The only one in crypto that offers you to actually call a phone number and get to someone with your questions just like that. Highly recommend! I have been a customer for more than a year and I have only good things to say about Celsius network.	The best in all that matters! It's a great platform, easy and simple to use, and beginner-friendly. I have been a customer for more than a year and I have only good things to say about Celsius network.
Celsius Network ROCKS!	If you are looking for the best #HomeForCrypto and where to earn steady yield then there is no better place than Celsius Network. The app is easy to use and understand with a recent update as well. The company is led by a great CEO who is fully engaged with the Celsius community! The mission of Celsius is to do good then do well and they live up to that very much. Rewards compound and pay out weekly!	If you are looking for the best #HomeForCrypto and where to earn steady yield then there is no better place than Celsius Network. The app is easy to use and understand with a recent update as well.
I despise it so much	I despise it so much. Transferring to other wallets is difficult, especially because you cannot swap your coins, and there is a long waiting period when you do send to someone. Verification within 24 hours. When compared to Crypto.com, I find it extremely difficult to gain access to your funds. I'm never going to use it again. My biggest blunder!	I despise it so much.

Humman Summary	***Customer Feedback***	***Predicted Summary***
Worst customer service and worst…	Worst customer service and worst company to deal with don't waste your time depositing your coins in there. You will get a run around. I am trying to withdraw coins from their platforms to another wallet and this been going on for over a month. Every time I call everyone is clueless and also they said that another country is in charge of the withdraw is all BS BE AWARE DONT WASTE YOUR TIME	Worst customer service and worst company to deal with don't waste your time depositing your coins in there.
Celsius is the BEST in Crypto	Celsius is the most transparent and responsive company I've seen in the crypto space. I have never experienced any issues using their products, and they have consistently paid higher rates on crypto than other yield platforms. It is my first recommendation to crypto newbies and veterans alike. You will try it once and never want to leave!	Celsius is the most transparent and responsive company I've seen in the crypto space. It is my first recommendation to crypto newbies and veterans alike.
Best place to store crypto	Best wallet to manage, store and get rewards with my crypto. Using the app since 2019, Celsius never miss weekly rewards deposited to my wallet. Best place to buy and swap between different crypto asses all this services and more with zero fees. Just go and try it.Leo	Best wallet to manage, store and get rewards with my crypto. Using the app since 2019, Celsius never miss weekly rewards deposited to my wallet. Best place to buy and swap between different crypto asses all this services and more with zero fees.
Loan Collateral Held Hostage	I took out a loan but now I cant pay the principle back. I thought I was going to be liquidated which is fine because I can take back the collateral left over and reinvest. Now I am being told that my funds will be held	I took out a loan but now I cant pay the principle back.

Humman Summary	*Customer Feedback*	*Predicted Summary*
	indefinitely until I pay the principle (which I cant)	
Celsius Network is one of the best…	Celsius Network is one of the best Crypto Apps and Websites I have ever used. I earn a passive income with ease and not have to look at my phone as to watch the market, if my BTC/ETH is up or down every 5 mins. I earn an APY % on my Crypto and know that Celsius has my best interest at heart! I love Celsius Netowrk...HODL CELSIUS!!!	Celsius Network is one of the best Crypto Apps and Websites I have ever used. I earn a passive income with ease and not have to look at my phone as to watch the market, if my BTC/ETH is up or down every 5 mins. I earn an APY % on my Crypto and know that Celsius has my best interest at heart! I love Celsius Netowrk...HODL CELSIUS!!
Underrated!	Celsius is one of the most underrated crypto related web sites. Their referral codes are great, they have reward programs which are better than most of their competitors. The community is awesome. The founders are trustworthy. I have been using them nearly for a year now, I had zero problems.	Celsius is one of the most underrated crypto related web sites.
Poor customer service! Atencion al cliente muy mala.	They suspended my account without reason. They didn't give me any opportunity to explain or at least answer their concerns, just shut down and I loose access to my coins!!! I've been waiting an answer to my ticket for more than a month! :(EDIT: The unlocked my account almost 3 moths later.Estoy super decepcionado, despues de casi 2 años como cliente de la nada me suspenden la cuenta. Pido explicaciones y nada, desastre.Tengo un ticket abierto hace más de 1 mes y no responden :(EDIT:	They suspended my account without reason. :(EDIT: The unlocked my account almost 3 moths later.Estoy super decepcionado, despues de casi 2 años como cliente de la nada me suspenden la cuenta.

Humman Summary	***Customer Feedback***	***Predicted Summary***
	desbloquearon mi cuent casi 3 meses despues.	
Ive being a celsiasn for over 2 years…	Ive being a celsiasn for over 2 years so excited to be part of the celsius comunity. I don't miss to participate in the live ask our CEO @Alexmachinski every Tuesday in open spaces Twitter and fridays AMA you tube being hodling and even uses the loan application for .75% love to share with others and do good to them while we do good strongly recomend to Hodl you're coins and gain yield helping other to reach 100 million users to a better financial freedom	Ive being a celsiasn for over 2 years so excited to be part of the celsius comunity.
Best home for your assets and passive income!	Been on the platform for almost three years and couldn't be happier with my experience. While my state is stricter then most and I can only use the yield part of the app I am fine earning passive income on my assets. Secure platform and a great, welcoming community. What more could you ask for?	Been on the platform for almost three years and couldn't be happier with my experience. Secure platform and a great, welcoming community.
Simply the Best!	Simply the Best!No fees ever for anything! Rewards paid weekly! Friendly customer service that you can actually call! Weekly updates from CEO on Celsius Network YouTube! 1% loans or less available! Great community supporting the company! Available in many nations! Great app design and top notch security!	Simply the Best!No fees ever for anything! Great app design and top notch security!
Great app, team and community!	Great app, team and community. Have been using Celsius since 2018 and never had a bad experience. I know the recent news about pausing withdrawals is concerning but they had to do restore liquidity and I am hopeful	Have been using Celsius since 2018 and never had a bad experience.

Humman Summary	***Customer Feedback***	***Predicted Summary***
	that they will resume withdrawals soon	
Celsius Network Promo Code	Took advantage of one of the Celsius Network promo codes. Waited 90 days for it to take effect. On the 90th day received a notification saying the funds had been credited to my account. Opened the app and didn't see the transaction until I scrolled down to when the transaction was initiated - 90 days ago. There it was. Very cool!	Took advantage of one of the Celsius Network promo codes. Opened the app and didn't see the transaction until I scrolled down to when the transaction was initiated - 90 days ago. Very cool!
Best yield and lending platform in…	Best yield and lending platform in crypto. Very consistent yield rates and the safest custodian for your assets. Fantastic customer service and they emphasise full transparency of how they distribute weekly rewards. Only company in crypto to have paid their community over $1Billion in yield. Oh and no fees!!!!!!	Best yield and lending platform in crypto. Fantastic customer service and they emphasise full transparency of how they distribute weekly rewards.
be aware of these scammers	be aware of these scammers. they get your money and dont let you withdraw it. for every withdrawal you should go through backgroubd checks!!! they dont release your fund and ask you to upload a video a picture with your id paper. another id verification background. hell of a nightmare. you put your money there then you should beg to get it back. be careful guys. i think they are running some scam.	be careful guys.
CeFi Project with Mediocre Rates	I have been with Celsius for several years. Never had problems… until recently. Seems like Celsius mismanaged their liquidity. They are currently unable to pay out their users.It might have been necessary and actually be in the interest of the community to disable	I have been with Celsius for several years. Seems like Celsius mismanaged their liquidity.

Humman Summary	***Customer Feedback***	***Predicted Summary***
	withdrawals for now. However, if it wasn't for the mismanagement of funds, Celsius would not have been in that situation in the first place.Nice bonus: Celsius pays the network fees for outgoing transactions… should they ever resume.	
Great app and community opening up opportunities	Great app, with a great community. I've been through the highs and have experienced the lows. This app and community gives you the opportunity to participate in financial tools not widely available. It is still up to you to know how to use these tools.	Great app, with a great community.
This is the worst wallet ever	This is the worst wallet ever. To send funds out, you have to wait 24 hours to get the approval on the address. I found out today that when you send a large amount out, it can take up to ANOTHER 24 hours for them to send. They are no better than a bank holding your funds. The Unbank Yourself shirt the CEO Mashinsky wears is BS and hypocritical. Their customer service is basically non-existent except for snail mail.	This is the worst wallet ever.
Very bad customer service	Very bad customer service, doesn't respond at all, forget of receiving any response from them (e.g. if you are a new user and something is not clear to you). They give you codes to attract you telling we give you USD 50 and you only have to buy or transfer 400 USD worth crypto and keep it for 30 days, then when you actually wanna do it, the minimum transaction is EUR 850, or you pay a high	Very bad customer service, doesn't respond at all, forget of receiving any response from them (e.g. They give you codes to attract you telling we give you USD 50 and you only have to buy or transfer 400 USD worth crypto and keep it for 30 days, then when you actually wanna do it, the

Humman Summary	***Customer Feedback***	***Predicted Summary***
	transaction fee, not practical and not user-friendly.	minimum transaction is EUR 850, or you pay a high transaction fee, not practical and not user-friendly.
Deserved top-notch reputation	Celsius has a the reputation to want to please its customers and I did not really believe it.But after 2 years as a customer I must recognize that it's true ! I've only had great experiences using their product and the customer service is so quick, efficient and good-willed ! I'm a customer for life now.	Celsius has a the reputation to want to please its customers and I did not really believe it.But after 2 years as a customer I must recognize that it's true ! I've only had great experiences using their product and the customer service is so quick, efficient and good-willed !
I have been with Celsius for years	I have been with Celsius for years. Celsius is a great platform with no fees even to withdraw you coins. Also Celsius pays you interest on your tokens you hold on the Celsius platform. What is better then that?Great Company, Great customer service!!	I have been with Celsius for years. Celsius is a great platform with no fees even to withdraw you coins. What is better then that?Great Company, Great customer service!
Worst customer service I've seen so…	Worst customer service I've seen so far. The verification process on other similar sites takes about 5 minutes. Here, you cannot complete it in hours. They do not accept your photos for hours and when you try to contact customer support, you get instantly spammed by their bot. Look for alternatives.	Worst customer service I've seen so far.
I accidentally sent funds to Celsius…	I accidentally sent funds to Celsius using the BEP20 chain instead of ERC20 (it's easily done, these chains weren't an option when I joined Celsius). They offered to help me recover the funds and would send them back to me but then didn't bother	I accidentally sent funds to Celsius using the BEP20 chain instead of ERC20 (it's easily done, these chains weren't an option when I joined Celsius).

Humman Summary	***Customer Feedback***	***Predicted Summary***
	communicating in about 6 months. How can they be trusted when they're holding funds like this? It's showing on the ledger in my wallet but no one will help.	
awful customer service	awful customer service, sent an inquiry about missing matic on november 1st and still have nothing to show for it. been trading crypto for a while now, so i know its not an error on my end. You know something shady is going on because when you apply for a loan they reply to you within seconds/minutes, but for a response about an error on their end for missing coins it takes weeks/months. It's sad to see that a man like AleX maschinsky lets this mediocrity slide in his company.	awful customer service, sent an inquiry about missing matic on november 1st and still have nothing to show for it. been trading crypto for a while now, so i know its not an error on my end.
Very bad company and communication. SKAM please avoid	Very bad company and communication. I have not received referral reward more than a month. They have told me that my referral have made successfully deposited but rewards are not here. More the a month and 15 emails I still receive answers like,, your request is at our tech department,,. Come on guys you can do better than this. At the end this is big SKAM please avoid.	Very bad company and communication. I have not received referral reward more than a month.
Their support to unfreeze my account…	Their support to unfreeze my account after suspicious activity has been NONexistent! I like the interest rates, but the support makes that all worthless if you can not get into your account. I had considered doubling my the amount I have in my account, but after the problems continuing for a week now, I don't think I will be doing that any time soon. So	I like the interest rates, but the support makes that all worthless if you can not get into your account.

Humman Summary	***Customer Feedback***	***Predicted Summary***
	much for their promised "48" hour policy .	
Horrible experience	Horrible experience, feels like a scam. I created account and sent little amount of crypto, everything seemed ok. I deposited some more. And... app logged out, I was trying to log in but it didn't allow to paste 2FA. I tried several times and they locked my account. Support doesn't answer. Now I don't know if I will get my crypto back	Horrible experience, feels like a scam.
3 reasons why I don't trust anyone else	1. Do you know any other company who is as transparent as Celsius?this means regulatory as well as CEO Alex Mashinsky's own availability for everyone to question2. $CEL token is deflationary3. The app and services offered are specifically built in the best interest for long term holders	1.