




BAB III

35462

-  Jurnal Repositor
-  Teknik Informatika
-  University of Muhammadiyah Malang

Document Details

Submission ID

trn:oid::1:3079759057

Submission Date

Nov 15, 2024, 8:35 AM GMT+7

Download Date

Nov 15, 2024, 9:31 AM GMT+7

File Name

35462.pdf

File Size

581.7 KB

10 Pages

3,519 Words

21,290 Characters




2% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text

Top Sources

- 2%  Internet sources
- 2%  Publications
- 2%  Submitted works (Student Papers)

Integrity Flags




0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Top Sources

- 2%  Internet sources
 - 2%  Publications
 - 2%  Submitted works (Student Papers)
-

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	
	ojs.unikom.ac.id	2%

Analisis Sentimen Ulasan Fitur Music Aplikasi Instagram Pada Google Play Store Menggunakan Metode Convolutional Neural Network (CNN)

Arolina Rinardi*¹, Christian Sri Kusuma Aditya S.Kom.,M.Kom², Didih Rizki Chandranegara S.Kom.,M.Kom³

¹Univerisitas Muhammadiyah Malang, ²Univerisitas Muhammadiyah Malang, ³Univerisitas Muhammadiyah Malang

e-mail: arolina0904@webmail.umm.ac.id*¹

Abstrak

Aplikasi Instagram merupakan sebuah medsos (media social) untuk berkomunikasi dari berbagai kalangan yang memiliki banyak fitur, khususnya fitur music. Pengguna berhak untuk menuangkan kritik ataupun saran dalam bentuk ulasan teks berkaitan cara kerja aplikasi Instagram pada Google Play Store. Dari masalah tersebut dapat dirumuskan masalah mengidentifikasi ulasan tentang fitur music Aplikasi Instagram yang mengandung sentiment negatif dan positif yang diambil pada Google Play Store dengan Teknik web screpping. Ulasan yang diambil untuk menjadi sebuah dataset sejumlah 2.260. Dataset dikumpulkan dan dilakukan pemberian label secara manual sejumlah 2.042 data negatif dan 218 data positif. Selain itu, penelitian ini bertujuan untuk mengukur performa model dengan menggunakan metode Convolutional Neural Network (CNN). Mengukur Tingkat performa model dengan menggunakan 2 percobaan yang masing-masing memiliki 2 skenario model. Hal yang membedakan kedua percobaan terletak pada hyperparameter yang digunakan pada model. Percobaan pertama scenario pertama memiliki nilai filters 32,16,8 dan scenario kedua 256,128,32 menggunakan nilai 32 sebagai nilai random state (K-Fold), 50 epoch dan 20 batch size. Percobaan kedua scenario pertama memiliki nilai filters 64,32,16 dan 128,64,32 untuk nilai filters scenario kedua, nilai 42 random state (K-Fold), 80 epoch dan 32 batch size. Hasil percobaan pertama model pertama dan kedua yaitu 93% dan 92%, sedangkan percobaan kedua model pertama dan kedua yaitu 97% dan 95%.

Kata kunci: Analisis Sentiment, Aplikasi Instagram, Google Play Store, Fitur Music, CNN

Abstract

Instagram application is a social media (social media) to communicate from various circles that has many features, especially music features. Users have the right to pour criticism or suggestions in the form of text reviews regarding how the Instagram application works on the Google Play Store. From this problem, the problem of identifying reviews about the music features of the Instagram application that contain negative and positive sentiments taken on the Google Play Store with web screpping technique can be formulated. Reviews taken to be a dataset of 2,260. The dataset is collected and manually labeled with 2,042 negative data and 218 positive data. In addition, this ponder points to degree the execution of the show by utilizing the Convolutional Neural Network (CNN). Measuring the execution level of the show by utilizing 2 tests, each of which has 2 show scenarios. The thing that distinguishes the two experiments lies in the hyperparameters used in the model. The first experiment of the first scenario has filter values 32,16,8 and the second scenario 256,128,32 using the value of 32 as the value of the random state (K-Fold), 50 epochs and 20 batch sizes. The second experiment of the first scenario has a filter value of 64,32,16 and 128,64,32 for the second scenario filter value, a value of 42 random states (K-Fold), 80 epochs and 32 batch sizes. The comes about of the primary try of the primary and moment models are 93% and 92%, whereas the moment explore of the primary and moment models are 97% and 95%.

Keywords: Sentiment Analysis, Instagram App, Google Play Store, Music Feature, CNN

1. Pendahuluan

Aplikasi Instagram merupakan media social yang dapat dilakukan oleh pengguna seperti melakukan interaksi komunikasi, dan mencari informasi[1]. Aplikasi ini memiliki banyak fitur yang disediakan, khususnya fitur music yang banyak diminati oleh pengguna. Pada penggunaan fitur music ini tentunya masih banyaknya kekurangan dan pengguna berhak untuk memberikan masukan kritik ataupun saran melalui toko online yaitu Google Play Store yang berupa ulasan teks[2]. Dari permasalahan tersebut perlu dilakukan untuk penelitian untuk menganalisis sentiment pada ulasan dataset teks mengukur performa model yang digunakan. Selain itu, mengidentifikasi sentiment positif dan negatif pada ulasan tersebut.

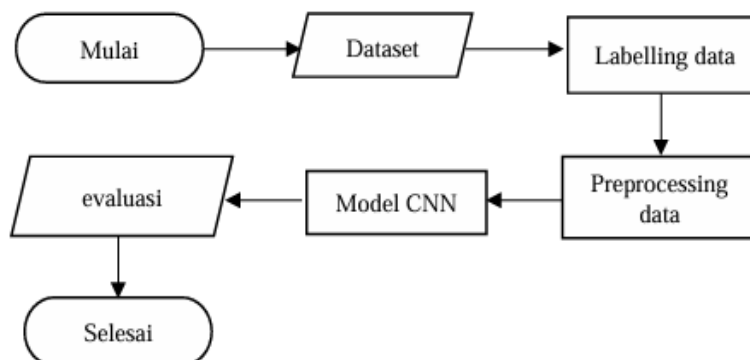
Pembahasan penelitian ini yang dilakukan dengan menggunakan metode Convolutional Neural Network (CNN) juga diimplementasikan ke pengujian yang telah dilakukan. Penelitian tentang pengembangan aplikasi berbasis web yang mengklasifikasikan ulasan pengguna terhadap aplikasi Zoom Cloud Meetings yang menggunakan metode CNN. Pada penelitian tersebut memiliki hasil akurasi sebesar 91,5% [3]. Pada penelitian lain yang membahas tentang analisis sentiment untuk mengevaluasi aplikasi E-Learning Ruang Guru dan Zenius selama pandemi, yang mana memiliki akurasi berjumlah 86%[4]. Dari penemuan penelitian sebelumnya memiliki hasil akurasi untuk penggunaan metode CNN lebih baik untuk dilakukan analisis sentiment pada ulasan dengan dataset teks.

Metode Convolutional Neural Network (CNN) merupakan algoritma untuk pengenalan pola jenis jaringan tiruan bisa dilakukan menggunakan data teks memiliki beberapa layer untuk melakukan modelnya. Metode ini biasanya dataset yang digunakan adalah dataset gambar sebagai inputannya [5]. Akan tetapi pada penelitian ini juga bisa dilakukan dataset teks sebagai inputannya. Pada pemilihan metode itu dilihat dari penelitian sebelumnya juga dapat dilakukan sebagai tugas NLP (Natural Language Processing). Disamping itu, penggunaan metode CNN juga dilakukan K-Fold Cross Validation agar hasil pengolahan data dalam model lebih akurat.

Penggunaan algoritma CNN dalam penelitian ini dilakukan untuk dilakukan 2 percobaan yang masing-masing memiliki 2 skenario model. Percobaan tersebut memiliki perbedaan terdapat pada penggunaan hyperparameter filters. Percobaan tersebut dilakukan untuk membandingkan percobaan untuk mengetahui tingkat performa yang terbaik pada setiap percobaan yang dilakukan. Perbandingan percobaan tersebut dilihat dari lama waktu eksekusi, plot loss, plot akurasi, classification report, confusion matrix yang dilakukan masing-masing percobaan.

2. Metode Penelitian

Pada metode pengujian ini yang menggunakan algoritma Convolutional Neural Network (CNN) ulasan teks yang mengandung fitur music aplikasi Instagram pada google play store. Penelitian ini dilakukan untuk mengidentifikasi sentiment negatif dan positif pada ulasan yang terkait. Selain itu juga untuk mengetahui tingkat performa yang dilakukan pada algoritma tersebut. Oleh karena itu, peneliti membuat alur pengujian sebagai berikut.



Gambar 1. Alur Pengujian

Tahapan alur pengujian dilakukan untuk proses analisis sentimen ulasan yang mengandung fitur music aplikasi instagram pada google play store. Ulasan atau dataset yang digunakan sebelumnya peneliti mengambil dengan teknik web screpping dan dilakukan labelling secara manual. Labelling dilakukan untuk membedakan sentimen negatif dan positif, lalu dilakukan

pengelohan data sesuai pada tahapan alur penelitian gambar diatas. Berikut penjelasan tahapan penelitian yang dilakukan.

2.1. Mengumpulkan data ulasan

Penelitian ini dilakukan dengan mengambil dataset dengan Teknik web screpping pada google play store. data yang diambil sejumlah 2.260 ulasan yang terkait fitur music aplikasi Instagram. Dataset yang diambil berbentuk csv agar memudahkan pengolahan data yang dilakukan. Setelah dilakukan screpping maka dataset diberi labelling pada penjelasan selanjutnya ini.

2.2. Labelling data

Pada tahapan setelah pengumpulan data maka dilakukan labelling untuk membedakan sentiment ulasan negatif dan ulasan positif secara manual. Pemberian label yang dilakukan ini agar dataset terfilter sesuai fitur music aplikasi Instagram. Pemberian label dilakukan dengan manual, karena dengan dilakukan begitu data dapat teruji kevalidan data dan dapat terlihat data setiap detailnya. Pada saat pemberian labelling setiap ulasan juga dilakukan pembersihan dataset atau dilakukan penghilangan kata yang tidak berguna, lalu selanjutnya dilakukan tahapan preprocessing lanjutan.

2.3. Preprocessing

Pada tahapan selanjutnya ini dilakukan preprocessing atau pembersihan dataset. Tahapan preprocessing ini dilakukan untuk menghilangkan noise kata atau kalimat yang tidak berguna. Pada tahap ini juga akan mempengaruhi tingkat peforma pada pengolahan data yang dilakukan selanjutnya pada model yang digunakan. Maka pada tahap ini sangat diperhatikan untuk setiap detail ulasan atau dataset yang harus bersih. Berikut tahapan preprocessing yang dilakukan pada penelitian ini.

2.3.1. Case Folding

Case Folding adalah bagian dilakukan untuk merubah seluruh kata pada dataset merubah huruf kecil atau lowercase [6].

2.3.2. Tokenize

Tokenize merupakan sebuah tahap pembersihan data yang digunakan untuk memecah kata menjadi bentuk token dan memisah kata dalam ulasan dataset[7].

2.3.3. Remove

Remove yang dilakukan pada proses preprocessing ini dilakukan 3 jenis yaitu remove hastags yang dilakukan untuk menghilangkan tanda pagar pada dataset. Remove punctuation dilakukan untuk menghilangkan tanda baca pada ulasan dataset. Remove links untuk menghilangkan kalimat pada dataset yang mengandung tautan atau links[8].

2.3.4. Stopword Removal

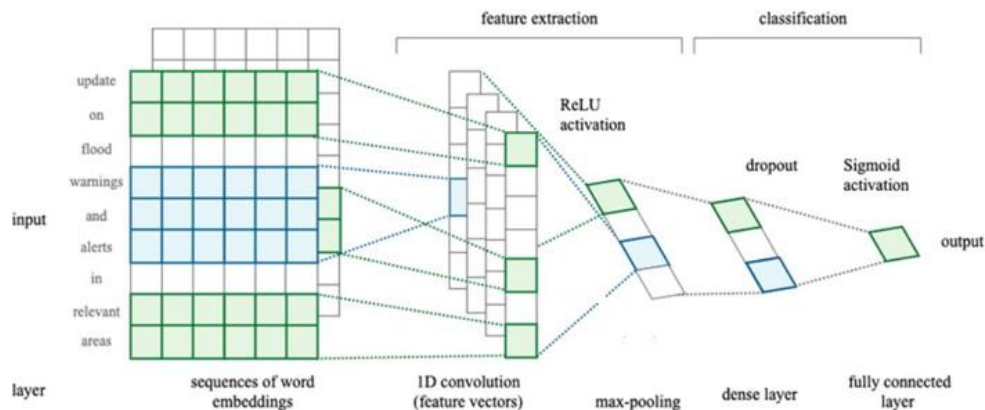
Stopword Removal yang dilakukan pada penelitian ini menggunakan kamus Bahasa Indonesia dan menghilangkan kata tidak memiliki arti pada ulasan dataset[9].

2.3.5. Lemmatization

Lemmatization merupakan tahap preprocessing yang dilakukan penelitian ini untuk mengubah kata menjadi bentuk dasar pada ulasan yang terkait[10].

2.4. Convolutional Neural Network (CNN)

pada tahap preprocessing selanjutnya dilakukan pemodelan dengan menggunakan algoritma Convolutional Neural Network (CNN). Algoritma ini merupakan sebuah deep learning yang memiliki beberapa layer antara lain embbeding,convolutional,dropout,pooling dan dense. Beberapa layer yang digunakan ini untuk memaksimalkan performa yang pada penggunaan model yang dijalankan. Pada pengujian ini juga dilakukan proses K-fold Cross Validation untuk membantu keakuratan hasil perfoma dari algoritma CNN. Nilai k yang digunakan adalah 5 serta 0,0001 sebagai angka learning rate. Berikut arsitektur Convolutional Neural Network (CNN).



Gambar 2. Arsitektur CNN[11]

2.5. Skenario percobaan

Pada skenario percobaan penelitian ini yang berisi Langkah-langkah dalam pemodelan percobaan dengan menggunakan Convolutional Neural Network (CNN). Percobaan dilakukan oleh pengujian ini sebanyak 2 kali masing-masingnya memiliki 2 skenario model yang diuji. Layer model yang di gunakan antara lain Embedding, Conv1D, MaxPooling1D, GlobalMaxPooling1D, Dense, dan Dropout. Perbedaan kedua percobaan terletak pada penggunaan hyperparameter Conv1D filters. Percobaan pertama skenario model pertama menggunakan filters 32,16,8. Percobaan pertama skenario model kedua menggunakan filters 256,128,32. Percobaan kedua skenario model pertama menggunakan filters 64,32,16. Percobaan kedua skenario model kedua menggunakan filters 128,64,32. Selain itu perbedaan lain juga terlihat dari penggunaan random state pada k-fold percobaan pertama 32, 50 epoch dan 20 batch size lalu 42 untuk percobaan kedua random state k-fold, 80 epoch dan 32 batch size percobaan kedua. Penggunaan optimizer adam untuk dilakukan pada percobaan pertama dan kedua skenario model pertama, dan RMSprop dilakukan untuk percobaan kedua skenario model kedua. Pada masing-masing percobaan dilakukan perbandingan hasil yang dilihat dari lama waktu eksekusi, plot loss, plot akurasi, classification report, dan confusion matrix. Dari hasil skenario percobaan yang dilakukan mendapatkan hasil yang terbaik sejumlah 97% dengan waktu eksekusi 39,42 seconds/detik dari percobaan kedua model pertama.

2.6. Evaluasi

Pada tahap evaluasi setelah dilakukan skenario percobaan, maka penelitian ini menghitung matrix evaluasi seperti akurasi, presisi, recall, dan F1-Score untuk melihat perolehan kinerja percobaan yang dilakukan pada tahap sebelumnya. Hasil evaluasi tersebut berbentuk Classification report, serta confusion matrix untuk menunjukkan pola kesalahan klasifikasi yang dibuat oleh percobaan pertama dan kedua. Selain itu juga menunjukkan plot loss dan plot akurasi pada setiap percobaannya untuk melihat uji coba pelatihan data pada setiap k-fold nya atau percobaan oleh kedua model tersebut.

3. Hasil Penelitian dan Pembahasan

Pada bab selanjutnya ditunjukkan perolehan dari pengujian dan pembahasan yang menggunakan metode CNN sebagai berikut.

3.1. Hasil pengumpulan dan preprocessing data

Pada hasil pengumpulan dataset yang dilakukan dengan menggunakan web scraping sejumlah 2.260 data dalam bentuk csv. Data tersebut terbagi menjadi 2 sentiment yaitu negative dan positif setelah dilakukan pada tahap preprocessing. Dataset bersentiment positif sejumlah 218 dan 2.042 bersentiment negatif. Akan tetapi dataset tersebut tidak seimbang maka dilakukan Teknik smote data agar data seimbang menjadi berjumlah 2.042 untuk masing-masing sentiment. Berikut perbedaan dataset setelah dan sebelum melakukan preprocessing.

Data sebelum proses preprocessing	Data setelah proses preprocessing	sentiment
Puas dengan aplikasi ini, karena banyak fitur menarik...Khususnya fitur musiknya pun sering di perbaharui	puas dengan aplikasi ini karena banyak fitur menarik khususnya fitur musiknya sering diperbaharui	1
Musiknya juga pas sudah dinaikin besoknya dibuat music tidak tersedia.. kecewa banget??..	musiknya sudah dinaikin besiknya dibuat musik tidak tersedia kecewa banget	0

Tabel 1. Hasil preprocessing data

3.2. Hasil Evaluasi Pengujian

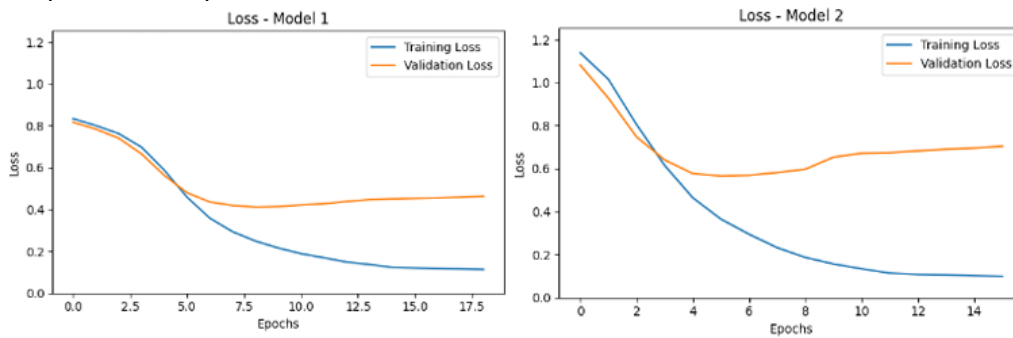
Pada penjelasan sub bab ini menjelaskan secara detail hasil evaluasi pengujian yang dilakukan oleh kedua percobaan. Selain itu, membandingkan tingkat performa yang terbaik pada model Convolutional Neural Network (CNN) yang dilakukan pengujian ini. Berikut tabel perbandingan kedua percobaan.

Percobaan	Skenario Model	Akurasi test	Loss	Filters	Random state (k-fold)	Epoch dan batch size
1	1	93%	24%	32,16,8	32	50 dan 20
	2	92%	34%	256,128,32		
2	1	97%	20%	64,32,16	42	80 dan 32
	2	95%	21%	128,64,32		

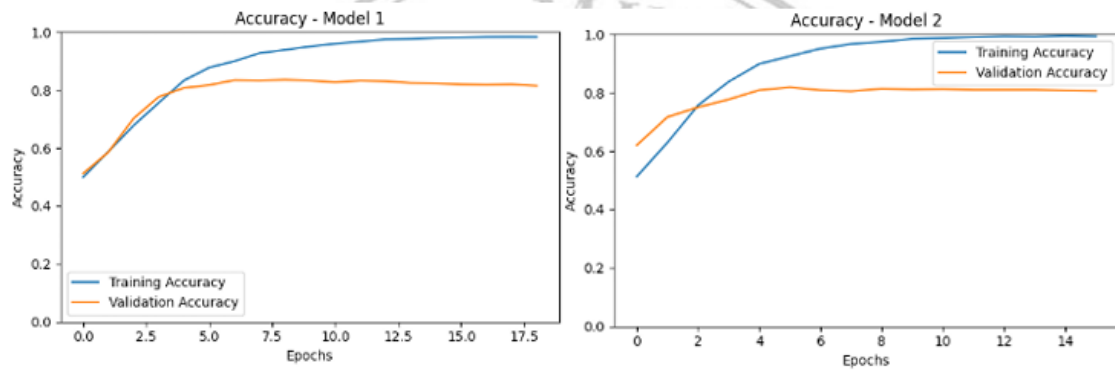
Tabel 2. Hasil perbandingan 2 percobaan

3.2.1. percobaan pertama

Pada percobaan pertama memiliki 2 skenario model, yang membedakan kedua skenario model pada percobaan ini adalah penggunaan filters model pertama menggunakan 32,16,8 dan 256,128,32 untuk model kedua. Penggunaan random state sejumlah 32, 50 epoch, dan 20 batch size. Berikut hasil pelatihan data dalam bentuk plot loss dan plot akurasi.



Gambar 3. Perbedaan plot loss percobaan pertama



Gambar 4. Perbedaan plot akurasi percobaan pertama

Pada kedua gambar 3 dan 4 plot diatas menunjukkan hasil garis yang stabil pada kedua skenario model pada percobaan pertama ini. Dikatakan stabil karena pada dasarnya garis plot loss seiring bertambahnya epoch yang menurun maka semakin baik kinerja model yang dilakukan, sedangkan garis plot akurasi akan meningkat seiring bertambahnya epoch maka kinerja model juga semakin baik. Kestabilan garis plot dipengaruhi oleh penggunaan tahap early stopping dan learning rate scheduler. Pada tahap tersebut digunakan untuk membantu menghindari terjadinya overfitting data saat pelatihan. Pada saat proses pelatihan data tersebut rata-rata waktu eksekusi yang digunakan pada percobaan pertama model pertama yaitu 44,17 seconds/detik dan untuk percobaan pertama model kedua yaitu 78,48 seconds/detik. Nilai plot loss untuk skenario model pertama 24% dan 34% untuk skenario model kedua. Hasil akurasi kedua skenario model mendapatkan hasil 93% dan 92%. Berikut hasil percobaan pertama dalam bentuk classification report.

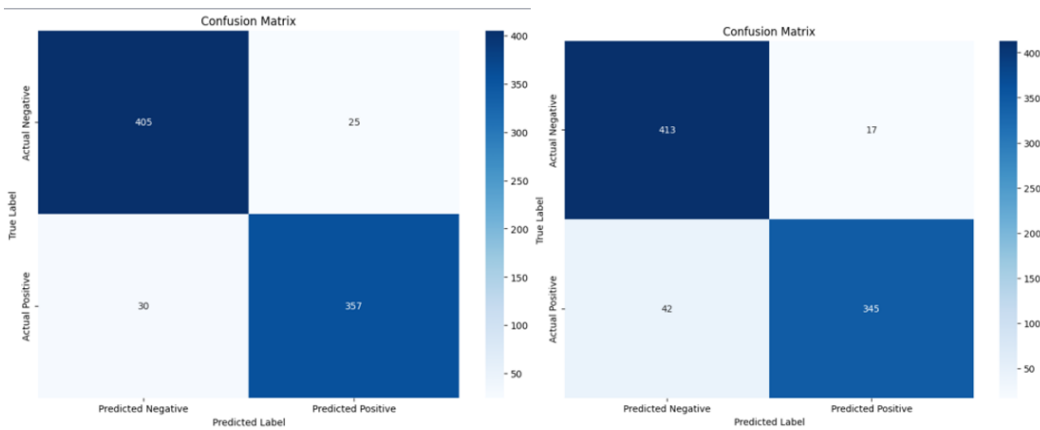
	precision	recall	F1-score	support
0	0,93	0,94	0,94	430
1	0,93	0,92	0,93	387
Akurasi			0,93	817

Tabel 3. Classification Report percobaan pertama model pertama

	precision	recall	F1-score	support
0	0,91	0,96	0,93	430
1	0,95	0,89	0,92	387
Akurasi			0,92	817

Tabel 4. Classification report percobaan pertama model kedua

Hasil tabel classification report percobaan pertama diatas menunjukkan hasil tersebut sudah cukup baik. Hasil akurasi percobaan tersebut 93% dan 92% untuk kedua skenario percobaan. hak tersebut dipengaruhi oleh tingkatan kebersihan data pada saat proses menjalankana tahap preprocessing. selain itu juga penggunaan K-fold cross validation sebanyak 5 untuk nilai k yang digunakan. Berikut hasil percobaan menggunakan confusion matrix.

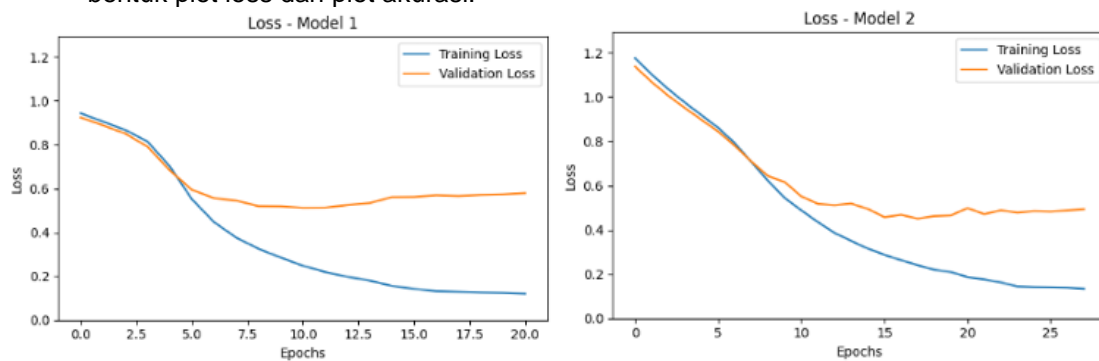


Gambar 5. Confusion Matrix Percobaan pertama

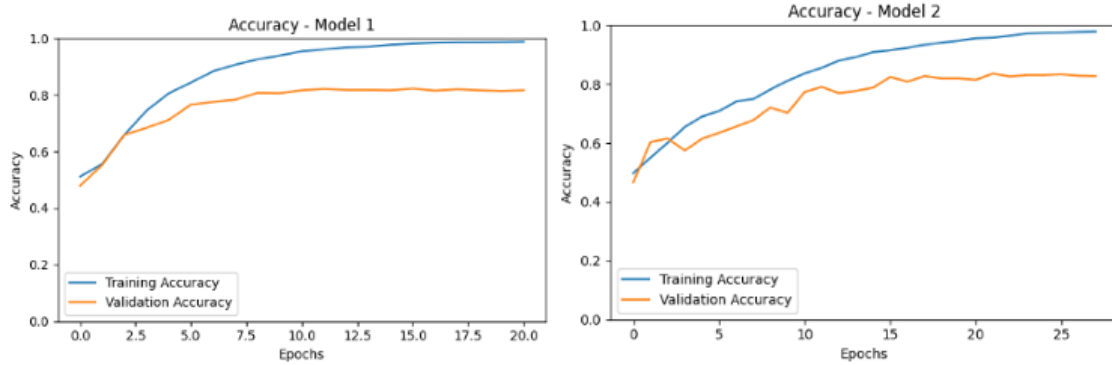
Pada hasil confusion matrix percobaan pertama yang ditunjukkan gambar di atas. Mengetahui kinerja model percobaan peratam untuk melihat pola kesalahan klasifikasi pada masing-masing kelas. Terdapat nilai True Negative (TN) sejumlah 405, artinya model memprediksi 405 ulasan negatif dengan benar. Nilai False Positives (FP) berjumlah 25, artinya model salah memprediksi 25 ulasan negatif sebagai positif. Nilai False Negatives (FN) sejumlah 30, artinya model salah memprediksi sejumlah 30 ulasan positif sebagai negatif. Nilai True Positives (TP) sejumlah 357, artinya model memprediksi 357 ulasan dengan benar. Sedangkan, confusion matrix pada model kedua memiliki nilai True Negative (TN) sejumlah 413, artinya model memprediksi 413 ulasan negatif dengan benar. Nilai False Positives (FP) dengan nilai 17, artinya model salah memprediksi 17 ulasan negatif sebagai positif.

3.2.2. Percobaan kedua

Pada percobaan kedua penelitian ini memiliki masing-masing 2 skenario model yang di uji. Perbedaan pada kedua skenario model pada penggunaan filters, skenario model pertama menggunakan filters 64,32,16 dan 128,64,32 untuk model kedua, jumlah epoch yang digunakan 80, 32 untuk ukuran batch, dan 42 untuk penggunaan random state k-fold cross validation. Penggunaan optimizer adam untuk skenario model pertama dan RMSprop untuk skenario model kedua. Berikut hasil pelatihan data dalam bentuk plot loss dan plot akurasi.



Gambar 6. Perbedaan plot loss percobaan kedua



Gambar 7. Perbedaan plot akurasi percobaan kedua

Pada gambar 6 dan 7 diatas dilakukan pelatihan data pada percobaan model pertama dan kedua yang memiliki sumbu y yang sama pada masing-masing model. Pada pelatihan gambar tersebut memerlukan waktu eksekusi untuk percobaan kedua model pertama yaitu 39,42 seconds/detik sedangkan percobaan kedua model kedua yaitu 66,79 seconds/detik. Dilihat dari waktu eksekusi tersebut model kedua percobaan ini lebih lama dibandingkan dengan model 2. Nilai loss validation model 1 dan 2 memiliki 20% dan 21%. Jika dilihat dari proses pelatihan percobaan kedua, maka percobaan kedua model pertama yang sudah mencukupi. Pada percobaan ini juga melakukan tahap seperti percobaan 2 yaitu melakukan early stopping dan learning rate scheduler yang sama, karena hal ini bisa melakukan monitoring untuk setiap langkah yang percobaan oleh peningkatan epoch. Setelah langkah tersebut maka dilakukan evaluasi yang ditunjukkan pada classification report percobaan kedua untuk model pertama dan kedua sebagai berikut.

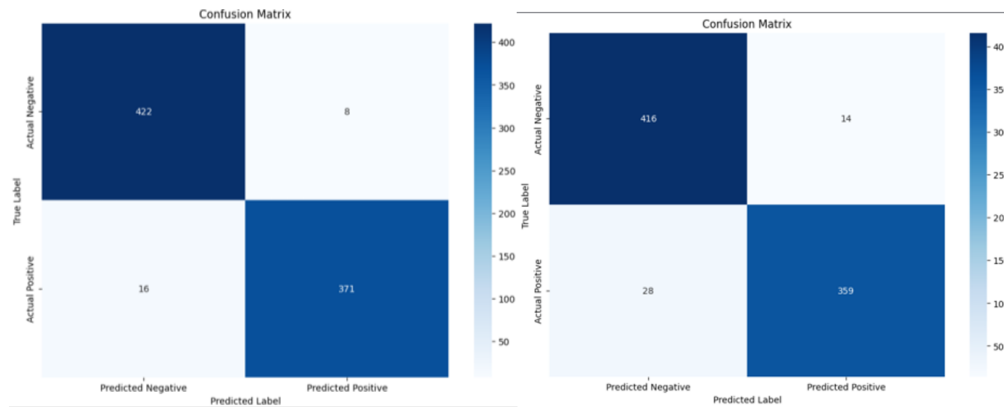
	precision	recall	F1-score	support
0	0,96	0,98	0,97	430
1	0,98	0,96	0,97	387
Akurasi			0,97	817

Tabel 5. Classification Report percobaan kedua model pertama

	precision	recall	F1-score	support
0	0,94	0,97	0,95	430
1	0,96	0,93	0,94	387
Akurasi			0,95	817

Tabel 6. Classification Report percobaan kedua model kedua

Pada tabel 5 dan 6 diatas menunjukkan bahwa hasil nilai akurasi percobaan kedua ini lebih baik dibandingkan dengan percobaan sebelumnya. Percobaan kedua ini memiliki nilai akurasi model pertama sejumlah 97% dan nilai loss sejumlah 20%. Model kedua memiliki jumlah akurasi 95% dan nilai loss sejumlah 21%. Selisih akurasi dari kedua model percobaan kedua ini adalah 2%. Dari proses pengolahan data yang dihasilkan percobaan kedua model pertama lebih baik atau mencukupi dibanding dengan percobaan pertama, dilihat dari hasil akurasi dan lamanya waktu eksekusi percobaan kedua model pertama yang paling cepat. Dapat di ketahui bahwa jika nilai akurasi semakin meningkat artinya performa yang dimiliki semakin baik, sedangkan untuk nilai loss semakin nilainya turun maka performa semakin baik. Performa percobaan kedua dikatakan lebih baik dari percobaan sebelumnya karena penggunaan nilai hyperparameter pada random state pada k-fold sejumlah 42 , nilai epoch sebanyak 80, dan 32 untuk ukuran batch yang digunakan.



Gambar 8. Classification Report percobaan kedua

pada gambar 8, perpecahan matrix dari percobaan kedua di atas. Dalam model pertama, nilai TP (True Positive) adalah 371, yang menunjukkan bahwa jumlah sampel positif yang diprediksi dengan benar adalah positif. Nilai FP (False Positive) adalah 8, yang menunjukkan bahwa jumlah sampel negatif yang diprediksi dengan benar sebagai positif. Nilai FN (False Negative) adalah 16, yang menunjukkan bahwa jumlah sampel positif yang diprediksi dengan benar sebagai negatif. Nilai TN (True Negative) adalah 422, menunjukkan bahwa jumlah sampel negatif yang di Nilai TN untuk 416 sampel negatif yang diprediksi benar, nilai FP untuk 14 sampel negatif yang diprediksi salah sebagai positif, dan nilai FN untuk 28 sampel positif yang diprediksi sebagai negatif. Total nilai TP adalah 416.

4. Kesimpulan

Pada Kesimpulan penelitian “Analisis Sentiment Ulasan fitur Musik Aplikasi Instagram Pada Google Play Store dengan menggunakan metode Convolutional Neural Network”. Penelitian dengan menggunakan dataset teks berjumlah masing-masing 2.042 setelah dilakukan Teknik smote atau imbalance data. Dengan menggunakan 2 percobaan yang masing-masing memiliki 2 skenario model. Hasil dari kedua percobaan, untuk percobaan pertama model pertama dan kedua nilai akurasi yaitu 93% dan 92%, memiliki selisih 1%. Nilai loss percobaan pertama ini 24% dan 34% dengan lama waktu eksekusinya 44,17 seconds/detik dan 78,48 seconds/detik. Percobaan kedua model pertama dan kedua memiliki nilai akurasi 97% dan 95%, memiliki selisih 2%. Nilai loss percobaan kedua ini 20% dan 21% dengan waktu eksekusi yaitu 39,42 seconds/detik dan 66,79 seconds/detik. Hasil perbandingan tersebut yang mencukupi yaitu penggunaan percobaan kedua model pertama karena memiliki akurasi yang tinggi dan waktu eksekusinya yang cepat.

Refrensi

Jurnal:

- [1] R. M. Candra and A. Nanda Rozana, “Klasifikasi Komentar Bullying pada Instagram Menggunakan Metode K-Nearest Neighbor,” *IT J. Res. Dev.*, vol. 5, no. 1, pp. 45–52, 2020, doi: 10.25299/itjrd.2020.vol5(1).4962.
- [2] A. Faadilah, “Analisis Sentimen Pada Ulasan Aplikasi Tokopedia di Google Play Store Menggunakan Metode Long Short Term Memory,” pp. 1–46, 2020.
- [3] R. Refianti and N. Anggraeni, “Sentiment Analysis Using Convolutional Neural Network Method to Classify Reviews on Zoom Cloud Meetings Application Based on Reviews on Google Playstore,” *Int. J. Eng. Sci. Inf. Technol.*, vol. 3, no. 3, pp. 7–16, 2023, doi: 10.52088/ijesty.v3i3.463.
- [4] G. A. Ramadhan, *Analisis sentimen ulasan aplikasi ruangguru dengan algoritma long short term memory*. 2023. [Online]. Available: [https://repository.uinjkt.ac.id/dspace/handle/123456789/71510%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/71510/1/GILANG AMBANG RAMADHAN-FST.pdf](https://repository.uinjkt.ac.id/dspace/handle/123456789/71510%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/71510/1/GILANG%20AMBANG%20RAMADHAN-FST.pdf)
- [5] I. Ayu Shafirra N, “Klasifikasi Sentimen Ulasan Film Indonesia dengan Konversi Speech-to-Text (STT) Menggunakan CNN,” *J. sains dan seni ITS*, vol. 9, no. 1, pp. 2301–9271, 2020.
- [6] R. Naquitasia, D. H. Fudholi, and L. Iswari, “Analisis Sentimen Berbasis Aspek pada

- 1 Wisata Halal dengan Metode Deep Learning,” *J. Teknoinfo*, vol. 16, no. 2, p. 156, 2022, doi: 10.33365/jti.v16i2.1516.
- [7] M. R. Nadhif, D. Wisnu Brata, and B. Rahayudi, “Analisis Sentimen Data Ulasan Pengguna Aplikasi TIX ID di Indonesia pada Google Play Store menggunakan Support Vector Machine,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 8, pp. 3932–3937, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [8] S. M. Fani, R. Santoso, and S. Suparti, “Penerapan Text Mining Untuk Melakukan Clustering Data Tweet Akun Blibli Pada Media Sosial Twitter Menggunakan K-Means Clustering,” *J. Gaussian*, vol. 10, no. 4, pp. 583–593, 2021, doi: 10.14710/j.gauss.v10i4.30409.
- [9] S. H. Badjrie, O. N. Pratiwi, and H. D. Anggana, “Analisis Sentimen Review Customer Terhadap Produk Indihome Dan First Media Menggunakan Algoritma Convolutional Neural Network Review Analysis Sentiment Customer Product Indihome And First Media Using Convolutional Neural Network,” *eProceedings ...*, vol. 8, no. 5, pp. 9049–9061, 2021, [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/15877>
- [10] R. Sovia, A. Syaifullah, F. Yenila, and R. Permana, “Penerapan Natural Language Processing Pada Sistem Chatbot Sebagai Helpdesk Obyek Wisata Menggunakan Metode Naïve Bayes,” *J. Infortech*, vol. 5, no. 2, pp. 210–218, 2023.
- [11] L. Styve, C. Navarra, J. M. Petersen, T. S. Neset, and K. Vrotsou, “A Visual Analytics Pipeline for the Identification and Exploration of Extreme Weather Events from Social Media Data,” *Climate*, vol. 10, no. 11, pp. 1–23, 2022, doi: 10.3390/cli10110174.