Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.2, April 2025, pp. 361~375
e-ISSN: 2548-6861

Aspect-Based Sentiment Analysis with LDA and IndoBERT Algorithm on Mental Health App: Riliv

Firda Ayu Dwi Aryanti [1]*, Ardytha Luthfiarta [2]*, Dennis Adiwinata Irwan Soeroso [3]*
* Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang, 50131, Indonesia
firdaaryanti65@gmail.com[1], ardytha.lutfhfiarta@dsn.dinus.ac.id[2], 111202113225@mhs.dinus.ac.id[3]

Article Info

Article history:

Received 2024-11-28
Revised 2025-01-07
Accepted 2025-01-21

Keyword:

ABSA,
IndoBERT,
LDA,
Mental Health,
Sentiment Analysis

ABSTRACT

Indonesia's mental health crisis in 2024 is increasing along with the high growth of internet users. Thus, high internet usage provides an opportunity for mobile applications including Riliv, a popular mental health application in Indonesia to become the most complete solution for overthinking, anxiety, and depression. This research aims to analyze the sentiment correlation of aspects based on App Store and Play Store reviews through scraping to effectively expose Riliv’s user satisfaction knowledge to developers using topic labeling with Latent Dirichlet Allocation (LDA) and sentiment labeling using Bidirectional Encoder Representations from Transformers (BERT) indobenchmark/indobert-base-p1 model on Aspect-Based Sentiment Analysis (ABSA). This study used 3068 reviews from September 2015 to December 2024. The main results obtained were 1) Identified the sentiment that positive is highest in 2020, neutral is highest in 2020, and negative is highest in 2018. 2) Identified 4 main aspects of the Riliv application: Access Support, Counseling Services, Meditation Features, and User Interface with LDA. 3) The majority distribution was negative on User Interface, neutral on Counseling Services, and positive on Meditation Features. 4) The effectiveness of IndoBERT compared to the non-transformer baseline algorithm. 5) The most optimal results were obtained with 70% training, 10% validation, and 20% testing, resulting in 95% accuracy.

I. INTRODUCTION

Mental health is a global concern, including in Indonesia. In 2022, it was found that around 15.5 million adolescents had mental health symptoms with the highest anxiety data, followed by hyperactivity, depression, behavioral problems, and post-traumatic stress reported from the I-NAMHS survey [1]. If left untreated, one of the most dangerous effects of mental illness is suicide attempts, which have occurred in various cases in students and anyone else [2]. The high prevalence of mental disorders is in line with the rapid increase in internet usage in Indonesia. They are reaching 221.536.479 people in 2024 with a penetration of 79.5%. The most commonly used devices are mobile phones or tablets, with the percentage of users reaching 89.44% [3]. The dominance of mobile devices opens up great opportunities for developing mobile-based digital applications to meet the community's needs in accessing information, services, and entertainment quickly and easily. One of them is the Riliv application. The Riliv application received an award from Google Japan 2015 for Best Impactful Android Application, from the Singapore International Foundation 2016 of 16 Best Young Social Entrepreneurs SEA, and Australian Awards 2017 [4]. It is Indonesia's first mental health application that provides various complete features including a mood tracker, meditation, counseling, journal, and journey. Until December 2024, more than 500 thousand users have downloaded it, with a total rating of 4.8 on the Play Store and 3,8 on the App Store. The good and bad ratings on the Riliv app make the app assessment less accurate.

Customer ratings often do not express the overall service provided [5]. Because the opinions given are usually in text form and are not relevant to the star rating on the rating feature [6] can be seen that the problem is the need for sentiment analysis of reviews by utilizing opinion representations from service users [7]. So it takes a very important analysis of sentiment to process a sentence.

Because a sentence must have aspects in it, either one or more. And leads to a certain purpose or goal. So it is necessary to classify Riliv reviews using aspect determination. Aspect-Based Sentiment Analysis (ABSA) is part of the sentiment analysis model, a branch of Natural Language Processing that can determine the number of aspects of the polarity of a sentence [8]. There are many methods in topic modeling, the most superior in performance is Latent Dirichlet Allocation (LDA) compared to others, which can be implemented for topic identification in scientific journals, classification, and clustering [9].

Previous research related to LDA and BERT performance by [10]. From Play Store users taken from GOJEK application reviews, user experience, service, and payment aspects are generated, with the highest accuracy in the Service aspect reaching 98.78%. For sentiment labeling from evaluation results influenced by the distribution of sentiment, the amount of data, data preprocessing, and evaluation models with the highest accuracy at 96.67%.

The use of IndoBERT in a study by [11] Showed significant effectiveness in dealing with the complexity of the Indonesian language, with the ability to overcome linguistic challenges such as complex morphology, diverse sentence structures, and distinctive use of particles and pronominal. IndoBERT can distinguish meanings between words that share the same root but different implications, such as 'raise' and 'raised', and identify the subject in a sentence even if it is not at the beginning. In addition, the model successfully captured sentiment nuances in analyzing tweets related to tax policy. This research also shows that IndoBERT effectively recognizes technical terms and loanwords frequently used in the Ministry of Finance, with the sentiment analysis results of 10,099 tweets showing the dominance of negative sentiments. Through Aspect-Based analysis (ABSA) and topic modeling using Latent Dirichlet Allocation (LDA), this research provides valuable insights for the Ministry of Finance to understand public perception and areas that require special attention in public communication and policy.

This research aims to explore the sentiment distribution of user reviews of the Riliv application, identify the aspects most frequently discussed by users, analyze the temporal trend of sentiment from 2015 to 2024, analyze sentiment on aspects, and predict review sentiment using the Non-Transformer and Transformer algorithms. It is expected that the results of this study can provide strategic insights for the development of mental health digital services in Indonesia, as well as an understanding of certain aspects affecting user satisfaction.

II. RESEARCH METHOD

This research has an illustrative flow of each stage shown in Figure 1. First, Data Collecting, Manual Labeling, Exploratory Data Analysis (ABSA), Pre-Processing, SMOTE, Data Transformation, Topic Clustering with LDA, Classification with Non-Transformer or Transformer and Evaluation (ABSA) are performed.

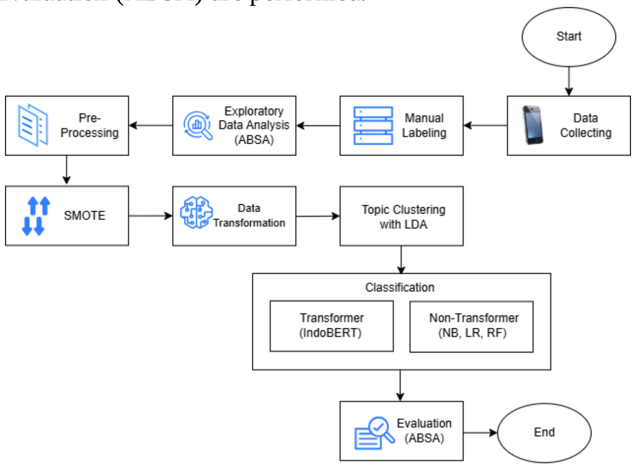


Figure 1. Flow of Research Methods

A. Data Collecting

The first stage that the author did in this research was to collect data related to journals and relevant websites. The data source focuses on the mental health application "Riliv" uploaded from September 2015 to December 2024. Data collection is done through web scraping using the app_store_scraper and google_play_scraper library with the Python programming language in Google Colab. Web scraping is the process of obtaining information automatically from a website by retrieving and extracting data of a certain size [12].

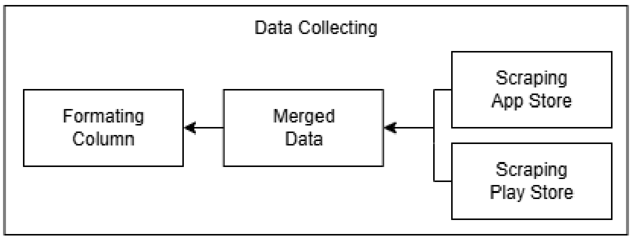


Figure 2. Flow of Data Collecting Methods

B. Manual Labeling

Manual labeling is done by the author to get the actual positive, neutral, and negative sentiments, using data from Riliv as in the following Figure 3.

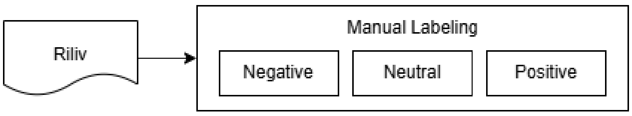


Figure 3. Flow of Manual Labeling Methods

C. Exploratory Data Analysis (ABSA)

Exploratory Data Analysis (EDA) is very important in this research. Conducting in-depth data exploration on how the dataset obtained can be processed for Aspect-Based Sentiment Analysis (ABSA). Check data type and basic statistics to describe the overall exploration of the data and whether it is to the needs of the analysis. Check sentiment distribution to determine the amount of sentiment distribution. Check cross-tabulate sentiment and score to validate the correctness of the sentiment the rating score parameter shows similarity with the actual label. Check the distribution of review lengths to find out how many times each review length is counted. Check review length by sentiment to find out the sentiment towards review length. Check yearly sentiment counts over time to analyze the number of negative, neutral, and positive sentiments in each year from 2015 to 2024. And check average ratings over time to analyze the average rating for each year from 2015 to 2024.

D. Pre-Processing

Pre-processing is an important step in transforming raw data text into clean and relevant for modeling analysis. In this research, the preprocessing process is carried out through five main stages.

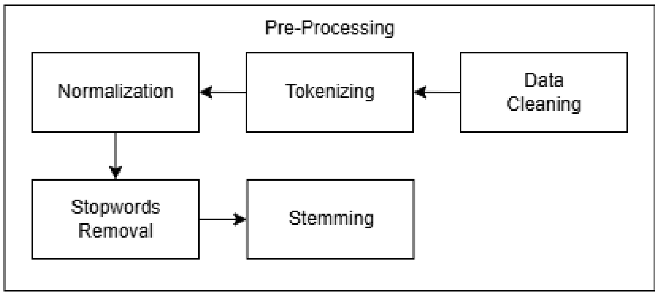


Figure 4. Flow of Pre-Processing Methods

1) *Data Cleaning*

The first step is removing missing values and duplicates from all columns to eliminate bias in the data process. Followed by case folding, remove special characters re.sub(r'[!"#$%&\'()*+: <=>?@\[\\\]^_`{|}~'\d]+', ' ', text), then remove repeated words next to each other re. sub(r'\b(\w+)\s+\1\b', r'\1', text), then remove single letters re.sub(r'\b\w\b', '', text) then replace commas, periods, and slashes with spaces re.sub(r'[,/.-]', '', text), then remove excess spaces.

2) *Tokenizing*

The process of text is broken down into tokens or phrase units.

3) *Normalization*

Word correction in reviews can change non-standard words to standardized ones and can describe certain abbreviations. Using over 1400 additional normalization words created by the author.

4) *Stopwords Removal*

Stop the list of words not complying with the common language source standards. Can improve text mining performance [13]. This research uses Sastrawi and an additional 280 stopwords from the author.

5) *Stemming*

The process of finding the base word or root word from the filtering results of each word [14]. This research uses the Sastrawi, Indonesian complex library. After that, it is combined again into a sentence.

E. Synthetic Minority Over-sampling Technique (SMOTE)

Applying resampling strategies to obtain a more balanced data distribution is an effective solution to the problem of imbalance. Handling imbalanced data on the Riliv review dataset using the Synthetic Minority Over-sampling Technique (SMOTE) can manage the balance of multiple classes. If there is a minority class, SMOTE can re-sample the minority class for a balance [15].

F. Data Transformation

Data transformation is divided into data division and data weighting. Data division of LDA and Classification with Non-Transformer, training data, and testing data. After that, it is weighted by extracting features into vectors using the TfidVectorizer module. The word frequency-inverse text frequency algorithm, or TF-IDF algorithm, is a popular weighting technique for text mining and information retrieval. The main argument is that a word's value is directly correlated with its frequency in the article and inversely correlated with its frequency in the corpus, excluding stop words.

Term frequency (TF) is the number of times a word appears in the document it is in. It is calculated using Equation (1), where $N\omega$ is the total number of terms in the text and reflects the number of times the word ω occurs in the text.

$$TF_{\omega} = \frac{N_{\omega}}{N} \quad (1)$$

The frequency of a word appearing throughout the entire document is reflected in the IDF (Inverse Document Frequency). Y is the total number of documents in the corpus, and $Y\omega$ is the number of documents that contain the term ω. Equation (2) is the method used to calculate IDF [16]

$$IDF_{\omega} = log\, \frac{Y}{Y_{\omega}+1} \quad (2)$$

Equation (3) illustrates that TF-IDF is the product of the values of TF and IDF. When extracting keywords from a text, a word with a larger TF-IDF value is considered more significant and is therefore more likely to be retained. In contrast, a word with a smaller value is more likely to be eliminated.

$$TF - IDF = TF_{\omega} * ID \quad (3)$$

While the Classification with Transformer data division uses training data, validation data, and testing data and weighting uses a Transformer-based embedding representation.

G. Topic Clustering with LDA (Latent Dirichlet Allocation)

1) Apply LDA

Three layers of words, topics, and documents make up the LDA (Latent Dirichlet Allocation) model for document topic generation. Word segmentation is used to ascertain the word frequency of each word in each document, after which the "document word" matrix is produced. This is the LDA topic model algorithm's fundamental concept. After obtaining the document-word matrix, "subject-word" matrix, and "document-topic" matrix through training, the document's topics are analyzed. The probability that each word will appear in the text is shown as follows:

$$\mathrm{P(word|doc)} = \sum_{\mathrm{topic}} \mathrm{p(word|topic)} \times \mathrm{p(doc|topic)} \quad (4)$$

The specific procedure used in this paper to improve the traditional LDA topic model based on the TF-IDF algorithm is as follows: Following sentiment classification, positive and negative data sets for each of the three categories of goods were obtained. To create a dictionary and corpus, use the gensim library in conjunction with the TF-IDF feature word extraction and corpora module. Apply the model. The number of topic topics for LDA model training is specified by LdaModel().

2) Topic Interpretation

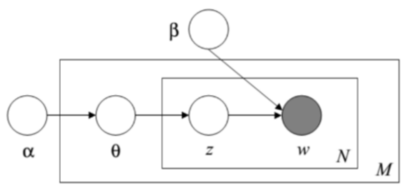


Figure 5. Linear Dirichlet Allocation Model Representation

M is the number of documents in the corpus, N is the number of words in the document, K is the number of topics to be found in the corpus, Alpha (α) is a parameter for the distribution of topics (θ) in documents, while beta (β) is a parameter for the distribution of words (φ) in topics. Z is the selected topic for a particular word in the document. And W as a particular word in the document. Using the LDA model to group documents into specific topics and provide easy-to-understand labels for each topic [10].

3) Results of Each Aspect

Helps to understand how sentiment is spread across each cluster, aspects tend to be positive, neutral, or negative.

H. Classification

1) Classification With Non-Transformer (Baseline)

This study compares Naive Bayes, Logistic Regression, and Random Forest as simple and frequently used classical machine learning models as a baseline with IndoBERT as a more complex transformer-based deep learning model.

1.1) Naïve Bayes

Naïve Bayes is a classification method focused on the classification of textual data, which has a feature where the results obtained from each class are independent because one document with another document has no relationship to get pure results only processed from the document itself. The equation can be seen as follows, calculated using Equation (4).

$$P(c|d) \propto P(c) \prod_{i=1}^{n_d} P(w_I|c) \quad (4)$$

Description :

P(c|d) = Probability of a class in a document/text d
P(c) = Prior probability c
$P(w_i|c)$ = Probability of a word in class c [17]

1.2) Logistic Regression

Logistic Regression is a classification method for text data that models the relationship between independent variables (features) and dependent variables (labels) using a logistic function. Calculated using Equation (5).

$$g_j(x) = \beta_{j0} + \beta_{j1}x_1 + \beta_{j2}x_2 + \cdots + \beta_{jp}x_p \quad (5)$$

Description :

g(x) = Logit function for class j
βj0 = Intercept for class j
$\beta j1, \beta j2, \ldots, \beta j$p = coefficient for each predictor variable ($x1, x2, \ldots, x$p) in class j [18]

1.3) Random Forest

Random Forest is a data classification method whose final result is based on a tree formed from major voting. To determine the informative level of an attribute node and the information gain value in Equation (6), find the impurity in Equation (7). And averaging gini in Equation (8).

$$Gini = 1 - \sum_{i=1}^{n} (Pi)^2 \quad (2.1) \quad (6)$$

$$\textit{Average gini impurity} = \frac{n}{i} \times \textit{gini} \quad (7)$$

$$\textit{Information Gain} = \textit{Gini impurity} - \textit{Average Gini Impurity} \quad (8)$$

Description :

n = Number of terms in the class
Pi = Probability of term occurrence in the document
i = Number of documents [19]

2) Classification With Transformer

IndoBERT is one of the BERT (Bidirectional Encoder Representations from Transformers) methods for Natural Language Processing that works in two ways: pre-training and finetuning. Using a corpus source of more than 220 million Indonesian languages.

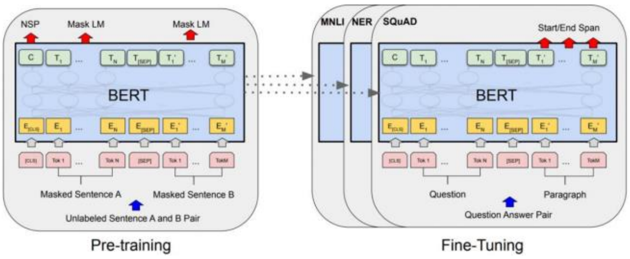


Figure 6. IndoBERT Pre-Training and Fine-Tuning Stages

In pre-training, the model learns using unlabeled data, while in fine-tuning, the model further adapts to pre-training parameters using labeled data. Both architectures are similar but have different output layers. The left image is the pre-training process which has the task of predicting random vague words called the Masked Language Model, represented by the green box labeled "Mask LM'. Then the next task is to predict whether the given sentence B follows Sentence A in the original text or is called the Next Sentence Prediction, illustrated by the red box labeled "NSP". Unlabeled sentences (sentences A and B) are processed by BERT during pre-training. The right image is the fine-tuning process which has the task of performing further refinements such as MNLI/NLI for the prediction of relationships between sentence pairs, NER for identification and categorization of entities in the text, SQuAD for the prediction of the start and end ranges of answers to questions based on paragraphs involving labeled data in each process [20]. The author uses epoch 15, lr 1e-5, weight with Adam optimizer.

I. Evaluation (ABSA)

TABLE I
CONFUSION MATRIX

Class	Prediction Negative	Prediction Neutral	Prediction Positive
Actual Negative	TNg	FNt2	FP1
Actual Neutral	FNg2	TNt	FP2
Actual Positive	FNg1	FNt1	TP

Description :

TNg = True Negative or predicted negative by the model and actual in the negative class

TP = True Positive or predicted positive by the model and actual in the positive class.

TNt = True Neutral or predicted neutral by the model and actual in the neutral class

FNg1/2 = False Negative or predicted negative by the model and actual in the positive or neutral class

FP1/2 = False Positive or predicted positive by the model and actual in the negative or neutral class

FNt1/2 = False Neutral or predicted neutral by the model and actual in the negative or positive class

Based on the 6 possibilities, the accuracy value is generated as the total number of correct values in Equation (9).

$$\frac{TP+TNg+TNt}{TP+FNt2+\ldots+FNt1+TNg} \tag{9}$$

Equation (10) produces Precision as how often the prediction is correct when the model predicts positive neutral or negative.

$$\text{Positive} = \frac{TP}{TP+FP1+FP2}$$
$$\text{Neutral} = \frac{TNt}{TNt+FNt1+FNt2}$$
$$\text{Negative} = \frac{TNg}{TNg+FNg1+FNg2} \tag{10}$$

Then Equation (11) produces Recall as how often the model predicts positive when the actual class is positive predicts neutral when the actual class is neutral or predicts negative when the actual class is negative.

$$\text{Positive} = \frac{TP}{FNg1+FNt1+TP}$$
$$\text{Neutral} = \frac{TNg}{FP1+FNt2+TNg}$$
$$\text{Negative} = \frac{TNt}{FNg2+FP2+TNt} \tag{11}$$

Equation (12) produces F1-Score as the harmonic mean of the Precision and Recall values [19]

$$2 \times \frac{Precison \times Recall}{Precision + Recall} \tag{12}$$

III. RESULT AND DISCUSSION

This research discusses the details of what has been previously outlined in the methods section.

A. Data Collecting

Scraping is taken from 2 sources namely App Store and Play Store using lang parameters in the form of id and country in Indonesia. Both have different column names.

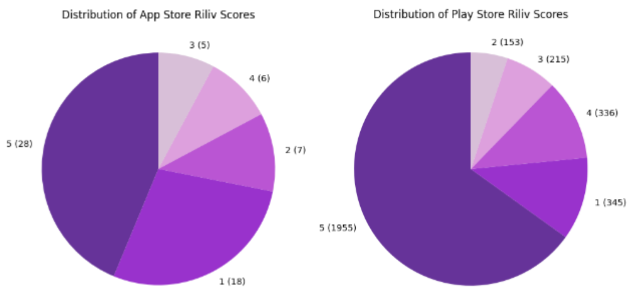


Figure 7. Distribution of App Store and Play Store Riliv Scores

Rename the columns on the App Store to be the same as the Play Store. That is review: content, date: at, and rating: score. This research collects several attributes from the dataset, the explanation of each attribute is in Table 2 below:

TABLE II
ATTRIBUTES FROM SCRAPING THE RILIV DATASET

Attributes	Description
userName	The identity of the author of the Riliv review, which can be a username or anonymous such
content	Review in the form of a series of words or sentences as an expression of the assessment of the use of the Riliv application.
score	The value range is a 1 to 5-star rating.
at	Time information when the review was uploaded. In the form of year, month, date, and hour

From the 4 columns as attributes that were obtained. A total of 3068 datasets were successfully collected from September 2015 to December 2024. After merging data or combination, the distribution of both can be seen in Figure 9.

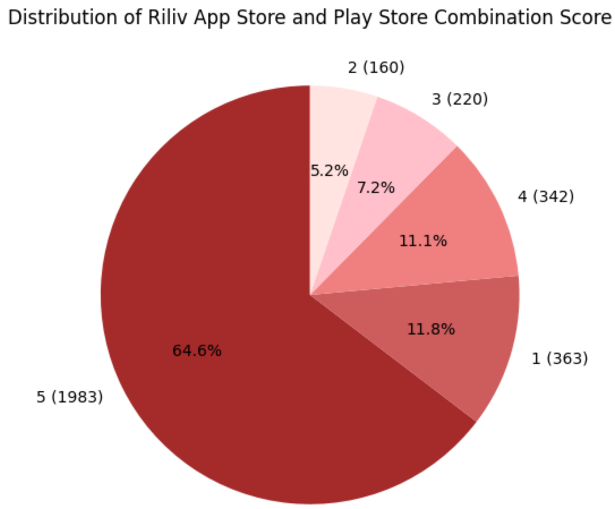


Figure 8. Distribution of App Store and Play Store Combination Score

B. *Manual Labeling*

The author determines the actual positive, neutral, and negative sentiments one by one, which is in Table 3 raw with manual labeling.

TABLE III
RILIV REVIEW BEFORE PREPROCESSING

userName	content	score	at	sentiment
Quine channel	Halah apaan gw udh bayar 100rb. Di tgl 31 juli. Gw pake riliv karna ngerasa bner2 butuh konsul ke psikolog saat itu juga karna gw udh ngedown parah. Terus dikasih pilihan jadwal konsul ama psikolog nya di tgl 4 atau 5 agustu. Gw minta di tgl nya dimajuin gk bisa katanya. Gila aja woy klo orng yg udh stress parah disuruh nunggu 4-5 hari udh keburu bunuh diri itu orang. Orng udh bayar itu tandanya dia lagi bner2 butuh saat itu juga. Klo disuruh nunggu 1 hari sih oke. Lah ini 4-5 hari. Situ mikir??	1	2021-08-01 01:14:43	negative
Pengguna Google	Aku pikir walau aku curhat ke aplikasi ini ga akan dibales atau walaupun dibales cuma seadanya ternyata engga sama sekali. Makasih riliv, its amazing. Aplikasi ini membantu banget terutama buat kalian yang lagi galau2nya sama semua orang. Tapi aku rasa ada item yang kurang deh, tolong dong adminnya riliv tambahin pemberitahuan dalam aplikasinya. Kaya di akun sosmed lainnya biar kita juga gampang tau kalo ada balesann. Makasih 😱☺☺	5	2017-09-06 14:05:35	neutral
Pengguna Google	Kenapa aplikasinya tidak bisa konek ke internet? Padahal semalam baikbaik saja. Saat saya buka pavi ini saya selalu mendapatkan tulisan " no internet found, please check your internet connection" padahal koneksi internet saya bagus sekali. Tolong dong diperbaiki. Terimakasih	3	2016-10-24 07:28:46	negative
Pengguna Google	Wow , udah di respon , bagus banget ni aps , punya masalah ? Curhat ke psikolog harus bayar , curhat ke temen takut di sebar , curhat ke ortu yang ada kena omel , akhirnya masalah ngendap tanpa soslusi akhirnya depresi ,	5	2016-03-03 12:12:01	positive

	ngerikan ?? Aps ini bikin kita bisa curhat + dapet solusi + dari psikolog pula ,, menang banyak kan :D thanks riliv 👍👍😊			
3068	…	…	…	…

C. *Exploratory Data Analysis (ABSA)*

After labeling data, ensure duplicate data and empty data from rows or columns. Deleting using Dropna missing value is obtained as null, from 3068 to 2912 with the distribution of data diversity in the following Table 4.

TABLE IV
NUMBER OF RILIV DATA PER YEAR

Year	**Total**
2015	67
2016	297
2017	303
2018	265
2019	295
2020	565
2021	533
2022	289
2023	170
2024	128
=	2912

Exploratory Data Analysis has several stages including. Check data type and basic statistics to describe the overall exploration of the data and whether it is to the needs of the analysis.

```
content             object
at          datetime64[ns]
userName            object
score                int64
sentimen            object
dtype: object
       content                             at          userName        score  \
count     2912                           2912              2912  2912.000000
unique    2688                            NaN              1557          NaN
top      Bagus                            NaN   Pengguna Google          NaN
freq        37                            NaN              1352          NaN
mean       NaN  2020-02-13 08:33:58.584821248               NaN     4.085508
min        NaN            2015-09-06 11:32:02               NaN     1.000000
25%        NaN     2018-02-18 10:38:36.500000               NaN     3.000000
50%        NaN            2020-06-20 11:25:32               NaN     5.000000
75%        NaN  2021-09-30 03:45:27.750000128               NaN     5.000000
max        NaN            2024-12-22 15:46:05               NaN     5.000000
std        NaN                            NaN               NaN     1.429400

        sentimen
count       2912
unique         3
top     positive
freq        1812
mean         NaN
min          NaN
25%          NaN
50%          NaN
75%          NaN
max          NaN
std          NaN
```

Figure 9. Data Type and Basic Statistic

Check sentiment distribution to determine the amount of sentiment distribution.

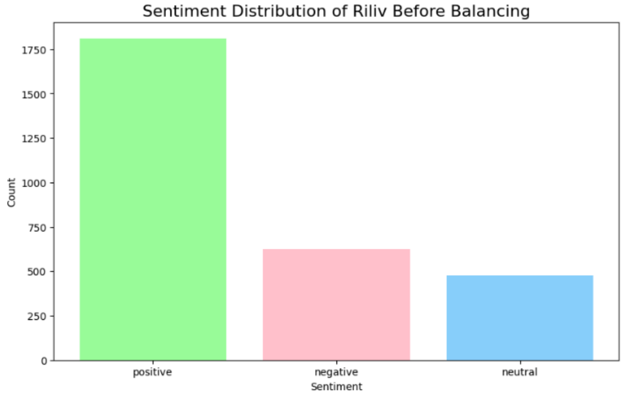


Figure 10. Sentiment Distribution of Riliv Before Balancing

Check the cross-tabulate score and sentiment to validate the correctness of the sentiment the rating score parameter shows similarity with the actual label.

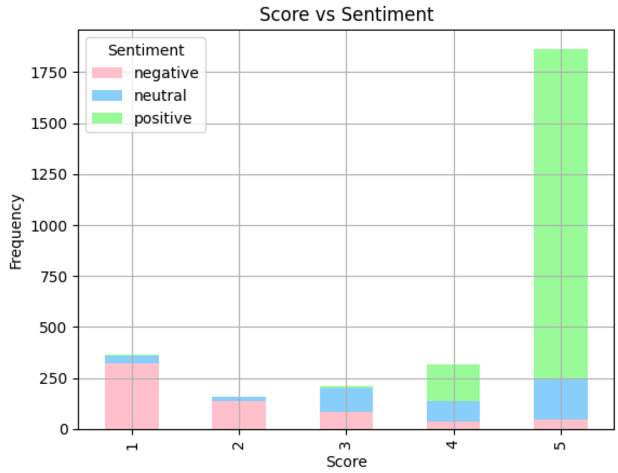


Figure 11. Cross-tabulate Score and Sentiment

The proportions are formatted into percentages in the following Table 5.

TABLE V
SCORE AND SENTIMENT PERCENTAGE

Score X Sentiment	**1**	**2**	**3**	**4**	**5**
Negative	51.36%	21.67%	13.32%	5.94%	7.70%
Neutral	8.60%	5.03%	24.32%	21.38%	40.67%
Positive	0.06%	0.00%	0.66%	9.77%	89.51%

From a score of 1: 51.36% of reviews with a score of 1 have negative sentiment, 8.60% of reviews with a score of 1 have neutral sentiment, and 0.06% of reviews with a score of 1 have positive sentiment. Up to a score of 5: 7.70% of reviews with a score of 5 have negative sentiment, 40.67% of reviews with a score of 5 have neutral sentiment, and 89.51% of reviews with a score of 5 have positive sentiment. Low scores tend to be negative and high scores tend to be positive, but there are still some classes that do not match. Therefore sentiment analysis is important in this research.

Check the distribution of review lengths to find out how many times each review length is counted.

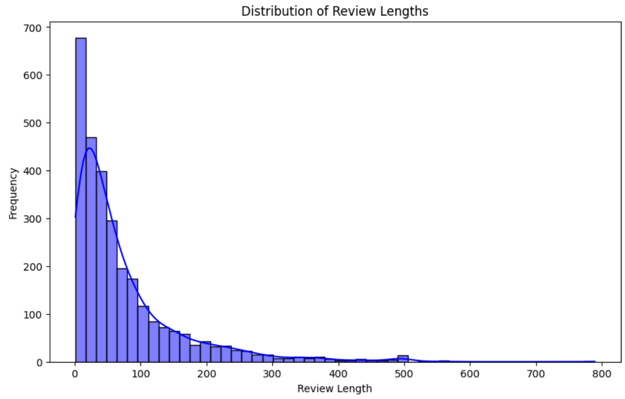


Figure 12. Distribution of Review Length

Check review length by sentiment to find out the sentiment towards review length.

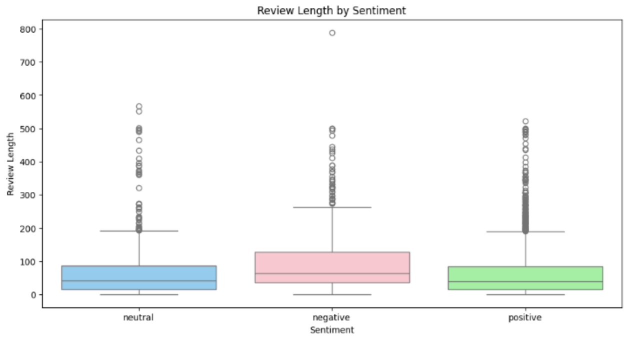


Figure 13. Review Length by Sentiment

Negative sentiment has a larger range of review lengths than other sentiments. This can be seen from the larger boxplot size and the presence of some outliers with much higher review lengths. Positive sentiment has a more centered distribution, with a smaller interquartile range than negative sentiment. Neutral sentiment has the narrowest distribution of review lengths, suggesting reviews tend to be short.

Check yearly sentiment counts over time to analyze the number of negative, neutral, and positive sentiments in each year from 2015 to 2024. Positive illustrations had the highest score in 2020 and the lowest in 2015. Then the neutral illustration has the highest value in 2020 and the lowest in 2015. And negative illustrations have the highest value in 2018 and the lowest in 2015.

And check the average ratings over time to analyze the average rating for each year from 2015 to 2024. The resulting reviews have the best average rating in 2024 and the worst rating in 2018. Illustrated in Figures 14 and 15 below.

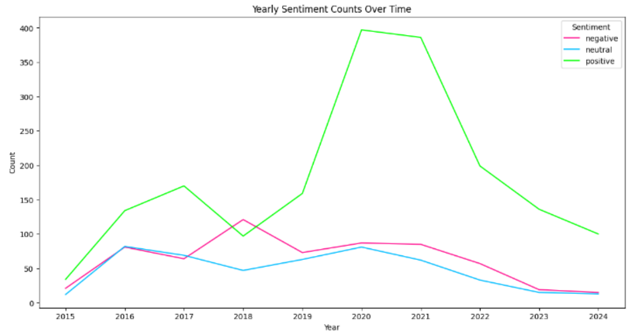


Figure 14. Yearly Sentiment Counts Over Time

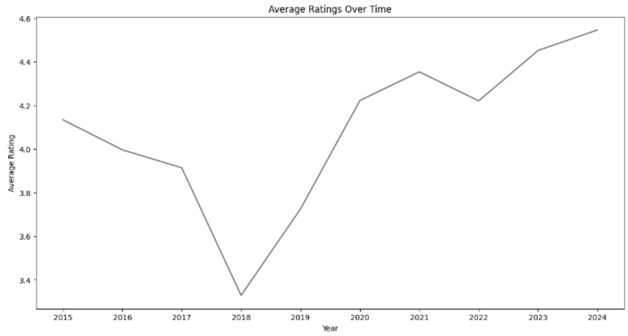


Figure 15. Average Ratings Over Time

D. *Pre-Processing*

1) Data Cleaning

To tidy up text that is difficult to read, a cleaning process is carried out with cleaning steps so that there is clarity in the text. The differences before and after the text is cleaned are shown in Table 6 below.

TABLE VI
COMPARISON OF SAMPLE DATA BEFORE AND AFTER DATA CLEANING

Before Cleaning	After Cleaning
Halah apaan gw udh bayar 100rb. Di tgl 31 juli. Gw pake riliv karna ngerasa bner2 butuh konsul ke psikolog saat itu juga karna gw udh ngedown parah. Terus dikasih pilihan jadwal konsul ama psikolog nya di tgl 4 atau 5 agustu. Gw minta di tgl nya dimajuin gk bisa katanya. Gila aja woy klo orng yg udh stress parah disuruh nunggu 4-5 hari udh keburu bunuh diri itu orang. Orng udh bayar itu tandanya dia lagi bner2 butuh saat itu juga. Klo disuruh nunggu 1 hari sih oke. Lah ini 4-5 hari. Situ mikir??	halah apaan gw udh bayar rb di tgl juli gw pake riliv karna ngerasa bner butuh konsul ke psikolog saat itu juga karna gw udh ngedown parah terus dikasih pilihan jadwal konsul ama psikolog nya di tgl atau agustu gw minta di tgl nya dimajuin gk bisa katanya gila aja woy klo orng yg udh stress parah disuruh nunggu hari udh keburu bunuh diri itu orang orng udh bayar itu tandanya dia lagi bner butuh saat itu juga klo disuruh nunggu hari sih oke lah ini situ mikir

2) Tokenization

After data cleaning, tokenization is performed to refine the data. The difference before and after the text is tokenized following Table 7.

TABLE VII
COMPARISON OF SAMPLE DATA BEFORE AND AFTER TOKENIZATION

Before Tokenization	After Tokenization
halah apaan gw udh bayar rb di tgl juli gw pake riliv karna ngerasa bner butuh konsul ke psikolog saat itu juga karna gw udh ngedown parah terus dikasih pilihan jadwal konsul ama psikolog nya di tgl atau agustu gw minta di tgl nya dimajuin gk bisa katanya gila aja woy klo orng yg udh stress parah disuruh nunggu hari udh keburu bunuh diri itu orang orng udh bayar itu tandanya dia lagi bner butuh saat itu juga klo disuruh nunggu hari sih oke lah ini situ mikir	[‘halah’, ‘apaan’, ‘gw’, ‘udh’, ‘bayar’, ‘rb’, ‘di’, ‘tgl’, ‘juli’, ‘gw’, ‘pake’, ‘riliv’, ‘karna’, ‘ngerasa’, ‘bner’, ‘butuh’, ‘konsul’, ‘ke’, ‘psikolog’, ‘saat’, ‘itu’, ‘juga’, ‘karna’, ‘gw’, ‘udh’, ‘ngedown’, ‘parah’, ‘terus’, ‘dikasih’, ‘pilihan’, ‘jadwal’, ‘konsul’, ‘ama’, ‘psikolog’, ‘nya’, ‘di’, ‘tgl’, ‘atau’, ‘agustu’, ‘gw’, ‘minta’, ‘di’, ‘tgl’, ‘nya’, ‘dimajuin’, ‘gk’, ‘bisa’, ‘katanya’, ‘gila’, ‘aja’, ‘woy’, ‘klo’, ‘orng’, ‘yg’, ‘udh’, ‘stress’, ‘parah’, ‘disuruh’, ‘nunggu’, ‘hari’, ‘udh’, ‘keburu’, ‘bunuh’, ‘diri’, ‘itu’, ‘orang’, ‘orng’, ‘udh’, ‘bayar’, ‘itu’, ‘tandanya’, ‘dia’, ‘lagi’, ‘bner’, ‘butuh’, ‘saat’, ‘itu’, ‘juga’, ‘klo’, ‘disuruh’, ‘nunggu’, ‘hari’, ‘sih’, ‘oke’, ‘lah’, ‘ini’, ‘situ’, ‘mikir’]

3) Normalization

After tokenization, normalization converts abbreviations and nonstandard slang words into standard words with regex from additional list normalization. The difference before and after the text is normalized follows Table 8.

TABLE VIII
COMPARISON OF SAMPLE DATA BEFORE AND AFTER NORMALIZATION

Before Normalization	After Normalization
[‘halah’, ‘apaan’, ‘gw’, ‘udh’, ‘bayar’, ‘rb’, ‘di’, ‘tgl’, ‘juli’, ‘gw’, ‘pake’, ‘riliv’, ‘karna’, ‘ngerasa’, ‘bner’, ‘butuh’, ‘konsul’, ‘ke’, ‘psikolog’, ‘saat’, ‘itu’, ‘juga’, ‘karna’, ‘gw’, ‘udh’, ‘ngedown’, ‘parah’, ‘terus’, ‘dikasih’, ‘pilihan’, ‘jadwal’, ‘konsul’, ‘ama’, ‘psikolog’, ‘nya’, ‘di’, ‘tgl’, ‘atau’, ‘agustu’, ‘gw’, ‘minta’, ‘di’, ‘tgl’, ‘nya’, ‘dimajuin’, ‘gk’, ‘bisa’, ‘katanya’, ‘gila’, ‘aja’, ‘woy’, ‘klo’, ‘orng’, ‘yg’, ‘udh’, ‘stress’, ‘parah’, ‘disuruh’, ‘nunggu’, ‘hari’, ‘udh’, ‘keburu’, ‘bunuh’, ‘diri’, ‘itu’, ‘orang’, ‘orng’, ‘udh’, ‘bayar’, ‘itu’, ‘tandanya’, ‘dia’, ‘lagi’, ‘bner’, ‘butuh’, ‘saat’, ‘itu’, ‘juga’, ‘klo’, ‘disuruh’, ‘nunggu’, ‘hari’, ‘sih’, ‘oke’, ‘lah’, ‘ini’, ‘situ’, ‘mikir’]	[‘halah’, ‘apaan’, ‘saya’, ‘sudah’, ‘bayar’, ‘ribu’, ‘di’, ‘tanggal’, ‘juli’, ‘saya’, ‘pakai’, ‘riliv’, ‘karena’, ‘merasa’, ‘benar’, ‘butuh’, ‘konsultasi’, ‘ke’, ‘psikolog’, ‘saat’, ‘itu’, ‘juga’, ‘karena’, ‘saya’, ‘sudah’, ‘ngedown’, ‘parah’, ‘terus’, ‘dikasih’, ‘pilihan’, ‘jadwal’, ‘konsultasi’, ‘sama’, ‘psikolog’, ‘nya’, ‘di’, ‘tanggal’, ‘atau’, ‘agustus’, ‘saya’, ‘minta’, ‘di’, ‘tanggal’, ‘nya’, ‘dimajukan’, ‘susah’, ‘katanya’, ‘gila’, ‘aja’, ‘woy’, ‘kalau’, ‘orang’, ‘yang’, ‘sudah’, ‘stress’, ‘parah’, ‘disuruh’, ‘tunggu’, ‘hari’, ‘sudah’, ‘keburu’, ‘bunuh’, ‘diri’, ‘itu’, ‘orang’, ‘orang’, ‘sudah’, ‘bayar’, ‘itu’, ‘tandanya’, ‘dia’, ‘lagi’, ‘benar’, ‘butuh’, ‘saat’, ‘itu’, ‘juga’, ‘kalau’, ‘disuruh’, ‘tunggu’, ‘hari’, ‘sih’, ‘oke’, ‘lah’, ‘ini’, ‘situ’, ‘mikir’]

4) Stopwords Removal

After normalization, stopword removal keeps important words like no, which affect the opposite meaning. The difference before and after the text is stopwords removal in the following Table 9.

TABLE IX
COMPARISON OF SAMPLE DATA BEFORE AND AFTER STOPWORDS REMOVAL

Before Stopwords Removal	After Stopwords Removal
[‘halah’, ‘apaan’, ‘saya’, ‘sudah’, ‘bayar’, ‘ribu’, ‘di’, ‘tanggal’, ‘juli’, ‘saya’, ‘pakai’, ‘riliv’, ‘karena’, ‘merasa’, ‘benar’, ‘butuh’, ‘konsultasi’, ‘ke’, ‘psikolog’, ‘saat’, ‘itu’, ‘juga’, ‘karena’, ‘saya’, ‘sudah’, ‘ngedown’, ‘parah’, ‘terus’, ‘dikasih’, ‘pilihan’, ‘jadwal’, ‘konsultasi’, ‘sama’, ‘psikolog’, ‘nya’, ‘di’, ‘tanggal’, ‘atau’, ‘agustus’, ‘saya’, ‘minta’, ‘di’, ‘tanggal’, ‘nya’, ‘dimajukan’, ‘susah’, ‘katanya’, ‘gila’, ‘aja’, ‘woy’, ‘kalau’, ‘orang’, ‘yang’, ‘sudah’, ‘stress’, ‘parah’, ‘disuruh’, ‘tunggu’, ‘hari’, ‘sudah’, ‘keburu’, ‘bunuh’, ‘diri’, ‘itu’, ‘orang’, ‘orang’, ‘sudah’, ‘bayar’, ‘itu’, ‘tandanya’, ‘dia’, ‘lagi’, ‘benar’, ‘butuh’, ‘saat’, ‘itu’, ‘juga’, ‘kalau’, ‘disuruh’, ‘tunggu’, ‘hari’, ‘sih’, ‘oke’, ‘lah’, ‘ini’, ‘hari’, ‘situ’, ‘mikir’]	[‘halah’, ‘apaan’, ‘bayar’, ‘ribu’, ‘tanggal’, ‘juli’, ‘pakai’, ‘riliv’, ‘merasa’, ‘benar’, ‘butuh’, ‘konsultasi’, ‘psikolog’, ‘ngedown’, ‘parah’, ‘dikasih’, ‘pilihan’, ‘jadwal’, ‘konsultasi’, ‘psikolog’, ‘tanggal’, ‘agustus’, ‘minta’, ‘tanggal’, ‘dimajukan’, ‘susah, ‘katanya’, ‘gila’, ‘stress’, ‘parah’, ‘disuruh’, ‘tunggu’, ‘hari’, ‘keburu’, ‘bunuh’, ‘diri’, ‘bayar’, ‘tandanya’, ‘benar’, ‘butuh’, ‘disuruh’, ‘tunggu’, ‘hari’, ‘situ’, ‘mikir’]

5) Stemming

After stopwords removal, the words in the sentence with affixes are converted into basic words without affixes by careful stemming.
The difference before and after the text is stemmed can be seen in Table 10.

TABLE X
COMPARISON OF SAMPLE DATA BEFORE AND AFTER STEMMING

Before Stemming	After Stemming
[‘halah’, ‘apaan’, ‘bayar’, ‘ribu’, ‘tanggal’, ‘juli’, ‘pakai’, ‘riliv’, ‘merasa’, ‘benar’, ‘butuh’, ‘konsultasi’, ‘psikolog’, ‘ngedown’, ‘parah’, ‘dikasih’, ‘pilihan’, ‘jadwal’, ‘konsultasi’, ‘psikolog’, ‘tanggal’, ‘agustus’, ‘minta’, ‘tanggal’, ‘dimajukan’, ‘susah, ‘katanya’, ‘gila’, ‘stress’, ‘parah’, ‘disuruh’, ‘tunggu’, ‘hari’, ‘keburu’, ‘bunuh’, ‘diri’, ‘bayar’, ‘tandanya’, ‘benar’, ‘butuh’, ‘disuruh’, ‘tunggu’, ‘hari’, ‘situ’, ‘mikir’]	halah apa bayar ribu tanggal juli pakai riliv rasa benar butuh konsultasi psikolog down parah kasih pilih jadwal konsultasi psikolog tanggal agustus minta tanggal maju susah kata gila stress parah suruh tunggu hari buru bunuh diri bayar tanda benar butuh suruh tunggu hari situ mikir

E. Synthetic Minority Over-sampling Technique (SMOTE)

TABLE XI
BEFORE AND AFTER SMOTE

Before SMOTE		
Positive	**Neutral**	**Negative**
1812	477	623
After SMOTE		
Positive	**Neutral**	**Negative**
1812	1812	1812

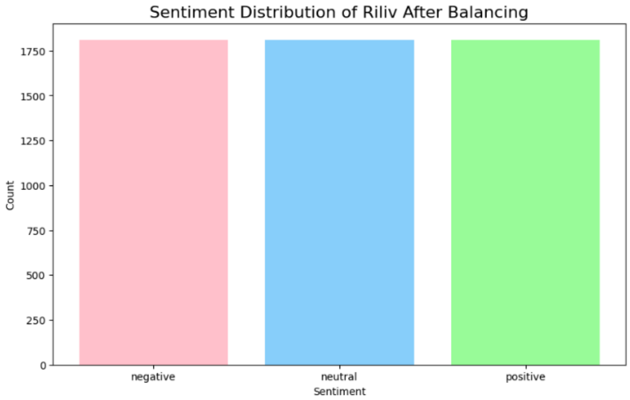


Figure 16. Sentiment Distribution of Riliv After Balancing

F. Data Transformation

Splitting the data ratio for LDA:

```
# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

For Classification with Non-Transformer:

```
test_sizes = [0.2, 0.3, 0.4]
ratios = ["80:20", "70:30", "60:40"]
```

As for the results of weighting data from TF-IDF with the calculation of how often tokens appear in documents and produce a value so that it can be used as a weight as follows.

	aamiin	ada	adalah	admin	agak	agar	agustus	ahli	ajak	ajar	...	work	worth	wujud	yaitu	yakin	yaudah	your	youtube	yuk	yuu
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.207122	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

5 rows × 1000 columns

Figure 17. TF-IDF Output

For Classification with Transformer :

```
ratios = [(8, 1, 1), (7, 1, 2), (7, 2, 1), (6, 3, 1), (6, 1, 3), (6, 2, 2)]
all_histories = {}
```

Then weigh the data with transformers.

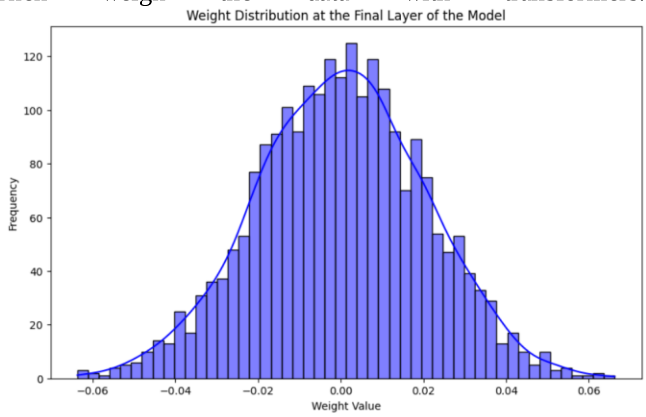


Figure 18. Transformer Final Layer Output

G. Topic Clustering with LDA (Latent Dirichlet Allocation)

1) Apply LDA

```
Top words for each topic:
Topic 1 :  ['pakai', 'cara', 'kode', 'mantap', 'gimana', 'tidak', 'tolong', 'tingkat', 'boleh', 'curhat']
Topic 2 :  ['bagus', 'aplikasi', 'bantu', 'suka', 'terimakasih', 'riliv', 'tenang', 'manfaat', 'kasih', 'terima']
Topic 3 :  ['bayar', 'coba', 'dulu', 'baru', 'bintang', 'gratis', 'tidak', 'psikolog', 'konseling', 'curhat']
Topic 4 :  ['bantu', 'tidak', 'buka', 'aplikasi', 'kenapa', 'susah', 'daftar', 'keren', 'unduh', 'gagal']
```

Figure 19. Top Words for Each Topic

2) Topic Interpretation

The following maps the topic value to the specified label.

```
# Map topic values to the specified labels
topic_labels = {
    0: "ACCESS SUPPORT",
    1: "MEDITATION FEATURES",
    2: "COUNSELING SERVICES",
    3: "USER INTERFACE"
}
```

3) Results of Each Aspect

Print documents with assigned topics.

TABLE XII
ASPECT-BASED REVIEW

ACCESS SUPPORT
sayang kakak suka bikin live tiktok rajin jawab baca tanya komentar satu cinta banget
MEDITATION FEATURES
aplikasi bantu buat tidur nyenyak stabil emosi sedikit bagus terimakasih riliv
COUNSELING SERVICES

semua cara umum kecuali konsultasi jam masuk lambat tidak profesional jawab kesan malah salah mending cari aplikasi konsultasi psikolog online untung cuma ambil paket asa rugi pakai uninstall
USER INTERFACE
buruk baru isi jurnal niat nulis selesai sampai page lama nulis pencet next page lambat selesai tiba lambat akhir tidak simpan hari mood hancur tambah nulis jurnal belum tidur gagal makin rumit

Here are the results of the distribution of the 4 aspects. The most is the User Interface and the least is Access Support.

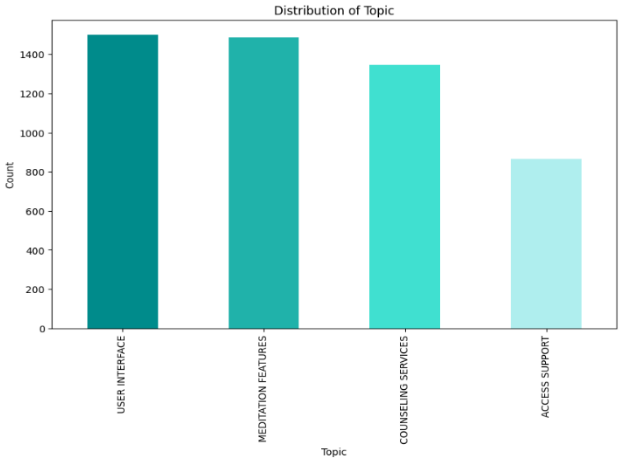


Figure 20. Distribution of Topic

H. Classification

1) Classification with Non-Transformer

TABLE XIII
CLASSIFICATION WITH NON-TRANSFORMER

Classification Report	**Naïve Bayes**			
	Train : Test (8 : 2)			
	Accuracy	Precision	Recall	F1-Score
Negative	85%	89%	85%	87%
Neutral	85%	77%	85%	81%
Positive	85%	89%	85%	87%
	Train : Test (7 : 3)			
Negative	84%	91%	83%	86%
Neutral	84%	76%	84%	80%
Positive	84%	86%	85%	85%
	Train : Test (6 : 4)			
Negative	84%	91%	83%	87%
Neutral	84%	77%	84%	80%
Positive	84%	85%	85%	85%
	Logistic Regression			
	Train : Test (8 : 2)			
Negative	87%	92%	88%	90%
Neutral	87%	78%	89%	83%
Positive	87%	92%	84%	88%
	Train : Test (7 : 3)			
Negative	87%	94%	87%	90%
Neutral	87%	78%	89%	83%
Positive	87%	90%	84%	87%
	Train : Test (6 : 4)			
Negative	86%	93%	87%	90%
Neutral	86%	78%	89%	83%
Positive	86%	88%	83%	86%
	Random Forest			
	Train : Test (8 : 2)			
Negative	**94%**	**97%**	**97%**	**97%**
Neutral	**94%**	**87%**	**99%**	**93%**
Positive	**94%**	**98%**	**85%**	**91%**
	Train : Test (7 : 3)			
Negative	93%	97%	97%	97%
Neutral	93%	87%	98%	92%
Positive	93%	96%	85%	90%
	Train : Test (6 : 4)			
Negative	91%	96%	94%	95%
Neutral	91%	85%	97%	90%
Positive	91%	95%	83%	89%

2) Classification with Transformer

TABLE XIV
CLASSIFICATION WITH TRANSFORMER

Classification Report	**IndoBERT**			
	Train : Validation : Test (8 : 1: 1)			
	Accuracy	Precision	Recall	F1-Score
Negative	95%	99%	97%	98%
Neutral	95%	87%	99%	93%
Positive	95%	99%	91%	95%
	Train : Validation : Test (7 : 1: 2)			
Negative	**95%**	**99%**	**98%**	**98%**
Neutral	**95%**	**89%**	**99%**	**94%**
Positive	**95%**	**98%**	**88%**	**93%**
	Train : Validation : Test (7 : 2: 1)			
Negative	95%	99%	97%	98%
Neutral	95%	89%	99%	94%
Positive	95%	97%	89%	93%
	Train : Validation : Test (6 : 1 : 3)			
Negative	94%	98%	97%	98%
Neutral	94%	86%	98%	91%
Positive	94%	99%	87%	92%

	Train : Validation : Test (6 : 3: 1)			
Negative	93%	99%	91%	95%
Neutral	93%	89%	93%	91%
Positive	93%	90%	94%	92%
	Train : Validation : Test (6 : 2: 2)			
Negative	94%	98%	96%	97%
Neutral	94%	90%	95%	92%
Positive	94%	95%	92%	93%

I. Evaluation (ABSA)

Figure 21. Wordcloud for Positive Sentiment

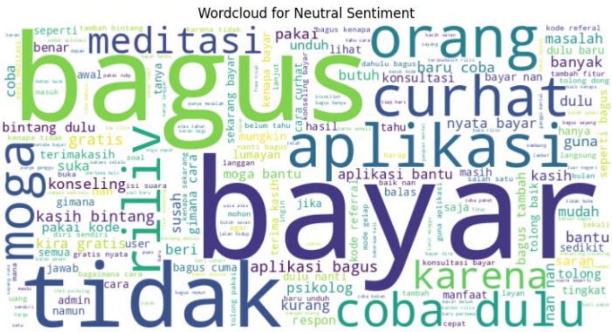


Figure 22. Wordcloud for Neutral Sentiment

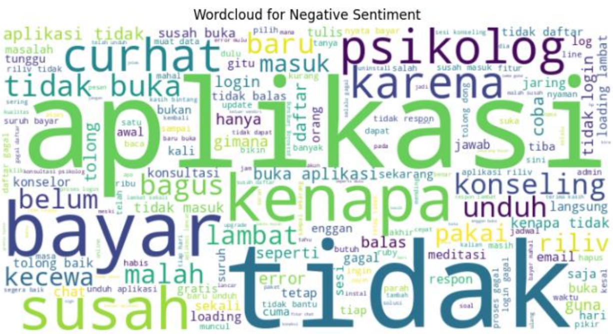
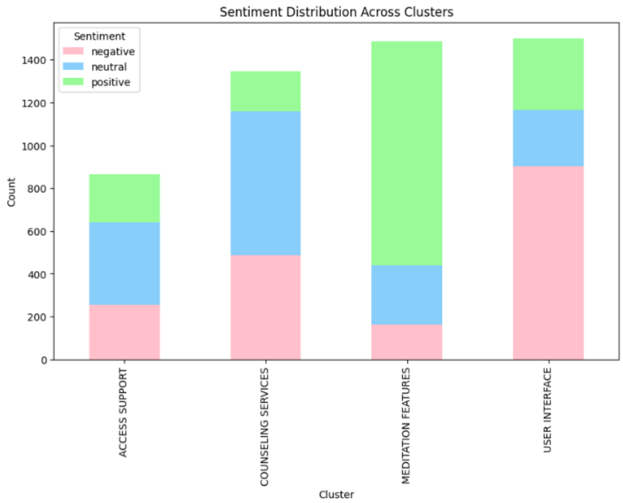


Figure 23. Wordcloud for Negative Sentiment



sentimen	negative	neutral	positive
aspect			
ACCESS SUPPORT	255	386	224
COUNSELING SERVICES	488	670	187
MEDITATION FEATURES	162	276	1049
USER INTERFACE	904	261	336

Figure 24. Sentiment Distribution Across Cluster

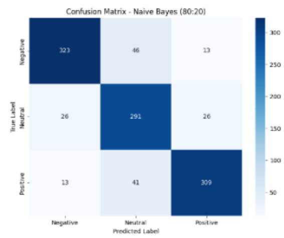


Figure 25. Confusion Matrix Naïve Bayes (80 : 20)

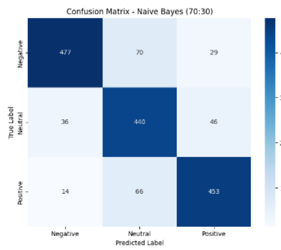


Figure 26. Confusion Matrix Naïve Bayes (70 : 30)

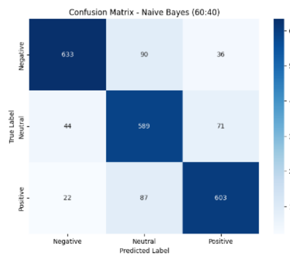


Figure 27. Confusion Matrix Naïve Bayes (60 : 40)

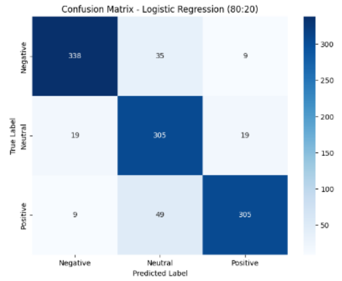


Figure 28. Confusion Matrix Logistic Regression (80 : 20)

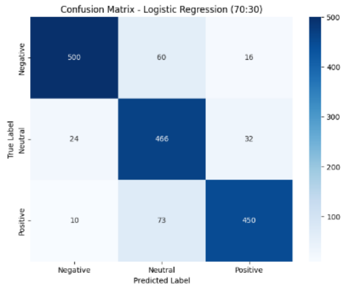


Figure 29. Confusion Matrix Logistic Regression (70 : 30)

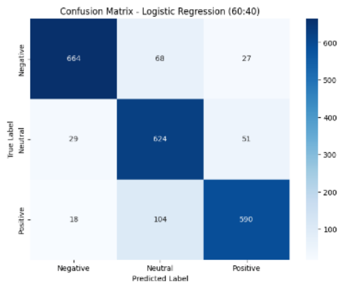


Figure 30. Confusion Matrix Logistic Regression (60 : 40)

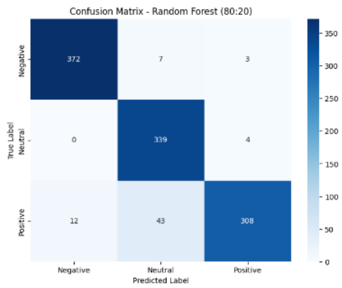


Figure 31. Confusion Matrix Random Forest (80 : 20)

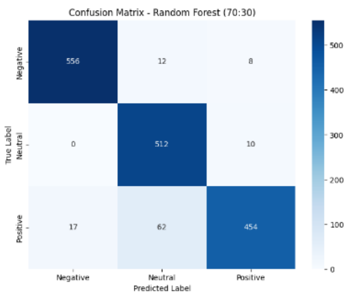


Figure 32. Confusion Matrix Random Forest (70 : 30)

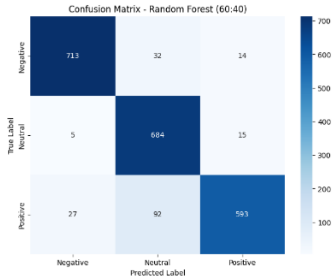


Figure 33. Confusion Matrix Random Forest (60 : 40)

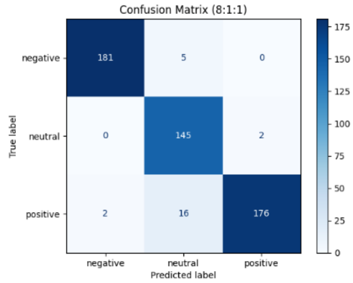


Figure 34. Confusion Matrix IndoBERT (8 : 1 : 1)

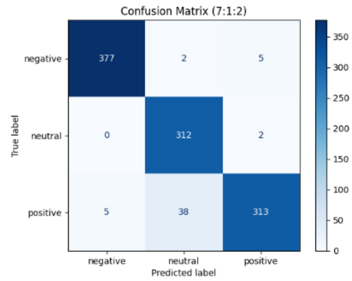


Figure 35. Confusion Matrix IndoBERT (7 : 1 : 2)

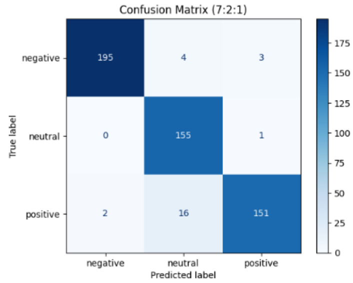


Figure 36. Confusion Matrix IndoBERT (7 : 2 : 1)

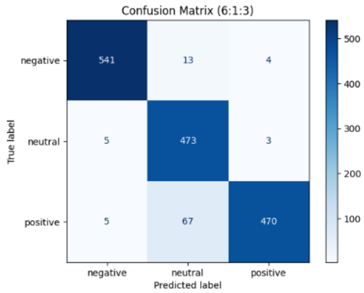


Figure 37. Confusion Matrix IndoBERT (6 : 1 : 3)

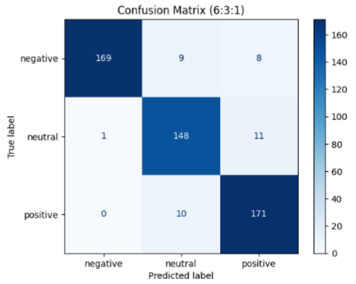


Figure 38. Confusion Matrix IndoBERT (6 : 3 : 1)

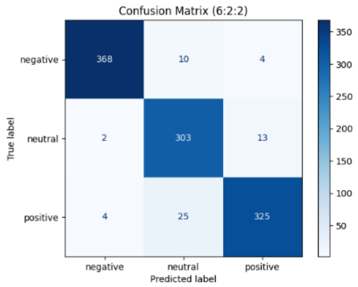


Figure 39. Confusion Matrix IndoBERT (6 : 2 : 2)

IV. CONCLUSION

The results of research on Aspect-Based Sentiment Analysis can be concluded that from 3068 App Store and Play Store reviews of the Riliv Application from September 2015 to December 2024, 4 aspects were obtained, namely Access Support, Counseling Services, Meditation Features, and User Interface. The majority of the results were negative on User Interface, neutral on Counseling Services, and positive on Meditation Features based on the results of applying the LDA (Latent Dirichlet Allocation) algorithm. As for the results of sentiment classification of user review data using a classification comparison scenario on Non-Transformers (Naïve Bayes, Logistic Regression, and Random Forest) with Transformers (IndoBERT), the best performance is generated by the IndoBERT algorithm at a training data division ratio of 70%, validation data 10%, and testing data 20% with Accuracy (negative 95%, neutral 95%, positive 95%), Precision (negative 99%, neutral 89%, positive 98%), Recall (negative 98%, neutral 99%, positive 88%), and F1-Score (Negative 98%, neutral 94%, and positive 93%), train loss 0.0542, validation loss 0.1244. By using epoch 15, lr 1e-5, weight with Adam optimizer.

From the confusion matrix that has 1054 samples, the correct negative prediction result is 377 the rest are neutral, the correct neutral prediction is 312 the rest are positive and negative, and the correct positive prediction is 313 the rest are negative and neutral.

The information results from the classification are obtained about the satisfaction and complaints faced by users of the Riliv App. Positive sentiment refers to user satisfaction. Based on the words that appear on the word cloud, users are *aplikasi, bantu, bagus, riliv, meditasi, konseling, nyaman, tenang, suka, banyak*. Users feel helped by counseling services or meditation features that are good and steady and feel that there are friends to confide in who are comfortable and calm. Then the user is happy with the Riliv application so thank you very much.

Neutral sentiments are based on words that appear *coba, dulu, tidak, aplikasi, bagus, bayar, orang, curhat, riliv, moga. R*efer to satisfaction but there are few complaints and there are users who just want to try and give a rating first before trying. The few complaints build on the polite use of help, please, after praise.

Negative sentiments are based on words that appear *aplikasi, tidak, bayar, susah, psikolog, unduh, kecewa, login, lambat, loading,* to refer to complaints from disappointed users. Users regret that there are paid premium provisions for Riliv access, are disappointed that payment transactions are difficult, have difficulty downloading applications and using applications that load slowly even though they have just downloaded, failed registration, psychologists or counselors who do not enter or do not reply to counseling sessions, and login errors so it is difficult to open the application.

Therefore, the good things are maintained, and the bad things can be improved and improved by Riliv. So Riliv can help reduce mental health cases to be more optimal and can consistently become the No.1 mental health application in Indonesia.

REFERENCES

[1] Queensland Centre for Mental Health Research, “I-NAMHS: Indonesia-National Adolescent Mental Health Survey,” May 2023. Accessed: Jan. 01, 2025. [Online]. Available: https://qcmhr.org/~teoqngfi/outputs/reports/12-i-namhs-report-bahasa-indonesia

[2] N. Wusqa and S. Novitayani, “Risiko Bunuh Diri Pada Mahasiswa Tingkat Akhir,” *Jurnal Ilmiah Mahasiswa Fakultas Keperawatan*, vol. 6, no. 2, 2022.

[3] APJII, “Survei Penetrasi Internet Indonesia 2024,” *Asosiasi Penyelenggara Jasa Internet Indonesia*, 2024.

[4] R. Saphira, “Aplikasi Riliv Berbasis Synchronous Sebagai Pelayanan Dalam Cyber Counseling Di Era Digital,” *Jurnal Ilmu Pendidikan, Psikologi, Bimbingan dan Konseling*, vol. 3, no. 2, Dec. 2022, doi: 10.52802/crm.v3i2.483.

[5] E. H. Muktafin, Pramono, and Kusrini, “Sentiments analysis of customer satisfaction in public services using K-nearest neighbors algorithm and natural language processing approach,” *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, vol. 19, no. 1, pp. 146–154, Feb. 2021, doi: 10.12928/TELKOMNIKA.v19i1.17417.

[6] S. Kushwah and S. Das, “Sentiment Analysis of Big-Data in Healthcare: Issue and Challenges,” *IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, pp. 658–663, 2020, doi: 10.1109/ICCCA49541.2020.9250841.

[7] E. Saad *et al.*, “Determining the Efficiency of Drugs under Special Conditions from Users’ Reviews on Healthcare Web Forums,” 2021, *Institute of Electrical and Electronics Engineers*. doi: 10.1109/ACCESS.2021.3088838.

[8] S. A. P. Perdana, T. B. Aji, and R. Ferdiana, “Aspect Category Classification dengan Pendekatan Machine Learning Menggunakan Dataset Bahasa Indonesia (Aspect Category Classification with Machine Learning Approach Using Indonesian Language Dataset),” *Jurnal Nasional Teknik Elektro dan Teknologi*

Informasi, vol. 10, no. 3, pp. 229–235, 2021, doi: 10.22146/jnteti.v10i3.1819.

[9] P. Zambrano *et al.*, "Technical Mapping of the Grooming Anatomy Using Machine Learning Paradigms: An Information Security Approach," *nstitute of Electrical and Electronics Engineers*, vol. 7, pp. 142129–142146, 2019, doi: 10.1109/ACCESS.2019.2942805.

[10] R. A. Rahman, V. H. Pranatawijaya, and N. N. K. Sari, "Analisis Sentimen Berbasis Aspek pada Ulasan Aplikasi Gojek," *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 4, no. 1, Jun. 2024, doi: 10.24002/konstelasi.v4i1.8922.

[11] P. T. Sejati, F. Al Zami, A. Marjuni, H. Indrayani, and I. D. Puspitarini, "Aspect-Based Sentiment Analysis for Enhanced Understanding of 'Kemenkeu' Tweets," *Journal of Applied Informatics and Computing (JAIC)*, vol. 8, no. 2, pp. 487–498, 2024.

[12] D. D. Ayani, H. S. Pratiwi, and H. Muhardi, "Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace," *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 7, no. 4, pp. 45–54, 2019, doi: 10.26418/justin.v7i4.30930.

[13] R. Rani and D. K. Lobiyal, "Performance evaluation of text-mining models with Hindi stopwords lists," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6 Part A, pp. 2771–2786, 2022, doi: 10.1016/j.jksuci.2020.03.003.

[14] A. Luthfiarta, J. Zeniarja, and A. Salam, "Integrasi Peringkas Dokumen Otomatis Dengan Algoritma Latent Semantic Analysis (LSA) Pada Peringkas Dokumen Otomatis Untuk Proses Clustering Dokumen," *Techno.COM*, vol. 13, no. 1, pp. 61–68, 2014, doi: 10.33633/tc.v13i1.543.

[15] N. Sulistiyowati and M. Jajuli, "Integrasi Naive Bayes Dengan Teknik Sampling SMOTE Untuk Menangani Data Tidak Seimbang," vol. 14, no. 1, 2020, doi: 10.25134/nuansa.v14i1.2411.

[16] D. Septiani and I. Isabela, "Analisis Term Frequency Inverse Document Frequency (TF-IDF) Dalam Temu Kembali Informasi Pada Dokumen Teks," vol. 1, no. 2, Mar. 2022.

[17] N. A. C.A, D. H. Citra, W. Purnama, C. Nisa, and A. R. Kurnia, "The Implementation of Naïve Bayes Algorithm for Sentiment Analysis of Shopee Reviews on Google Play Store," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 2, pp. 47–54, Apr. 2022, doi: 10.57152/malcom.v2i1.195.

[18] A. M. Mantika, A. Triayudi, and R. T. Aldisa, "Sentiment Analysis on Twitter Using Naïve Bayes and Logistic Regression for the 2024 Presidential Election," *SaNa: Journal of Blockchain, NFTs and Metaverse Technology*, vol. 2, no. 1, pp. 44–55, Feb. 2024, doi: 10.58905/sana.v2i1.267.

[19] T. F. Basar, D. E. Ratnawati, and I. Arwani, "Analisis Sentimen Pengguna Twitter terhadap Pembayaran Cashless menggunakan Shopeepay dengan Algoritma Random Forest," *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, vol. 6, no. 3, pp. 1426–1433, Mar. 2022.

[20] P. Sayarizki, Hasmawati, and H. Nurrahmi, "Implementation of IndoBERT for Sentiment Analysis of Indonesian Presidential Candidates," *Ind. Journal on Computing*, vol. 9, no. 2, pp. 61–72, 2024, doi: 10.34818/indojc.2024.9.2.934.