ISSN 2085-4579

# Aspect-Based Sentiment Analysis on Application Review using CNN

(Case Study : Peduli Lindungi Application)

Putri Arta Aritonang[1], Monika Evelin Johan[2], Iwan Prasetiawan[3]
[1,2,3] Information System Department, Faculty of Information and Technology, Universitas Multimedia Nusantara, Tangerang, Indonesia
[1] putri.arta@student.umn.ac.id
[2] monika.evelin@umn.ac.id
[3] iwan.prasetiawan@lecturer.umn.ac.id

Accepted 1 July 2022
Approved 28 July 2022

***Abstract*—As an obligatory application during the COVID-19 pandemic by Indonesians, PeduliLindungi must have provided outstanding quality services to its users. However, as of December 2021, users' sentiment toward the quality and service of the PeduliLindungi application was still low, with an application rating of 3.6 out of 5 on the Google Play Store. This study uses text mining techniques for the Aspect-Based Sentiment Analysis (ABSA) task in the PeduliLindungi application review, a sentiment analysis task based on the aspect category of the application. This study aims to classify the users' sentiment on aspects of the application and provide insight and knowledge to improve the quality of the PeduliLindungi application. The ABSA method used in this study is the classification of aspects and sentiments using the Convolutional Neural Network (CNN) algorithm. The results showed that the CNN model could produce such good performance with an f1 score of 92.23% in the aspect classification and 95.13% in the sentiment classification. The results of user sentiment modelling showed the dominance of negative sentiment in the eight aspects of the application, namely Visual Experience, Scan – Check-in/out, Vaccine Certificate, eHac, COVID Test, Register/Login, Performance and Stability, and Privacy, Data, and Security.**

***Index Terms*—Aspect-Based Sentiment Analysis, Convolution Neural Network, PeduliLindungi, Text Classification, Text Mining.**

## I. Introduction

COVID-19 (Coronavirus Disease) first identified in Wuhan, China, in December 2019, has spread throughout the world until now [1]. In identifying cases and preventing the spread of the virus, many types of mobile applications have been developed. The first COVID-19 mobile application to be developed and widely published was an application for contact tracing created to notify users if they met another person infected with COVID-19 [2]. In Indonesia, the application developed to assist government in tracking to prevent the spread of COVID-19 is the PeduliLindungi application [3].

First released on March 28, 2020, the PeduliLindungi application has a tracking function by relying on community involvement to share location data to trace contact history with COVID-19 patients, patients under supervision, and people under supervision can be carried out.

Not only as a contact tracing application, PeduliLindungi also continues to grow and has many additional features. In September 2021, responding to the policy for the Implementation of Restrictions on Community Activities, commonly known as PPKM, the PeduliLindungi application became a mandatory application for public access, according to the rules in the Instruction of the Ministry of Home Affairs Number 42 of 2021 [4]. This rule has led to an increase in the use of the PeduliLindungi application.

As of December 2021, PeduliLindungi is the number 1 application in Indonesia in the medical category on the Google Play Store. It has been downloaded by more than 50,000,000 people and has a 3.6 out of 5 rates on the Google Play Store [5]. The rating is still relatively low, considering that PeduliLindungi, as an obligatory application, should provide excellent quality and service to its users.

Various reviews, as well as good and bad, are inevitable. However, this can be used to improve the quality of the application based on the analysis results from user reviews. By knowing the sentiments of aspect reviewed by users, developers can improve the quality of the relevant aspects of the application.

Sentiment analysis extracts sentiments, opinions, or judgments on products or services [6]. Most sentiment analysis is carried out at the sentence level, so it does not provide sufficiently important information for decision-making. However, this information can be obtained by conducting sentiment analysis at the sub-sentence level or aspect level [6].

If a reviewer reviews a product, the thing being reviewed relates to the aspects that exist in the product. It does not mean the reviewers like or dislike the product as a whole, but in certain aspects. This concept sparked the Aspect-Based Sentiment Analysis (ABSA), which aims to discover people's sentiment about aspects of an entity [6]. The ABSA process is mainly done by classifying aspects and sentiments. The algorithm model used will classify the text into category aspects and then determine the sentiment [7].

Research conducted by [8] compared several deep learning algorithms in performing ABSA on hotel reviews with target classification aspects: price, hotel, room, location, service, restaurant, and sentiment classification: positive and negative. The study showed the CNN model algorithm has an accuracy of 90.4% for sentiment classification and 87.2% in aspect classification.

Researchers [9] compared the CNN algorithm model with Naïve Bayes in conducting ABSA on online marketplace reviews with target classification aspects: accuracy, quality, service, price, packaging, delivery, and sentiment classification: positive and negative. The CNN algorithm has a higher average accuracy of 91.98% for aspect classification and 93.07% for sentiment classification. No other journal sources are identical to the topic of ABSA for the Pedulilindungi application.

This research uses the CNN algorithm for the text classification task and aims to build the CNN model to classify aspects and sentiments on the PeduliLindungi application review, discover the model's performance, and compare the sentiment per an aspect of the application in versions 4.0.2 and 4.0.5.

In comparison to previous research, this research uses the Pedulilindungi application review as the research object. It will also classify unlabeled data using the CNN model built to compare the sentiments of each aspect sentiment on reviews of different application version. Both research conducted by [8] and [9] also used more general target classification aspects, such as quality, service, and price. Meanwhile, this research followed a series of aspect categories standards by Android and curated some aspects directly related to the function of the application content.

## II. Literature Review

### A. Text Mining

Text mining is a process of mining text data from an unstructured format to a structured format to identify existing patterns [10]. The main goal is to obtain and extract useful information from the text for use in further tasks. Text mining requires structuring the text used as input because it has an unstructured format. Therefore, text pre-processing must be carried out to clean and convert text into a structured format.

The pre-processing stages are divided as follows:

1) *Case Folding*

A common approach to deal with inconsistent capitalization in text is to generalize all characters by using the same letter, which is lowercase [11]. In addition, removing punctuation, numbers, extra spaces, and single characters is required to reduce noise.

2) *Tokenization*

Tokenization is breaking long sentence text into words, called tokens [10]. This process investigates each sentence and creates a list of tokens that can be used as input for the following algorithm [12]. The main objective is to investigate the words in a sentence [11].

3) *Normalization*

This process aims to normalize non-standard languages to the appropriate word in the KBBI.

4) *Filtering*

This process includes steps such as removing words with no information or are unnecessary (stop words). With this, the dimensionality of the text can be reduced without reducing the text content [12].

5) *Stemming*

This process aims to search for stem words by transforming words that have affixes or suffixes to the root words [12].

### B. Aspect-Based Sentiment Analysis

Aspect-Based Sentiment Analysis, or ABSA, is a type of sentiment analysis that aims to determine sentiment in each specified aspect [8]. ABSA processes information at the sub-sentence level or aspect level.

In several studies, the process in ABSA is divided into two tasks, namely the task of aspect extraction and estimating the polarity/rating [6]. Aspect extraction aims to extract words/aspects from product reviews and group synonyms for each aspect because each person can use different phrases that refer to the same aspect [6]. The second task is polarity estimation which aims to determine the sentiment on an aspect, whether positive, negative, or neutral [6]. In this method, aspects are extracted first and then classified as positive or negative [13].

ABSA is also carried out in other method because the aspect extraction process requires a lot of resources [8]. The ABSA process can be done by classifying aspects and sentiments. The model used will classify text documents into category aspects and sentiment tendencies [7]. For example, in the review sentence "The food price is quite high", the ABSA model will classify the sentence into price aspects and negative sentiment classes [7]. This method requires labeled text data to train the model used in ABSA.

### C. Convolutional Neural Network

Commonly used in computer vision and image processing, such as image classification and object detection, Convolutional Neural Network or CNN has been proven effective in Natural Language Processing (NLP) and has achieved good results in semantic text classification task [14].

The following is an explanation of each layer used on CNN for text classification task:

1) Embedding layers
This layer functions to map input in vocabulary indices into low-dimensional vectors [16]. The maximum sentence length determines the vocabulary size. After the words are transformed into vectors, it will be fed to the convolutional layer [15].
2) Convolutional layers
This layer is the main processing layer of the model, which carries out the convolution process for inputs and filters [9]. When the input enters this layer, a convolution operation involving a filter is applied to the word window to generate a new feature. The filter is applied repeatedly to each word window in the sentence to produce a feature map [14].
3) Pooling Layer
This layer gradually reduces the number of parameters, the computational complexity of the model, and control overfitting [16]. Max-over-time pooling is often applied to feature maps to retrieve the most important feature (feature with highest value) for each map [14].
4) Fully-connected Layer or Dense Layer
This layer forms one-dimensional neurons and consists of neurons interconnected with neurons in the previous and subsequent layers [9]. In this layer, regularization can be done with a dropout function to keeps the neurons in a probability value between 0 and 1, making it easier to classify output classes [17]. This layer will also output the specified number of classes using SoftMax activation.

## III. Methodology

### A. *Overview of Research Object*

The object of this research is user reviews on the PeduliLindungi application. PeduliLindungi is an application developed to help government agencies in tracking to prevent the spread of COVID-19 [3]. Since September 2021, PeduliLindungi has been an obligatory application for several activities and public access.

This study uses two application version reviews, namely versions 4.0.2 and 4.0.5. The selection of the application version is based on the amount of review data that is adequate for this study and adjusts the data collection period.

1) PeduliLindungi Version 4.0.2
PeduliLindungi version 4.0.2 was updated on October 19, 2021. The following is the list of the menus accessible to users in this version:
- Vaccine Certificate
- COVID-19 Test Results
- E-Hac
- Scan QR Code
- Check-in History
- Travel Regulations
- Telemedicine
- Healthcare Facility
- COVID-19 Statistics
- Get Vaccine
- Account

2) PeduliLindungi Version 4.0.5
PeduliLindungi version 4.0.5 was updated on November 19, 2021. Several things are updated in this version, which listed on the PeduliLindungi page on Google Play, including:
- Changes in the UI/UX.
- Added Chinese, Japanese, Russian, Korean, and Spanish language options.
- Improved flow of the E-Hac menu.
- Added a CAPTCHA in certificate claim.
- Added FAQ regarding zoning color status.
- Eradication of bugs (errors).

### B. *Research Flow*

The research flow used in this study is an adaptation of a research journal [7], [8] with several adjustments. The following is the steps that describes the flow of this research:

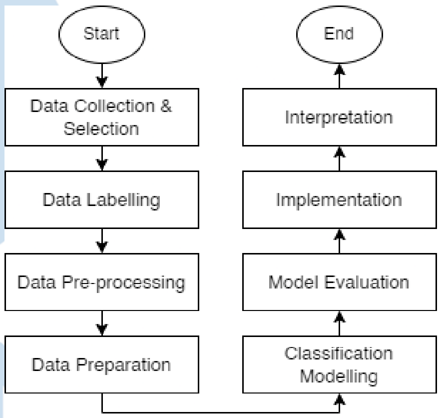



Fig. 1. Research Flow

a) Data Collection and Selection
The first step is to collect PeduliLindungi's review data from Google Play Store using the google_play_scraper library in Python, and then perform the data selection process.

b) Data Labelling
The next step is to manually label the review data for version 4.0.2 for the aspect and sentiment columns based on the review text. For aspect category, this research follows a series of aspect categories standards by Android, including Visual Experience (UI/UX), Functionality, Performance and Stability, and Privacy and Security [18]. For this research, the Functionality aspect will be expanded into features in the PeduliLindungi application. The aspect column has eight label targets, as shown in Table I. As for the sentiment, there are two

target labels, namely positive and negative sentiment, in Table II.

TABLE I. ASPECT CATEGORIES

| No | Aspect | Description |
|---|---|---|
| 1 | Visual experience (UI/UX) | The visuals of the application, user journey, content accessibility, and navigation. |
| 2 | Functionality – Scan QR, Check-in or Check-out | The function to check in or check out a place using QR scan, and location detection. |
| 3 | Functionality – Vaccine Certificate | The functions in the vaccine certificate menu. |
| 4 | Functionality – e-Hac | The functions in the e-Hac menu: to create, and view travel documents. |
| 5 | Functionality – COVID-19 Test | The functions in the COVID-19 test results menu. |
| 6 | Functionality – Register / Login | The application's ability to perform the registration and login into the application. |
| 7 | Performance and stability | The application's stability, performance (loads quickly, gives feedback to the user), and battery usage. |
| 8 | Privacy, Data, and Security | Access or permissions to support the applications, such as location permissions, accessing data, data storage, and the ability to display reliable data. |

TABLE II. SENTIMENT CATEGORIES

| No | Aspect | Description |
|---|---|---|
| 1 | Positive | The reviews contain kind words, positive emotions, and support both implicitly and explicitly. |
| 2 | Negative | The reviews contain bad words, negative emotions, and do not support either implicitly or explicitly. |

c) Data Pre-processing

At this stage, we clean the application review data in versions 4.0.2 and 4.0.5 and convert it from unstructured text data into structured ones.

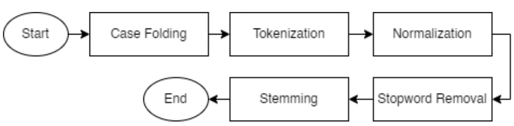



Fig. 2. Data Pre-processing Flow

d) Data Preparation

This stage is intended to prepare text data that is acceptable as the input of neural network.

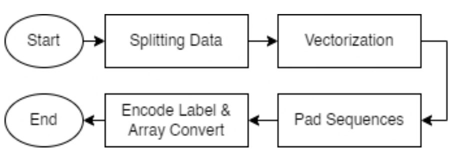



Fig. 3. Data Preparation Flow

e) Classification Modelling

Modeling is divided into two: modeling for aspect classification and sentiment classification. A separate CNN model will be created using parameters that show the best accuracy results in the hyperparameter tuning process. After initiating the model, training will be carried out using the data train.

f) Model Evaluation

We evaluate the model for its accuracy and loss using validation data at this stage. After that, we test the model using test data and evaluate the classification performance using metrics, namely accuracy, precision, recall, and F1 score.

g) Implementation

This stage aims to classify the unlabeled data (review version 4.0.5), using the CNN model for aspect classification and sentiment classification. The output is the classified aspects and sentiment of 4.0.5 review.

h) Interpretation

This stage aims to explain the results of the aspect and sentiment classification and compare the sentiments in each aspect on version 4.0.2 with version 4.0.5. The comparison is intended to determine whether the sentiment in each aspect significantly changed accordingly to the version update.

## IV. RESULTS AND DISCUSSION

### A. Data Collection and Selection

We collect the user reviews from Google Play Store using google_play_scraper. There are 2,320 reviews of version 4.0.2 and 1,031 reviews of version 4.0.5.

### B. Data Labelling

Table III shows the example of review data of version 4.0.2 with its labeled aspect and sentiment.

TABLE III. EXAMPLE OF REVIEW DATA

| Review | Aspect | Sentiment |
|---|---|---|
| Hasil ehac beda swaktu dibuat, sudah diedit tetap gak berubah padahal butuh untuk terbang. | eHac | Negative |
| Checkin lambat, gak bisa baca QR, padahal udah nyalain GPS. | Scan, Checkin/out | Negative |
| Suka skali dngn update sekarang. Simpel dan color coded. | Visual Experience | Positive |
| Kerja bagus saya sudah bisa cek sertifikat vaksin. | Vaccine Certificate | Positive |

The distribution of data labels, as seen in Table IV, shows the data is label imbalance. The aspect with the most reviews is Performance and Stability with 408 reviews, while lowest is the COVID-19 Test with only 73 reviews. Most reviews are negative sentiments with 1840 reviews and positive sentiment with 480 reviews.

TABLE IV. ASPECT AND SENTIMENT LABEL DISTRIBUTION

| Aspect | Positive | Negative | Total |
|---|---|---|---|
| Visual Experience | 107 | 281 | 388 |
| Scan, Checkin/out | 71 | 336 | 407 |
| Vaccine Certificate | 87 | 248 | 335 |
| eHac | 17 | 94 | 111 |
| COVID-19 Test | 10 | 63 | 73 |
| Register/Login | 40 | 362 | 402 |
| Performance and Stability | 105 | 303 | 408 |
| Privacy, Data, and Security | 43 | 153 | 196 |
| **Total** | 480 | 1840 | 2.320 |

## *C. Data Pre-processing*

This stage covers cleaning and converting unstructured text data into a structured format by filtering the terms of unnecessary things and normalizing them to a more uniform sequence. The result of data pre-processing is shown in Figure 4.

```
0    [tampil, dukung, buat, smartphone, rasio, layar, lama, tombol, agree, t...
1    [hasil, ehac, beda, waktu, buat, edit, kali, kali, tetap, ubah, padahal...
2    [sertifikat, vaksin, muncul, telepon, beberapa, kali, alas, sistem, sed...
3    [suka, sekali, update, sekarang, simpel, color, coded, bahkan, ibu, sul...
4    [checkin, susah, checkout, lebih, susah, aplikasi, bahkan, sering, mati...
Name: content, dtype: object
```

Fig. 4. Data Pre-processing Result

## *D. Data Preparation*

We perform 4 steps in this stage, including:

a. Splitting Data

We divide the review of version 4.0.2 into 70% train data, 15% validation data, and 15% test data with random_state of 42. Thus, 1.624 data trains, 348 validation data, and 348 test data.

b. Vectorization

This stage aims to convert the text into a unique integer list form where each integer represents a unique word in the dictionary. Figure 5 shows the results of vectorization. The vocabulary size is 1.911, which indicates there are 1.911 unique words in the data.

```
Vocab size         : 1911
Original data      : buat ehac pilih airplane ubah jadi car terus
After vectorization: [9, 38, 81, 235, 73, 8, 162, 14]
```

Fig. 5. Vectorization Result of Review Data

c. Pad Sequences

We transform each review to the same length so it can enter the neural network. The max sentence length parameter will determine how long each sequence is based on the longest sentence in the review data. To have the inputs with the same length, we fill the empty slot in the sequence with 0.

```
Max sentence length : 54
Original data       : buat ehac pilih airplane ubah jadi car terus
After vectorization and pad sequences:
[  9  38  81 235  73   8 162  14   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
```

Fig. 6. Pad Sequences Result on Review Data

d. Encode Label and Array Convert

We perform the label encoding at this stage to convert the string type of aspect and sentiment label to a unique integer. The aspect label is then converted into a binary matrix.

## *E. Classification Modelling*

1) Aspect Classification Modelling

The CNN model in this study uses CNN sequential, with each layer stacked linearly from end to end. Each layer and parameter used can be seen in Table V.

TABLE V. LAYER STRUCTURES OF ASPECT CNN MODEL

| Layer | Specification |
|---|---|
| Embedding | Input_dim = 1.911<br>Output_dim = 150<br>Input_length = 54 |
| Conv1D | Filters = 128<br>Kernel_size = 9<br>Activation = relu |
| GlobalMaxPooling1D | - |
| Dropout | Rate = 0.2 |
| Dense | Units = 8<br>Activation = Softmax |

We compile the aspect model using an Adam optimizer with a learning rate of 0.0001 and categorical cross-entropy for loss type.

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding_1 (Embedding)      (None, 54, 150)           286650
_________________________________________________________________
conv1d_1 (Conv1D)            (None, 46, 128)           172928
_________________________________________________________________
global_max_pooling1d_1 (Glob (None, 128)               0
_________________________________________________________________
dropout_1 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_1 (Dense)              (None, 8)                 1032
=================================================================
Total params: 460,610
Trainable params: 460,610
Non-trainable params: 0
```

Fig. 7. CNN Model Initiation for Aspect Classification

We train the initiated aspect model using data train, while the validation process uses the validation data.

2) Sentiment Classification Modelling

The sentiment classification model also uses CNN sequential. Each layer and parameter used can be seen in Table VI.

TABLE VI. LAYER STRUCTURES OF SENTIMENT CNN MODEL

| Layer | Specification |
|---|---|
| Embedding | Input_dim = 1.911<br>Output_dim = 50<br>Input_length = 54 |
| Conv1D | Filters = 128<br>Kernel_size = 3<br>Activation = relu |
| GlobalMaxPooling1D | - |
| Dropout | Rate = 0.2 |
| Dense | Units = 1<br>Activation= Sigmoid |

We also compile the sentiment model using an Adam optimizer with a learning rate of 0.001 and binary cross-entropy for loss type.

```
Model: "sequential_2"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding_2 (Embedding)      (None, 54, 50)            95650
_________________________________________________________________
conv1d_2 (Conv1D)            (None, 52, 128)           19328
_________________________________________________________________
global_max_pooling1d_2 (Glob (None, 128)               0
_________________________________________________________________
dropout_2 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_2 (Dense)              (None, 1)                 129
=================================================================
Total params: 115,107
Trainable params: 115,107
Non-trainable params: 0
```

Fig. 8. CNN Model Initiation for Sentiment Classification

Then, we train the initiated sentiment model using data train, and the validation process uses the validation data.

### F. *Model Evaluation*

1) Evaluate and Test Aspect Classification Model

Figure 9 shows the accuracy and loss graph during the training and validation using the CNN model for aspect classification.

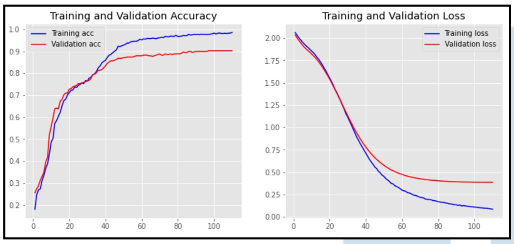



Fig. 9. Accuracy and Loss of CNN Model of Aspect

We can see the CNN model for aspect classification is well trained with training accuracy of 98.5% and validation accuracy of 90.2%. There is no overfitting indicated by the validation loss that continues to decrease in each epoch.

Then we apply the model into the test data to classify the aspect of the review. Figure 10 shows the confusion matrix of the aspect classification using test data.

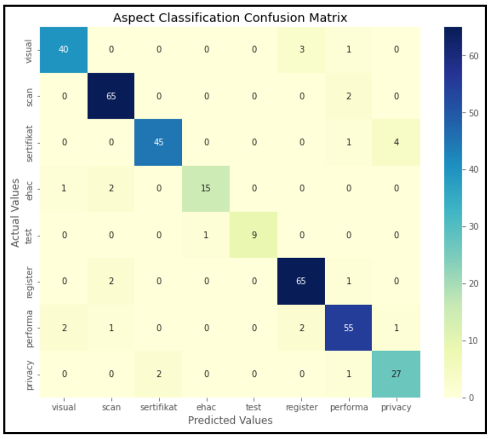



Fig. 10 Aspect Classification Confusion Matrix

The CNN aspect model got an overall accuracy of 0.9224, precision of 0.9234, recall of 0.9224, and f1 score of 0.9223. The performance of each label can be seen in Table VII. The Scan Check-in or check-out aspect has the highest F1 score of 0.9489, while the aspect with the lowest F1 score is the Privacy, Data, and Security aspect with 0.8710.

TABLE VII. ASPECT CLASSIFICATION REPORT

| Aspect | Precision | Recall | F1 |
|---|---|---|---|
| Visual Experience | 0.9302 | 0.9091 | 0.9195 |
| Scan, Checkin/out | 0.9286 | 0.9701 | 0.9489 |
| Vaccine Certificate | 0.9574 | 0.9000 | 0.9278 |
| eHac | 0.9375 | 0.8333 | 0.8824 |
| COVID-19 Test | 1.0000 | 0.9000 | 0.9474 |
| Register/Login | 0.9286 | 0.9559 | 0.9420 |
| Performance and Stability | 0.9016 | 0.9016 | 0.9016 |
| Privacy, Data, and Security | 0.8438 | 0.9000 | 0.8710 |
| **Accuracy** | 0.9224 | | |

2) Evaluate and Test Sentiment Classification Model

Figure 11 shows the accuracy and loss graph during the training and validation of the CNN model for sentiment classification.

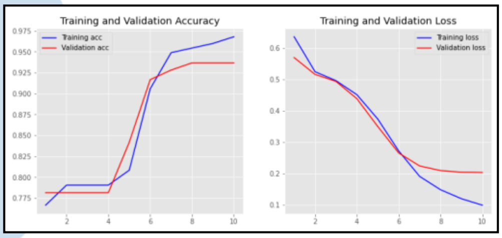



Fig. 11. Accuracy and Loss of CNN Model of Sentiment

The CNN model for sentiment classification is also well trained with a training accuracy of 97.4% and validation accuracy of 93.7%. Also, there is no overfitting indicated.

We apply the sentiment model into test data to classify the sentiment of the user review. Figure 12 shows the confusion matrix of the sentiment classification using test data.

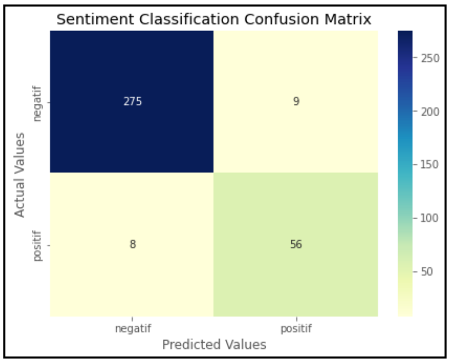



Fig. 12. Confusion Matrix of Sentiment Classification

The CNN sentiment model got an overall accuracy of 0.9510, precision of 0.9514, recall of 0.9510, and f1 score of 0.9513. The performance of each label can be seen in Table

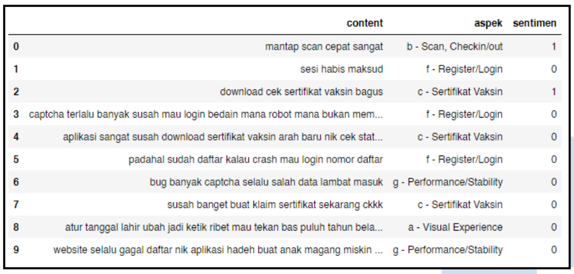
VIII. Negative sentiment has the highest F1 score of 0.97, while positive sentiment has an F1 score of 0.8682.

TABLE VIII. SENTIMENT CLASSIFICATION REPORT

| Sentiment | Precision | Recall | F1 |
|---|---|---|---|
| Negative | 0.9717 | 0.9683 | 0.9700 |
| Positive | 0.8615 | 0.8750 | 0.8682 |
| **Accuracy** | 0.9510 | | |

## G. Implementation

From the model evaluation, we concluded that the CNN model for both classification of aspects and sentiments has a good performance. After that, we perform the classification predictions to the unlabeled data using both CNN aspect and sentiment model.



| | content | aspek | sentimen |
|---|---|---|---|
| 0 | mantap scan cepat sangat | b - Scan, Checkin/out | 1 |
| 1 | sesi habis maksud | f - Register/Login | 0 |
| 2 | download cek sertifikat vaksin bagus | c - Sertifikat Vaksin | 1 |
| 3 | captcha terlalu banyak susah mau login bedain mana robot mana bukan mem... | f - Register/Login | 0 |
| 4 | aplikasi sangat susah download sertifikat vaksin arah baru nik cek stat... | c - Sertifikat Vaksin | 0 |
| 5 | padahal sudah daftar kalau crash mau login nomor daftar | f - Register/Login | 0 |
| 6 | bug banyak captcha selalu salah data lambat masuk | g - Performance/Stability | 0 |
| 7 | susah banget buat klaim sertifikat sekarang ckkk | c - Sertifikat Vaksin | 0 |
| 8 | atur tanggal lahir ubah jadi ketik ribet mau tekan bas puluh tahun bela... | a - Visual Experience | 0 |
| 9 | website selalu gagal daftar nik aplikasi hadeh buat anak magang miskin ... | g - Performance/Stability | 0 |

Fig. 13. Classified Aspect and Sentiment of 4.0.5 Review

Table IX shows the distribution of the classified aspects and sentiments on 4.0.5 application review.

TABLE IX. LABEL DISTRIBUTON OF CLASSIFIED 4.0.5 REVIEWS

| Aspect | Positive | Negative | Total |
|---|---|---|---|
| Visual Experience | 28 | 63 | 91 |
| Scan, Check-in or check-out | 12 | 141 | 153 |
| Vaccine Certificate | 17 | 177 | 194 |
| eHac | 1 | 18 | 19 |
| COVID-19 Test | 0 | 34 | 34 |
| Register/Login | 8 | 169 | 177 |
| Performance and Stability | 55 | 263 | 318 |
| Privacy, Data, and Security | 6 | 39 | 45 |
| **Total** | 127 | 904 | 1.031 |

## H. Interpretation

Figure 14 compares sentiment in each aspect in versions 4.0.2 and 4.0.5, where red represents negative sentiment and blue represents positive sentiment.

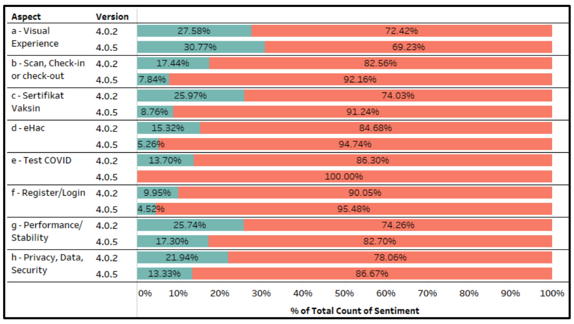



Fig. 14. Comparison of Sentiment in Each Aspects

User sentiment on both version of PeduliLindungi application is dominated by negative sentiment. Figure 15 shows that if we explored it per aspect, almost all the aspects experienced an increase in negative sentiment in version 4.0.5, except for the Visual Experience aspect.

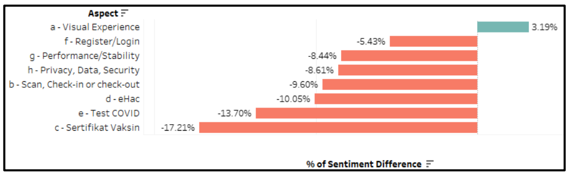



Fig. 15. Sentiment Percentage Difference in Version 4.0.2 to 4.0.5

Vaccine Certificate aspect experienced the highest increase in negative sentiment in version 4.0.5 by 17.21%. It was caused by an additional CAPTCHA that did not function well in claiming vaccine certificate; as in the review, "Setiap mau klaim sertifikat vaksin gak bisa, captcha ngulang terus" (every time (I) want to claim a vaccine certificate, it fails, the captcha keeps repeating).

The COVID-19 Test aspect also experienced an increase of 13.70% due to the number of test results that did not available in the application, such as in the review "Udah test PCR di RS AK Gani Palembang, tapi hasil di PeduliLindungi belum keluar padahal besok terbang" ((I) had a PCR test at the AK Palembang hospital, but the results PeduliLindungi have not been released, even though tomorrow (is the schedule to) fly).

Furthermore, the third-highest aspect is the eHac aspect, increasing 10.05% due to the eHac creation flow becoming more complex, as in the review "Pengisian eHac yang terbaru tidak praktis." (The latest (version of) eHac filling is impractical).

From the comparison between sentiment per aspect on version 4.0.2 and 4.0.5, it can be concluded that the application updates made an increase of negative sentiment in 7 aspects of the application. It is necessary to improve the PeduliLindungi application based on the cause. This analysis can also be used to prioritize aspects for corrective action in the following application update.

Research conducted by [8] and [9] has proven that CNN performs well for the ABSA task in classifying aspects and sentiments of user reviews. As for this research, we also performed the classification task to unlabeled data and compared each sentiment per aspect on different review per application version to expose whether an application update significantly changed the user's sentiment. Furthermore, this research's application of the CNN model gave better results with no overfitting indicated. The targeted aspects in this research will also give a better insight into the application usage since we use more detailed aspects than in the previous research.

## V. Conclusion

This study performed good results of Aspect-Based Sentiment Analysis (ABSA) using CNN model on aspect classification and sentiment of review data. The results showed that the CNN model could produce such good performance with an f1 score of 92.23% in the aspect classification and 95.13% in the sentiment classification.

User sentiment on the eight aspects of the application: Visual Experience, Scan – Checkin/Out, Vaccine Certificate, eHac, COVID-19 Test, Register/Login, Performance and Stability, and Privacy, Data, and Security is dominated by negative sentiment. As for the application version 4.0.5, the sentiment given to each aspect increased in negative sentiment, except for the Visual Experience aspect.

In version 4.0.5, the Vaccine Certificate aspect increased 17.21% due to the CAPTCHA feature that did not function properly. It was then followed by COVID Test aspect by 13.70% due to the large number of test results not released in the application, and the eHac aspect of 10.05% due to the impractical flow of the eHac filling.

Since the performance of the CNN model in this study has proven to be good, it can be continued with the development of applications that will facilitate the monitoring of user sentiment on every aspect of the application review. For further research, exploration of word embedding options also can be carried out using pre-trained word embedding such as Word2Vec or Glove to improve the word representation with semantic meaning.

## References


[1] Tim Detikcom, “Kapan Sebenarnya Corona Pertama Kali Masuk RI?,” Apr. 26, 2020. https://news.detik.com/berita/d-4991485/kapan-sebenarnya-corona-pertama-kali-masuk-ri (accessed Nov. 15, 2021).

[2] H. John Leon Singh, D. Couch, and K. Yap, “Mobile Health Apps That Help With COVID-19 Management: Scoping Review,” JMIR Nursing, vol. 3, no. 1, p. e20596, Aug. 2020, doi: 10.2196/20596.

[3] PeduliLindungi, “Apa itu PeduliLindungi?” https://www.pedulilindungi.id/#tentang (accessed Nov. 22, 2021).

[4] C. G. Asmara, “Selama PPKM, Kegiatan Ini Wajib Pakai Aplikasi PeduliLindungi,” Sep. 01, 2021. https://www.cnbcindonesia.com/news/20210901152445-4-272941/selama-ppkm-kegiatan-ini-wajib-pakai-aplikasi-pedulilindungi (accessed Nov. 16, 2021).

[5] BBC News, “Sanksi pelanggar aplikasi Pedulilindungi menuai kritik: ‘negara ini senang sekali menghukum warganya,’” Dec. 23, 2021. https://www.bbc.com/indonesia/indonesia-59759975 (accessed Mar. 07, 2022).

[6] Z. M. Zohreh Madhoushi, A. R. Hamdan, and S. Zainudin, “Aspect-Based Sentiment Analysis Methods in Recent Years,” Asia-Pacific Journal of Information Technology & Multimedia, vol. 08, no. 01, pp. 79–96, Jun. 2019, doi: 10.17576/apjitm-2019-0801-07.

[7] P. Rizki Amalia and E. Winarko, “Aspect-Based Sentiment Analysis on Indonesian Restaurant Review Using a Combination of Convolutional Neural Network and Contextualized Word Embedding,” IJCCS (Indonesian Journal of Computing and Cybernetics Systems, vol. 15, no. 3, pp. 285–294, 2021, doi: 10.22146/ijccs.67306.

[8] S. Cahyaningtyas, D. Hatta Fudholi, and A. Fathan Hidayatullah, “Deep Learning for Aspect-Based Sentiment Analysis on Indonesian Hotels Reviews,” Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control, Aug. 2021, doi: 10.22219/kinetik.v6i3.1300.

[9] M. T. Ari Bangsa, S. Priyanta, and Y. Suyanto, “Aspect-Based Sentiment Analysis of Online Marketplace Reviews Using Convolutional Neural Network,” IJCCS (Indonesian Journal of Computing and Cybernetics Systems), vol. 14, no. 2, p. 123, Apr. 2020, doi: 10.22146/ijccs.51646.

[10] IBM Cloud Education, “Text Mining,” Nov. 16, 2020. https://www.ibm.com/cloud/learn/text-mining (accessed Nov. 17, 2021).

[11] K. Kowsari, K. J. Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, “Text classification algorithms: A survey,” Information (Switzerland), vol. 10, no. 4. MDPI AG, 2019. doi: 10.3390/info10040150.

[12] A. S. Nayak, A. P. Kanive, N. Chandavekar, and B. R, “Survey on Pre-Processing Techniques for Text Mining,” International Journal Of Engineering And Computer Science, Jun. 2016, doi: 10.18535/ijecs/v5i6.25.

[13] H. Peng, Y. Ma, Y. Li, and E. Cambria, “Learning multi-grained aspect target sequence for Chinese sentiment analysis,” Knowledge-Based Systems, vol. 148, pp. 167–176, May 2018, doi: 10.1016/j.knosys.2018.02.034.

[14] Y. Kim, “Convolutional Neural Networks for Sentence Classification,” Aug. 2014. doi: 10.3115/v1/D14-1181.

[15] M. Z. Amin and N. Nadeem, “Convolutional Neural Network: Text Classification Model for Open Domain Question Answering System,” 2018.

[16] K. O’Shea and R. Nash, “An Introduction to Convolutional Neural Networks,” Nov. 2015, [Online]. Available: http://arxiv.org/abs/1511.08458

[17] E. N. Arrofiqoh and H. Harintaka, “IMPLEMENTASI METODE CONVOLUTIONAL NEURAL NETWORK UNTUK KLASIFIKASI TANAMAN PADA CITRA RESOLUSI TINGGI,” GEOMATIKA, vol. 24, no. 2, p. 61, Nov. 2018, doi: 10.24895/jig.2018.24-2.810.

[18] Android Developers, “Core app quality,” May 17, 2021. https://developer.android.com/docs/quality-guidelines/core-app-quality (accessed Nov. 28, 2021).