Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)

Journal Homepage: http://journal.lembagakita.org/index.php/jtik

Sentimen Analisis dengan *Long Short-Term Memory* dan *Synthetic Minority Over Sampling Technic* Pada Aplikasi Digital Perbankan

Ali Ahmad [1], Windu Gata [2*], Supriadi Panggabean [3]

[1,2,3] Program Studi Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Nusa Mandiri, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta, Indonesia.*

a r t i c l e i n f o

Article history:
Received 28 February 2024
Received in revised form
14 May 2024
Accepted 30 July 2024
Available online October
2024.

DOI:
https://doi.org/10.35870/jtik.v8i4.2320.

Keywords:
Sentiment Analysys; Text Mining; LSTM; Class Imbalance.

a b s t r a c t

In recent years, we have witnessed significant growth in digital banking transactions, supported by technological advancements. According to the latest data from the FinTech Association of Indonesia, digital banking transactions in Indonesia have increased by 35% from the previous year. In this context, the development of digital banking applications becomes increasingly important. However, to ensure the quality and success of these applications, feedback from users is crucial. One technique used by banks is sentiment analysis to gather feedback on their digital applications. This research aims to analyze user sentiment for two banking applications, DbankPro and M-BCA, through reviews on the Google Playstore. The method used is CRISP-DM, implementing the "Imbalance Data Handling with SMOTE" technique and LSTM model. The test results show the accuracy of sentiment analysis for M-BCA is 91.07%, while for DbankPro it is 89.82%. The implications of this research emphasize the importance of paying attention to user feedback in the development of digital banking applications to enhance their quality and meet user expectations.

Kata Kunci:
Analisa Sentimen; Text Mining; LSTM; Class Imbalance.

a b s t r a k

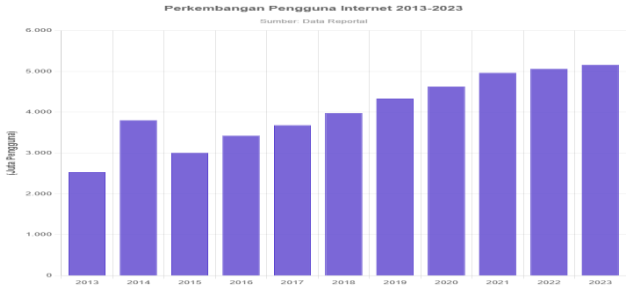
Dalam beberapa tahun terakhir, kita telah menyaksikan pertumbuhan yang signifikan dalam transaksi perbankan digital, didukung oleh kemajuan teknologi. Menurut data terbaru dari Asosiasi FinTech Indonesia, transaksi perbankan digital di Indonesia meningkat sebesar 35% dari tahun sebelumnya. Dalam konteks ini, perkembangan aplikasi perbankan digital menjadi semakin penting. Namun, untuk memastikan kualitas dan keberhasilan aplikasi tersebut, umpan balik dari pengguna sangat diperlukan. Salah satu teknik yang digunakan oleh bank adalah analisis sentimen untuk mendapatkan umpan balik terhadap aplikasi digital mereka. Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap dua aplikasi perbankan, yaitu DbankPro dan M-BCA, melalui ulasan di Google Playstore. Metode yang digunakan adalah CRISP-DM, dengan menerapkan teknik "Imbalance Data Handling with SMOTE" dan model LSTM. Hasil pengujian menunjukkan akurasi analisis sentimen untuk M-BCA sebesar 91,07%, sementara untuk DbankPro sebesar 89,82%. Implikasi dari penelitian ini adalah pentingnya memperhatikan umpan balik pengguna dalam pengembangan aplikasi perbankan digital untuk meningkatkan kualitasnya dan memenuhi harapan pengguna.

Corresponding Author. Email: windu.gata@nusamandiri.ac.id [2].

ACM Computing Classification System (CCS)
EBSCOhost
Communication and Mass Media Complete (CMMC)

1. Latar Belakang

Perkembangan tekhnologi dalam beberapa tahun terakhir mengalami peningkatan yang cukup signifikan, salah satu perkembangan yang mengalami percepatan yaitu perkembangan teknologi internet. Pertumbuhan ini mengalami peningkatan terutama pada saat pandemi Covid-19 melanda dunia. Penerapan Isolasi mandiri oleh pemerintah dan tuntutan untuk menjalankan aktifitas pekerjaan, pembelajaran dan belanja tetap berjalan dengan normal menuntut implementasi akselerasi percepatan pemanfaatan internen dan teknologi digital. Menurut data.goodstats.id melansir terdapat 5,16 Miliar pengguna internet secara global ini menggambarkan 64% dari populasi dunia sudah terhubung ke internet [1]. Seperti terlihat pada gambar 1 dibawah ini.



Gambar 1. Perkembangan Internet 2013 – 2023

Dalam era digital aplikasi *mobile* telah mengambil peranan penting dalam mendukung kegiatan sehari-hari. Terlebih aplikasi finansial yang banyak tersedia didalam platform *mobile* terkhusus aplikasi digital perbankan seperti D-Bank Pro dari Bank Danamon, Livin dari Bank Mandiri dan Bank Cetnral Asia (BCA) dengan BCA mobilenya. Aplikasi ini memudahkan dalam bertransaksi dalam keadaan mobilitas yang tinggi ataupun terbatas. Aplikasi digital tersebut bisa didapatkan melalui Aplikasi *Google Play Store*. Dalam hal ini aplikasi *Google Play Store* juga memberikan ruang untuk penggunanya memberikan review atau ulasan terhadap aplikasi yang telah digunakan. Ulasan ini dapat membantu penyedia aplikasi untuk berinovasi untuk memperbaiki aplikasi digital perbankan yang mereka miliki, sehingga proses analisa sentimen pada urain tersebut perlu dilakukan dengan harapan memudahkan bagi pemiliki aplikasi mendapatkan gambaran tingkat kepuasan pengguna aplikasi, saran dan masukan untuk pengembangan aplikasi yang lebih tepat sasaran.

Analisis sentimen diperlukan dengan tujuan untuk mengidentifikasi, mengekstrak, dan memahami sentimen atau opini yang terkandung dalam teks, sehingga dapat memberikan wawasan tentang bagaimana pengguna merespons dan menilai aplikasi tertentu [2]. Analisa sentimen memiliki masalah yang cukup mendasar diantaranya hasil data yang berupa uraian/review setelah setelah melalui proses normalisasi terdapat data yang terlalu dominan (Mayoritas) dibandingkan dengan data lainnya. Keadaan ini di sebut dengan 'Imbalance Data'. pada penelitian ini, peneliti akan membagi sentimen yang terdapat pada uraian aplikasi digital perbankan menjadi positif, netral dan negatif statemen. Contoh kondisi imbalance yaitu rentang jarak antara sentimen positif dengan netral terdapat gap yang cukup jauh. Dengan kondisi seperti ini kecenderungan program untuk memprediksi dengan lebih baik terdapat pada statemen yang positif (Mayoritas) dan sebaliknya kecenderungan kurang baik jika memprediksi pada statement yang netral (Minoritas). Pada riset tentang aplikasi digital perbankan pada salah satu bank digital (BNC) dengan menggunakan metode Support Vector Machine (SVM) Algorithm didapatkan hasil Accuracy sebesar 82.33% dengan beberapa scenario dengan hasil terbaik menggunakan skenario split data 90% data training dan 10% test data [3].

Pada penelitian lainnya terkait dengan sentiment analisis aplikasi digital perbankan pada salah satu bank konvensional besar di Indonesia yaitu Bank Central Asia (BCA) dengan nama aplikasi BCA Mobile didapatkan nilai TN Positif sebesar 44% dengan nilai akurasi sebesar 82% dengan algoritma yang digunakan adalah Naïve Bayes Algorithm [4]. Penggunaan data review dari google playstore juga digunakan oleh aplikasi Games yang penelitiannya dilakukan pada Games Candy Crush Saga dan Clashof the Clans, hasil dari penelitian menunjukkan aplikasi tersebut cukup popular dengan sejumlah *Negative Review* yang didapatkan. Dengan metode klasifikasi menggunakan polaritas sentimen dengan algoritma Logistic Regression didapatkan akurasi sebesar 81.1% [5]. Selain itu penelitian lain yang dilakukan oleh Rahman et all, terkait dengan Analisis Sentimen Tweet COVID-19 menggunakan *Word Embedding* dan Metode *Long Short-Term Memory* (LSTM) menunjukkan bahwa model LSTM menghasilkan akurasi yang lebih baik di bandingkan dengan model algorithma konvensional hal ini dilihat dari hasil akurasi yang di

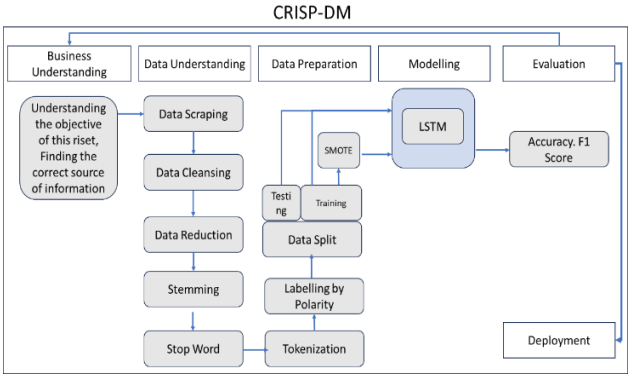
hasilkan yaitu 81% dengan selisih 8% jika di bandingkan dengan metode konvensional [6]. Berdasarkan penelitian yang telah dilakukan sebelumnya maka, penulis pada penelitian ini akan menggunakan Model LSTM yang di kombinasikan dengan SMOTE. Alasan utama penggunaan SMOTE adalah untuk mengatasi masalah ketidak seimbangan data antara 2 kelas atau lebih, sehingga data tersebut memiliki komposisi yang berimbang untuk dilanjutkan ke proses selanjutnya yaitu pengukuran akuras menggunakan model algorithma LSTM [7]. Teknik ini umum digunakan para peneliti ketika mereka mendapati ketidak seimbangan data yang cukup signifikan.

Tujuan dari penelitian ini untuk mendapatkan model dari machine learning yang sesuai dalam menghadapi masalah pada data yang memiliki kelas tidak seimbang, juga mampu memberikan performa prediksi dan klasifikasi yang bagus., bahkan saat dihadapkan pada dataset yang memiliki kelas yang tidak seimbang. Untuk mencapai tujuan tersebut, penelitian ini memiliki sasaran-sasaran berikut:

1) Melakukan pengumpulan data ulasan *google playstore* pada aplikasi digital perbankan Bank Danamon dan Bank BCA.
2) Menggunakan metode *preprocessing* dalam melakukan pengolahan data.
3) Melakukan komparasi terhadap akurasi dari hasil pelabelan menggunakan NTLK Sentimen Analyzer dengan hasil akurasi menggunakan Pelabelan secara manual.
4) Menerapkan teknik SMOTE (*over sampling*), penanganan keseimbangan data.
5) Membangun model dengan metode yang diajukan.
6) Melakukan evaluasi dari hasil yang kemudian bisa dijadikan masukan bagi pengembangan aplikasi berdasarkan ulasan dari *google playstore* tersebut.

2. Metode Penelitian

Pada penelitian ini penulis menggunakan metode CRISP-DM (*Cross-Industry Standard Process for Data Mining)* yang langkah-langkahnya telah di sesuaikan dengan tujuan dari penelitian ini. Seperti terlihat pada Gambar 2 dibawah ini:



Gambar 2. Metode CRISP-DM

Penjelasan dari Langkah-langkah ada pada gambar 2 terkait dengan yang dilakukan dalam penelitian ini yaitu:

1) *Business Understanding*

   Tahap ini adalah tahap awal dalam penelitian, mengetahui kebutuhan dan tujuan dari penelitian memberikan gambaran yang jelas mengenai proses dan hasil akhir yang di inginkan. Pada penelitian ini bertujuan untuk mendapatkan perspektif masyarakat Indonesia terhadap aplikasi digital perbankan pada dua bank besar di Indonesia yaitu D-Bank Pro dari Bank Danamon dan *Mobile Banking* dari Bank Central Asia
2) Data *Understanding*

   Pada tahap ini penulis mengumpulkan dataset mengenai ulasan dari pengguna aplikasi digital perbankan dari 2 Bank ternama yaitu Bank Danamon Indonesia dan Bank Central Asia yang diambil dari ulasan pada "*Google Playstore*" dan menganalisa fitur mana saja yang bisa digunakan untuk pemodelan ini. Pada tahap ini dilakukan beberapa langkah yaitu:

   a) Data *Scrapping*

      Proses pengambilan data dilakukan melalui "*Script Python*" dengan menggunakan compiler aplikasi "*Spyder*". Dengan objek data yang akan diambil adalah ulasan/*Review* yang diberikan oleh pengguna kedua aplikasi digital perbankan tersebut pada aplikasi "*Google Playstore*". Pada penelitian ini penulis mengambil data sebanyak 15000 baris data untuk masin masing aplikasi digital perbankan tersebut. Data set hasil scrapping ini memiliki 10 Features yaitu: "*username*, *userImage*, *content*, *score*, *thumbsUpCount*, *reviewCreated Version*, *at*, *replyContent*, *repliedAt*, *appVersion*".

b) Data Cleansing
Pada tahap ini, dilakukan proses pembersihan data teks terkait dengan teks yang mengandung *HTML Tags, Urls, mention*, *hashtag, Special Character, Digit* dan *punctuation.*

c) Data *Reduction*
Reduksi data adalah teknik yang digunakan dalam data minning dengan tujuan untuk memperkecil ukuran kumpulan data dengan tetap mempertahankan informasi terpenting dari data.

d) Data *Stemming*
Pada tahap ini, fitur yang akan di proses lebih lanjut adalah fitur "*Content*" yang merupakan uraian/review dari pengguna aplikasi digital perbankan. menjadi kata dasar. Teknik Stemming adalah teknik yang diperlukan dalam Text Minning untuk mengubah kata kembali kepada kata dasarnya. Misal" dipercayakan" akan menjadi "percaya", "berusaha" akan menjadi "usaha" dan seterusnya.

e) *StopWord*
Proses lanjutan dari *fase pre-processing* yang bertujuan untuk menghilangkan kata-kata yang sering muncul namun tidak memiliki nilai untuk dapat di olah di *Machine Learning*. salah satu contoh *StopWord* adalah kata hubungi misal "yang", dan ke dan seterusnya. Untuk menghilangkan *StopWord* diperlukan list dari kata yang di maksud sebagai *StopWord*. Pada penelitian ini penuis akan menggunakan list *StopWord* dari sastrawi.

3) Data *Preparation*
Tahapan data *preparation* adalah tahapan tahapan dalam proses persiapan data yang akan di olah, beberapa tahapan tersebut diantaranya adalah proses normalisasi yang dilakukan terhadap data, pada Text Mining teknik ini bertujuan untuk membersihkan teks dari karakter atau simbol sehingga data tersebut telah bersih dan siap di pergunakan untuk diolah menggunakan *Machine Learning/Deep Learning*. tahapan yang dilakukan pada phase ini yaitu:

a) Proses Tokenisasi adalah proses yang dilakukan dalam membagi teks yang berupa kalimat, paragraf maupun dokumen. Spasi dan tanda baca menjadi pemisah dan ciri dari token. Contoh tokenisasi yaitu: aku pergi ke pasar menjadi "Aku" "Pergi" "Ke" "Pasar".

b) *Labelling by Polarity / Manual Labeling*
Tahap ini merupakan tahap lanjutan setelah proses tokenisasi dan vektorisasi dilakukan. Proses *Labelling by Polarity* adalah teknik pelabelan pada kalimat dan kata apakah termasuk kedalam kalimat positif, negatif ataupun netral berdasarkan hasil nilai polarisasi yang di lakukan melalui metode *Lexicon*. Selain itu proses manual labeling juga dilakukan

c) Data *Split*
Tahapan ini data yang sudah diproses kemudian di pecah menjadi dua jenis kategori yaitu, data *Training* dan data *Testing*. Data *Training* adalah data yang merupakan kumpulan dataset yang telah di bentuk sebagai bahan pembelajaran oleh model yang akan digunakan untuk menemukan pola data yang tujuan akhirnya agar dapat digunakan untuk dijadikan patokan untuk data baru [8]. Sedangkan data *Testing* adalah sekumpulan dataset yang akan di pergunakan untuk menguji model yang telah di latih menggunakan data *Training*.

d) *Smote – Imbalance Data Handling*
Penelitian *Text Minning* khususnya penelitian yang menggunakan data ulasan atau *review* terkait dengan sentimen, banyak di temukan kondisi data yang tidak seimbang dalam satu kelas. SMOTE adalah salah satu teknik yang digunakan untuk menangani masalah ketidak seimbangan data [7][9].

4) *Modelling*
Tahap ini peneliti akan proses Pembangunan model dari *Machine Learning* yang kemudian di pergunakan untuk memproses data yang telah di olah sebelumnya. Pada penelitian hanya menggunakan satu model saja yaitu *Long ShortTerm Memory* (LSTM)). Sebagai model utamanya untuk kemudian akan dilakukan proses komparasi proses LSTM dengan metode SMOTE dan tanpa SMOTE.

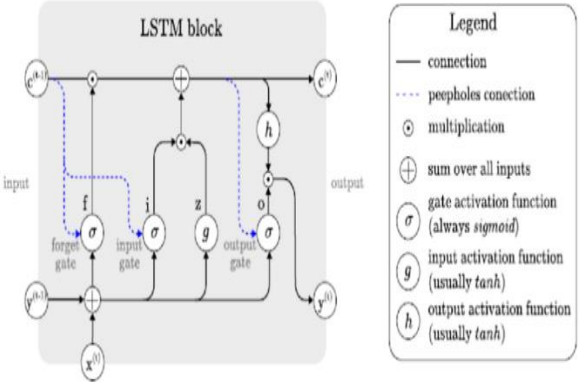
5) Model *Evaluation*
Proses penilaian kemampuan dan kinerja dari model yang digunakan untk memprediksi atau mengkalasifikasikan data yang telah di olah. Pada tahap ini pengetahuan terkait dengan seberapa baik model melakukan prediksi dan mengeneralisasi

untuk mendapatkan hasil yang akurat. Model akan melakukan pengujian untuk mendapatkan model yang menghasilkan performa terbaik. Untuk mendapatkan hasil dari performa model tesebut pada penelitian ini akan menggunakan *Accuracy* (persentase prediksi yang benar dari semua prediksi yang dilakukan oleh model), *Precision* (mengukur seberapa banyak dari prediksi positif yang sebenarnya benar), *Recall* (mengukur seberapa banyak dari kelas positif yang sebenarnya terdeteksi dengan benar oleh model) dan *F1-Score* (keseimbangan antara presisi dan *recall*)

6) *Deployment*
Merupakan tahap akhir dari model CRISP-DM, yaitu model yang telah dilakukan pengujian dan mendapatkan hasil yang terbaik akan dilanjutkan kedalam proses pembuatan aplikasi yang mengacu dan menggunakan model algoritma yang sudah diuji dan berhasil tersebut.

Long Short-Term Memory (LSTM)

Long Short- Term Memory (LSTM) adalah pengembangan dari algoritma *Recurrent Neural Network* (RNN) yang yang telah di kembangkan dengan penambahan memory cell yang memiliki fungsi menyimpan informasi atau data [10]. LSTM sesuai digunakan untuk menangani masalah terkait adanya *vanishing gradient* pada RNN yang terjadi pada saat data sequential yang banyak dan Panjang di proses [11]. Dampak dari adanya masalah vanishing gradient membuat RNN tidak berhasil mendapatkan long term *dependencies*. Isu terkait *vanishing gradient* menyebabkan menurunnya hasil dari akurasi suatu prediksi pada RNN [12].



Gambar 3. Arsitektur LSTM

3. Hasil dan Pembahasan

Business Understanding

Penelitian ini berfokus untuk mendapatkan ulasan dari masyarakan Indonesia terkait dua Aplikasi Perbankan yaitu M-BCA dan DBankPro terkait sentimen mereka terhadap kedua aplikasi tersebut.

Data Understanding

Data yang digunakan diambail dengan teknik *Scrapping* menggunakan bahasa pemrograman python dan *Spyder* sebagai aplikasi *program compiler*nya. Data yang di ambil menggunakan 15000 raw data. Dengan teknik dan parameter yang digunakan didapatkan total data ulasan untuk M-BCA sebanyak 15.000 data dan Dbank Pro sebanyak 7169. Total fitur dari data ini terdapat 11 fitur. Dan fitur yang digunakan adalah fitur *content* yang berisi ulasan/statemen dari pengguna kedua aplikasi tersebut.

Tabel 1. Ulasan pengguna M-BCA

Mobile Banking BCA	
Content	*Score*
Saya kasih bintang 1, saya TF ke BCA m-bangking saya 40.000, kenapa yang masuk cuman 4000, sisanya kemana, tolong dong di jawab!!!!!	1

Tabel 2. Ulasan Pengguna Dbank Pro

Dbank Pro Danamon	
content	*score*
Bentar2 mnta update. tiba2 ga bs login smpe bbrp hr, tlp call center sehari smpe 3x, cm suruh restart hp tp ttp ga bs, smpe install di hp berbeda.stlh komplain lumayan ngomel&nanya petugas IT br detail nanyanya& ga smpe 10 menit udh bs.kl udh diinfo dg metode yg sama tdk bs hrsnya segera follow up ga smpe brlarut2.sangat sangat sangat merugikan krn menghambat transaksi smpe bbrp hr& menghabiskn pulsa,wkt& tenaga.sangat mengecewakan	1

Pada tahap ini, peneliti juga melakukan tahap awal persiapan data, dengan tujuan untuk memudahkan proses ke tahap selanjutnya. Adapun tahapan yang dilakukan yaitu:

Data Reduction

Jumlah fitur pada dataset ini seperti dijelaskan pada poin diatas ada sekitar 11 fitur. Dari hasil *assessment* yang telah dilakukan dilakukan *data reduction* dengan teknik *feature selection*. Untuk penelitian ini hanya fitur content yang akan di proses lebih lanjut, karena berisi ulasan dari pengguna sesuai dengan objective dari penelitian ini [13].

Data Cleansing

Ulasan yang terdapat pada fitur *Content* masih berisikan kalimat, symbol dan karakter karakter yang tidak di perlukan untuk penelitian ini. Data cleansing adalah proses yang dilakukan untuk membersihkan kalimat / data dari symbol dan karakter yang tidak di perlukan [14]. Beberapa proses yang dilakukan yaitu:

1) *Text Cleansing*

Tahapan ini merupakan proses menghilangkan *Emoticon, Double Space, symbol*, angka dan termasuk didalamnya proses lowerisasi kata pada masing masing kalimat didalam dataset.

Content	score	case_folding	remove_unnecessary_char	remove_nonaplhanumeric	remove_numer
Saya kasih bintang 1,saya TF ke BCA m-bangking saya 40.000,kenapa yang masuk cuman 4000,,sisahnya kemana,,tolong dong di jawab!!!!!	1	saya kasih bintang 1,saya tf ke bca m-bangking saya 40.000,kenapa yang masuk cuman 4000,,sisahnya kemana,,tolong dong di jawab!!!!!	saya kasih bintang 1,saya tf ke bca m-bangking saya 40.000,kenapa yang masuk cuman 4000,,sisahnya kemana,,tolong dong di jawab!!!!!	saya kasih bintang 1 saya tf ke bca m bangking saya 40 000 kenapa yang masuk cuman 4000 sisahnya kemana tolong dong di jawab	saya kasih binta saya tf ke bca bangking sa kenapa ya masuk cum sisahnya kema tolong dong jawab

Gambar 4. Hasil *Text Cleansing* M-BCA

content	score	case_folding	remove_unnecessary_char	remove_nonaplhanumeric	numeric
Bank lain ada pnjman dana nya,msa ia Danamond ngga ada	1	bank lain ada pnjman dana nya,msa ia danamond ngga ada	bank lain ada pnjman dana nya,msa ia danamond ngga ada	bank lain ada pnjman dana nya msa ia danamond ngga ada	bank lain ada pnjman dana nya msa ia danamon ngga ada

Gambar 5. Hasil *Text Cleansing* DbankPro

2) Pemrosesan Kata Slang

Tahapan ini dilakukan untuk melakukan perubahan dari kata-kata yang tidak baku yang terdapat di dalam dataset menjadi kata baku yang bertujuan untuk menjaga konsistensi kata.

content	score	normalize_slang
Saya kasih Bintang 1,saya TF ke BCA m-bangking saya 40.000,kenapa yang masuk cuman 4000,,sisahnya kemana,,tolong dong di jawab!!!!!	1	saya kasih bintang saya transfer ke baca sama bangking saya kenapa yang masuk cuma sisahnya kemana tolong dong di jawab

Gambar 6. Hasil Kata Slang M-BCA

content	score	normalize_slang
Bank lain ada pnjman dana nya,msa ia Danamond ngga ada	1	bank lain ada pnjman dana nya masa ia danamond tidak ada

Gambar 7. Hasil Kata Slang DBankPro

3) Pemroresan Kata Negasi

Data yang berupa teks yang berupa ulasan atau review yang di pergunakan untuk analisa sentimen memiliki suatu kondisi yang jika tidak di proses dengan benar maka hasilnya akan menjadi tidak bagus. Contoh kata negasi yaitu kalimat positif yang diawali dengan kata 'tidak','belum' seperti 'belum bisa',' 'tidak bisa', Solusi untuk masalah ini yaitu menggabungkan kata 'tidak' tersebut dengan kalimat yang diikutinya. Contoh: 'tidak bisa' menjadi 'tidakbisa'.

content	score	Negasi
Kenapa gak bisa masuk sih gak ngerti gua	1	kenapa tidakbisa masuk sih tidakmengerti gue

Gambar 8. Hasil Negasi M-BCA

content	score	Negasi
Kenapa nggak bisa transaksi di valas ...layar memutih semua...transaksi yg lain bisa	3	kenapa enggakbisa transaksi di valas layar memutih semua transaksi yang lain bisa

Gambar 9. Hasil Negasi DBank Pro

4) *Stemming*

Proses ini adalah tahapan dalam mengubah kata yang memiliki sisipan dan imbuhan menjadi ke bentuk kata dasarnya [15]. Pada penelitian ini *Library* yang digunakan dalam tahapan ini adalah Sastrawi.

content	score	Stemming
Saya kasih bintang 1,saya TF ke BCA m-bangking saya 40.000,kenapa yang masuk cuman 4000,,sisahnya kemana,,tolong dong di jawab!!!!!	1	saya kasih bintang saya transfer ke baca sama bangking saya kenapa yang masuk cuma sisahnya mana tolong dong di jawab

Gambar 10. Hasil Stemming M-BCA

content	score	stemming
Bank lain ada pnjman dana nya,msa ia Danamond ngga ada	1	bank lain ada pnjman dana nya masa ia danamond tidakada

Gambar 11. Hasil Stemming Dbank Pro

5) *StopWord*

Tahapan ini adalah Proses yang dilakukan untuk menghapus kata-kata yang tidak memiliki makna dari dataset [16].

content	score	remove_stopword
Saya kasih bintang 1,saya TF ke BCA m-bangking saya 40.000,kenapa yang masuk cuman 4000,,sisahnya kemana,,tolong dong di jawab!!!!!	1	kasih bintang transfer baca bangking masuk sisahnya tolong

Gambar 12. Hasil StopWord M-BCA

content	score	remove_stopword
Bank lain ada pnjman dana nya,msa ia Danamond ngga ada	1	bank pnjman dana danamond tidakada

Gambar 13. Hasil StopWord DBank Pro

Data Preparation

Setelah proses normalisasi data dilakukan pada langkah sebelumnya, maka tahapan selanjutnya yaitu *Data Preparation*. Tahapan dalam *Text Minning* bertujuan untuk mempersiapkan data lebih lanjut agar dapat di proses oleh Model dari *Machine Learning* yang telah di tetapkan pada penelitian ini yaitu LSTM.

1) *Tokenization*

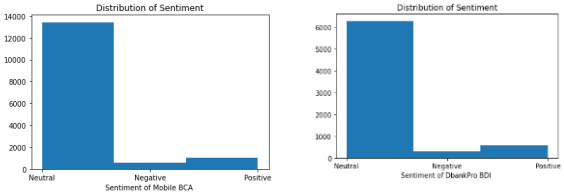
Proses memecah kalimat menjadi kata per kata. Penggunaan tokenisasi dibutuhkan saat ingin mendapatkan makna dari kata yang di ekstrak [17].

BCA Tokenize	Bank Danamon Tokenize
Review 14522 Tokens: ['kali', 'ganti', 'device', 'blokir', 'data', 'cabang', 'reset', 'mbanking', 'internet', 'banking']	Review 6715 Tokens: ['halo', 'selamat', 'bank', 'pro', 'bank', 'lowong', 'kerja', 'pt', 'bank', 'mandiri', 'cabang', 'olahraga', 'atletik', 'lompat', 'gaya', 'jongkok', 'main', 'game', 'pc', 'gratis', 'fitur', 'lowong', 'kerja', 'si', 'si', 'si', 'program', 'co', 'op', 'naung', 'pt', 'telkomsel', 'flash', 'unlimited', 'uang', 'online', 'kolanfee', 'hijab', 'tutorial', 'segi', 'ekonomi', 'pesan', 'inbox', 'scv', 'musik']
Review 14520 Tokens: ['canggih', 'aneh', 'kartu', 'gue', 'pegang', 'ganti', 'device', 'doang', 'verifikasi', 'wajah', 'susah', 'malas', 'pakai', 'gue', 'buang', 'kartu']	Review 6716 Tokens: ['sandi', 'lupa', 'aplikasi', 'lantar', 'sandi', 'solusi', 'ya']
Review 14524 Tokens: ['bos', 'habis', 'kirim', 'aktifikasi', 'belumdapat', 'kode', 'otp', 'ulang', 'kali', 'pulsa', 'habis']	Review 6718 Tokens: ['ganti', 'nomor', 'telfon', 'banking', 'nomor', 'telfon', 'banking', 'blokir', 'tidakbisa', 'uang', 'tidakbisa', 'terima', 'kode', 'verifikasi']
Review 14526 Tokens: ['kesini', 'tidakjelas', 'bilang', 'transaksi', 'gagal', 'ulang', 'kali', 'saldo', 'potong', 'tabung', 'habis', 'tidakkaruan', 'aduh']	Review 6719 Tokens: ['aneh', 'akun', 'tidakbisa', 'transfer', 'lihat', 'detail', 'transaksi']
Review 14528 Tokens: ['x', 'isi', 'pulsa', 'telkomsel', 'tidakmasuk', 'saldo', 'ribu', 'tolong', 'tanggung', 'terima', 'kasih']	Review 6732 Tokens: ['aplikasi', 'error', 'kontek', 'rekan', 'tidakbisa'] Review 6733 Tokens: ['tolong', 'bantu', 'buka', 'kunci', 'aplikasi', 'dbank', 'pro']

Gambar 14. Hasil Tokenisasi

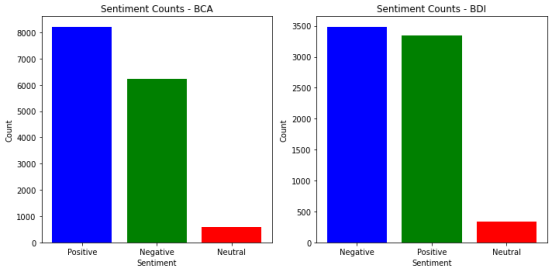
2) *Labelling By Polarity dan Manual Labelling*

Pada tahap proses Labelling by Polarity, merupaka proses pemberian label sentimen pada ulasan dengan melihat hasil dari dijalankannya library NLTK Sentimen Analyrzer dengan parameter yang telah ditentukan, yaitu jika hasil dari NLTK adalah = - 0,01 – 0,01 maka label sentimen yang diberikan adalah Netral, Jika <- 0.01 maka label sentiment yang diberikan adalah negative dan jika >0,01 maka label sentimen yang diberikan adalah positif. Untuk kemudian label tersebut kita masukkan kedalam dataset.



Gambar 15. *Labelling By Polarity*

Dari hasil *Labeling by Polarity* terlihat bahwa terjadi gap yang cukup besar antara sentimen netral dengan sentiment positif dan negatif. Dengan data seperti ini maka kondisi data tidak bisa digunakan untuk di proses lebih lanjut. Langkah selanjutnya adalah menggunakan teknik manual labelling dengan melihat kalimat per kalimat lalu memberikan label pada kalimat tersebut.

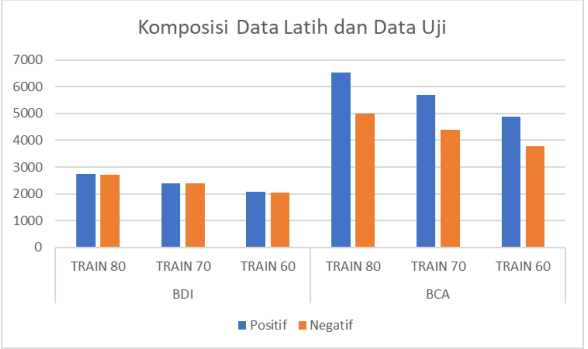


Gambar 16. Hasil *Manual Labelling*

Dari Hasil Pelabelan Manual pada gambar 4 diatas terlihat bahwa data sentiment lebih baik dibandingkan dengan data *by polarity*, oleh karena itu maka data label yang menggunakan manual yang akan digunakan untuk di proses lebih lanjut.

3) *Data Split*

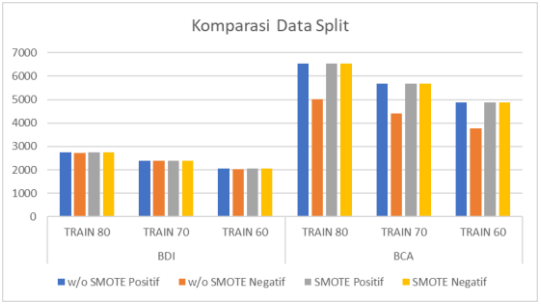
Tahapan ini dataset yang sudah diolah kemudian dibagi menjadi 2. Yaitu data latih dan data uji yang kemudian akan di terapkan ke model-model pada penelitian ini. Data Latih digunakan sebagai data pelatihan untuk melatih seluruh model dengan berbagai metode dan juga *Machine Learning* [18].



Gambar 17. Komposisi Data Training dan Uji

4) *SMOTE*

Proses SMOTE di perlukan jika pada pemrosesan data ditemukan adanya ketidak seimbangan yang cukup signifikan pada kelas yang akan di jadikan data latih pada penelitian ini. Namun untuk dataset yang bersifat string harus di lakukan vektorisasi terlebih dahulu, pada penelitian ini yang digunakan adalah TF-IDF Vectorizer.



Gambar 18. Data setelah SMOTE

Pada gambar 6, Proses SMOTE terlihat berhasil di aplikasikan untuk data minoritas.

Modelling

Proses pemodelan ini merupakan tahap eksperimen untuk membangun model *machine learning* dengan tujuan memperoleh hasil analisis sentimen yang optimal. Dalam penelitian ini, algoritma yang digunakan adalah *Long Short-Term Memory* (LSTM), yang telah terbukti efektif dalam menangani data sekuensial seperti ulasan pengguna. Beberapa teknik diterapkan untuk mendapatkan hasil terbaik, termasuk penggunaan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) untuk menangani ketidakseimbangan kelas data. Penelitian ini membandingkan performa model LSTM dengan dan tanpa penerapan SMOTE.

Tabel 3. Hasil Modeling dengan LSTM

Ulasan	Data Split	SMOTE	NO SMOTE	DIFF
	Data Train 80	86.2	87.64	1.44
M-BCA	Data Train 70	88	89.3	1.3
	Data Train 60	89.77	91.07	1.3
	Data Train 80	86.82	86.01	-0.81
DBankPro	Data Train 70	88.23	87.94	-0.29
	Data Train 60	89.82	89.82	0

Secara keseluruhan, model *Long Short-Term Memory* (LSTM) menunjukkan tingkat akurasi yang tinggi pada kedua aplikasi yang diuji, yaitu M-BCA dan DbankPro. Dalam penelitian ini, tujuan utama adalah untuk memahami bagaimana model LSTM dapat digunakan untuk analisis sentimen terhadap ulasan pengguna aplikasi perbankan digital dengan dan tanpa penerapan teknik *Synthetic Minority Over-sampling Technique* (SMOTE).

Untuk aplikasi M-BCA, akurasi tertinggi sebesar 91.07% dicapai dengan menggunakan komposisi data latih sebesar 60% tanpa SMOTE. Hasil ini menunjukkan bahwa LSTM mampu secara efektif memanfaatkan data latih yang lebih kecil ketika data tersebut sudah relatif seimbang. Tanpa perlu penyesuaian tambahan melalui teknik SMOTE, model ini dapat memprediksi sentimen dengan tingkat akurasi yang sangat baik. Hal ini menandakan bahwa distribusi kelas pada dataset ini tidak terlalu miring, sehingga teknik SMOTE tidak diperlukan untuk meningkatkan kinerja model.

Di sisi lain, aplikasi DbankPro juga menunjukkan hasil terbaik pada komposisi data latih 60%, di mana akurasi yang dicapai adalah 89.82%, baik dengan maupun tanpa penggunaan SMOTE. Ini mengindikasikan bahwa untuk dataset ini, penyeimbangan data tidak memberikan dampak signifikan terhadap peningkatan akurasi. Mungkin saja distribusi data pada aplikasi DbankPro lebih seimbang dibandingkan dengan M-BCA, atau karakteristik data tersebut tidak terlalu dipengaruhi oleh ketidakseimbangan kelas.

Namun, perbedaan antara penggunaan SMOTE dan tanpa SMOTE pada aplikasi M-BCA terlihat lebih signifikan dalam beberapa skenario. Misalnya, pada komposisi data latih 80%, penerapan SMOTE mampu meningkatkan akurasi sebesar 1.44%. Hal ini

mengindikasikan bahwa dalam beberapa kondisi, SMOTE dapat membantu mengatasi ketidakseimbangan data yang lebih besar, meskipun pengaruhnya dapat bervariasi tergantung pada distribusi awal data. Sebaliknya, pada aplikasi DbankPro, penerapan SMOTE tidak memberikan peningkatan akurasi yang berarti dan bahkan sedikit menurunkan performa pada beberapa skenario, seperti pada data latih 80% di mana terjadi penurunan akurasi sebesar 0.81%.

Pengujian lebih lanjut terhadap metrik lainnya, seperti *Recall*, *Precision*, dan *F1-Score*, mengonfirmasi bahwa penggunaan LSTM memberikan hasil yang baik dalam mengklasifikasikan sentimen. *Recall* mengukur kemampuan model untuk mendeteksi semua instance positif secara benar, sementara *Precision* menilai seberapa banyak prediksi positif yang benar-benar akurat. *F1-Score* adalah metrik yang menyeimbangkan antara *Precision* dan *Recall*, memberikan gambaran yang lebih lengkap tentang kinerja model. Detail metrik ini dapat dilihat pada Tabel 4, yang menunjukkan bagaimana setiap model mampu mempertahankan keseimbangan antara deteksi yang tepat dan keakuratan prediksi dalam kategori sentimen positif dan negatif.

Tabel 4. Hasil Recal, Precission & F1-Score

		SMOTE				w/o SMOTE		
Ulasan	Data Split	Class	Precission	Recall	F1-Score	Precission	Recall	F1-Score
	Data Train 80	0/1	0.8/0.92	0,90/0.83	0,85/0.88	0,85/0,90	0,86/0,89	0,85/0,89
M-BCA	Data Train 70	0/1	0.82/0.93	0,91/0,86	0,86/0,89	0,87/0,91	0,88/0,90	0,87/0,91
	Data Train 60	0/1	0,8/0,92	0,90/0,83	0,85/0,88	0,89/0,92	0,90/0,92	0,90/0,92
	Data Train 80	0/1	0,89/0,85	0,86/0,88	0,87/0,86	0,88/0,84	0,85/0,87	0,87/0,85
DBankPro	Data Train 70	0/1	0,89/0,87	0,88/0,88	0,89/0,88	0,89/0,87	0,88/0,88	0,89/0,87
	Data Train 60	0/1	0,91/0,89	0,89/0,90	0,90/0,90	0,91/0,89	0,90/0,90	0,90/0,89

Model Evaluation

Proses pengujian terhadap Model LSTM dengan menggunakan teknik SMOTE over sampling dan tanpa menggunakan SMOTE dengan komposisi data latih yang berbeda pada analisa sentimen terhadap aplikasi digital perbankan yaitu Mobile Banking BCA dari Bank Central Asia dan Dbank PRO dari Bank Danamon Indonesia. Dengan proses penentuan sentimen apakah positif, negatif atau Neutral dengan menggunakan library *NTLK Sentimen Analyzer*. Penggunaan NTLK *Sentiment Analyzer* dan Pelabelan secara manual. Terbukti dengan pelabelan manual memberikan hasil yang lebih baik dibandingkan penggunaan NTLK Sentimen Analyzer. hal ini dapat di sebabkan oleh beberapa factor diantaranya:

1) *Librar*y Bahasa Indonesia pada NTLK belum cukup banyak. Sehingga belum cukup untuk memberikan gambaran yang nyata mengenai pelabelan sentimen pada yang mengggunakan Bahasa Indonesia.
2) Bahasa yang umum di gunakan pada NTLK Sentimen Analyzer dan terbukti memberikan hasil yang baik adalah Bahasa Inggri, sehingga untuk hasil yang lebih baik perlu dilakukan proses translasi terlebih dahulu sebelum di proses menggunakan NTLK Sentiment Analyzer.

Pada penelitian dengan proses pelabelan manual didapatkan ulasan sentimen positif sebanyak 7836 dan 3305 ulasan, serta ulasan negatif sebanyak 3250 dan 6011. Pada penelitian ini penulis membatasi hanya kepada sentimen positif dan sentimen negatif yang dilakukan pengujian dikarenakan kondisi label yang sudah cukup untuk dapat dilanjutkan keproses selanjutnya untuk mendapatkan hasil akurasi, presisi dan recall dan F1-Score. Dengan kondisi data yang tidak seimbang maka dilakukan proses penyeimbangan data tersebut dengan menggunakan

SMOTE dengan teknik *Oversampling* sehingga didapatkan nilai positift dan negatif. Pada penelitian ini penulis melakukan teknik pengujian dengan 3 komposisi data latih yang berbeda yaitu 80:20, 70:30 dan 60:40 untuk mendapatkan komposisi ideal dan optimal untuk model yang akan di uji. Hasil SMOTE over sampling yang dapat dilihat pada gambar 4.3. pada gambar tersebut proses yang dilakukan yaitu meningkatkan data minorituas agar berimbang dengan mayoirtas, dan terbukti berhasil.

Berdasarkan teknik yang disebutkan diatas, hasil akurasi didaptkan untuk nilai terbaik ada pada komposisi data latih 60% untuk kedua ulasan dengan nilai akurasi terbaik yaitu: 91,07% untuk M-BCA dan 89,82% untuk aplikasi DbankPro dari Bank Danamon. Rentang peningkatan yang terjadi di rentang 0 sampai dengan 1,44%. Tidak berbeda jauh dengan hasil F1-Score, Precission dan Recal. Nilai terbaik didapatkan pada komposisi data latih di 60% dengan masing masing nilai yaitu:

a) M-BCA :F1 Score = Class 0 = 0,90 dan Class 1= 0,92, Precission=Class 0 = 0,89 dan Class 1= 0,92, Recall = Class 0 = 0,90 dan Clas 1=0,92
b) DBank Pro : F1 Score = Class 0 = 0,90 dan Class 1= 0,89, Precission=Class 0 = 0,91 dan Class 1= 0,89,Recall = Class 0 = 0,90 dan Clas 1=0,90

Deployment

Pada penelitian ini proses deployment tidak dilakukan namun model yang telah di hasilkan dari penelitian ini dapat di aplikasikan dengan membuat aplikasi yang dapat membantu memilah antara sentiment positif dan negative dari ulasan pengguna aplikasi pada google playstore yang kemudian dijadikan repository yang dapat digunakan oleh developer dari pembuat aplikasi M-BCA dan DBankPro sebagai acuan perbaikan dari aplikasi tersebut.

4. Kesimpulan

Pengaplikasiaan model LSTM terbukti cukup efektif dengan salah satu kelebihan dari LSTM adalah secara model dan arsitektur sudah cukup kompleks asehingga dari hasil penggunaan metode SMOTE atau teknik penanganan data tidak seimbang tidak terlalu signifikan, hanya di rentang 0 – 1,44% peningkatan yang terjadi. Adapun hasil terbaik yang didapatkan penggunaan LSTM Tanpa SMOTE dan pada data latih 60%. Ini dapat dilihat berdasarkan nilai Accuracy pada aplikasi Mobile BCA adalah 91,07% atau 1,3%, Precision 0,09 % lebih baik dan recall 0,09% dan untuk F1-Score 0,11% lebih baik. Sedangkan pada aplikasi DBank Pro hasil yang terbaik di dapatkan pada data latih 60% dengan hasil accuracy 89%, Precission dengan nilai 0,91 %, recall 0,90% dan F1-Score 0,90%. Terjadinya perbedaan hasil bisa disebabkan oleh dataset / data pattern yang digunakan bisa berbeda beda, mengingat penelitian yang dilakukan terkait dengan Sentimenent atau ulasan masyarakat. Ulasan yang diberikan, khususnya untuk ulasan yang negatif dapat dijadikan masukan bagi pengembang untuk memperbaiki aplikasi sehingga dapat memenuhi ekspektasi dari pengguna aplikasi tersebut. Adapun beberapa masukan yang dapat di pertimbangkan yaitu:

1) Stabilitas Jaringan dari Aplikasi
2) Proses update yang tidak berjalan dengan baik, dengan adanya kondisi error setelah proses upgrade.
3) Proses OTP yang ketika terjadi error mengurangi pulsa dari pengguna.

Ulasan-ulasan negatif tersebut tentunya dapat berdampak negatif terhadap korporasi dan berdampak kepada reputasi dari korporasi tersebut. Jika dalam jangka panjang tidak segera diperbaiki akan berdampak kepada kelangsungan dari korporasi. Agar penelitian ini dapat dikembangkan lebih lanjut dan dapat memberikan manfaat untuk khalayak luas, Berikut ini saran yang diusulkan:

1) Melakukan eksperimen dengan menggunakan parameter dan teknik Class weighting dan juga melakukan ujicoba dengan menggunakan beberapa algorithma transformer sehingga dapat ditemukan metodologi dan model yang lebih baik.
2) Mengunakan *Library* KBBI (Kamus Besar Bahasa Indonesia) atau yang sudah memiliki *library* bahasa Indonesia yang lebih banyak sehingga kata perkata dapat dinilai sesuai dengan bobot dari kata tersebut.
3) Melakukan pengembangan terhadap aplikasi lanjutan yang dapat memberikan informasi ataupun indikasi terhadap kepuasan dari pengguna layanan aplikasi M-BCA dan DbankPro yang kemudian akan di olah oleh administrator/tim developer dari kedua aplikasi tersebut sebagai acuan dalam perbaikan baik itu

fitur, stabilitas dan juga tampilan dari kedia aplikasi digital perbankan tersebut. Sehingga aplikasi perbankan digital M-BCA dan DbankPro dapat memenuhi ekspektasi dari penggunanya.

5. Daftar Pustaka

[1] Fauziyyah, A. K. (2020). Analisis sentimen pandemi Covid19 pada streaming Twitter dengan text mining Python. *Jurnal Ilmiah SINUS*, *18*(2), 31-42. DOI: http://dx.doi.org/10.30646/sinus.v18i2.491.

[2] Kusnawi, K., & Rahardi, M. (2023). Sentiment Analysis of Neobank Digital Banking using Support Vector Machine Algorithm in Indonesia. *JOIV: International Journal on Informatics Visualization*, *7*(2), 377-383. DOI: http://dx.doi.org/10.30630/joiv.7.2.1652.

[3] Sari, W. F., Rahim, R., & Adrianto, F. (2023). Analisis Sentiment Review Pengguna BCA Mobile Menggunakan Teks Mining. *Cakrawala Repositori IMWI*, *6*(2), 981-987. DOI: https://doi.org/10.52851/cakrawala.v6i2.295.

[4] Gupta, A., & Kamthania, D. (2021, April). Study of Sentiment on Google Play Store Applications. In *Proceedings of the International Conference on Innovative Computing & Communication (ICICC)*.

[5] Rahman, M. Z., Sari, Y. A., & Yudistira, N. (2021). Analisis Sentimen Tweet COVID-19 menggunakan Word Embedding dan Metode Long Short-Term Memory (LSTM). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, *5*(11), 5120-5127.

[6] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321-357. DOI: https://doi.org/10.1613/jair.953.

[7] Ulfa, M. A., Irmawati, B., & Husodo, A. Y. (2018). Twitter Sentiment Analysis using Na¨ ive Bayes Classifier with Mutual Information Feature Selection. *Journal of Computer Science and Informatics Engineering (J-Cosine)*, *2*(2), 106-111.

[8] Suhaeni, C., & Yong, H. S. (2023). Mitigating Class Imbalance in Sentiment Analysis through GPT-3-Generated Synthetic Sentences. *Applied Sciences*, *13*(17), 9766. DOI: https://doi.org/10.3390/app13179766.

[9] Sudriani, Y., Ridwansyah, I., & A Rustini, H. (2019, July). Long short term memory (LSTM) recurrent neural network (RNN) for discharge level prediction and forecast in Cimandiri river, Indonesia. In *IOP Conference series: Earth and environmental science* (Vol. 299, p. 012037). IOP Publishing. DOI 10.1088/1755-1315/299/1/012037.

[10] Witantoa, K. S., ERa, N. A. S., Karyawatia, A. E., Arya, I. G. A. G., Kadyanana, I., & Astutia, L. G. (2022). Implementasi LSTM pada Analisis Sentimen Review Film Menggunakan Adam dan RMSprop Optimizer. *Jurnal Elektronik Ilmu Komputer Udayana p-ISSN*, *2301*, 5373.

[11] Cahyadi, R., Damayanti, A., & Aryadani, D. (2020). Recurrent neural network (rnn) dengan long short term memory (lstm) untuk analisis sentimen data instagram. *JIKO (Jurnal Informatika dan Komputer)*, *5*(1), 1-9. DOI: http://dx.doi.org/10.26798/jiko.v5i1.407.

[12] Yan, H., Ma, M., Wu, Y., Fan, H., & Dong, C. (2022). Overview and analysis of the text mining applications in the construction industry. *Heliyon*, *8*(12). DOI:https://doi.org/10.1016/j.heliyon.2022.e12088.

[13] Widowati, T. T., & Sadikin, M. (2020). Analisis Sentimen Twitter terhadap Tokoh Publik dengan Algoritma Naive Bayes dan Support Vector Machine. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, *11*(2), 626-636. DOI: https://doi.org/10.24176/simet.v11i2.4568.

[14] Permana, Y., & Emarilis, A. (2021, March). Stemming analysis indonesian language news text with Porter algorithm. In *Journal of Physics: Conference Series* (Vol. 1845, No. 1, p. 012019). IOP Publishing. DOI 10.1088/1742-

6596/1845/1/012019.

[15] Sarica, S., & Luo, J. (2021). Stopwords in technical language processing. *Plos one*, *16*(8), e0254937. DOI: https://doi.org/10.1371/journal.pone.0254937.

[16] Rojas-Barahona, L. M. (2016). Deep learning for sentiment analysis. *Language and Linguistics Compass*, *10*(12), 701-719. DOI: https://doi.org/10.1111/lnc3.12228.

[17] Sailasya, G., & Kumari, G. L. A. (2021). Analyzing the performance of stroke prediction using ML classification algorithms. *International Journal of Advanced Computer Science and Applications*, *12*(6).

[18] Halim, A. M., Dwifebri, M., & Nhita, F. (2023). Handling Imbalanced Data Sets Using SMOTE and ADASYN to Improve Classification Performance of Ecoli Data Sets. *Building of Informatics, Technology and Science (BITS)*, *5*(1), 246-253. DOI: https://doi.org/10.47065/bits.v5i1.3647.