p–ISSN: 2723 - 6609 e-ISSN: 2745-5254
Vol. 5, No. 8 August 2024 http://jist.publikasiindonesia.id/

Classification of Indonesia False News Detection Using Bertopic and Indobert

Veren Prisscilya[1*], Abba Suganda Girsang[2]
Universitas Bina Nusantara, Indonesia
Email: veren.prisscilya@binus.ac.id[1*], agirsang@binus.edu[2]

*Correspondence

Keywords: hoax, deep learning, per topic, indoor, text classification.

ABSTRACT

In the current global era, the development of technology and information is very rapid, so it is very easy to get information/news from the internet. Because of the ease of getting this information, there is a lot of circulating fake news (hoaxes), the news is not filtered so anyone can spread news that is not clear in content. This can lower a person's credibility in the professional world, cause division, threaten physical and mental health, and can also result in material losses. Based on this, to stop the spread of hoaxes is to detect them as early as possible and block them. This detection can use deep learning methods which are also one of the architectures of transformers, namely a combination of BERTopic which is used to find important words from the news narrative, then the words are combined into the narrative and classified using Indo Bidirectional Encoder Representation from Transformer (IndoBERT). For experiments, the author uses a dataset taken from the kaggle.com website entitled Indonesia False News (HOAX) dataset. This study uses a learning rate of 1e-5, a batch size of 16 and using 5 epochs so that the f1-Score results are 92% for validation data and 91% for testing data.

Introduction

False News has become a serious problem in Indonesia, especially in the social media era. False News can have a significant negative impact on society, from influencing public opinion to triggering social conflict (Chien, 2017).

Hoaxes are defined as fake news that aims to convince readers to believe the fake news. Every day and even every minute a lot of news appears and spreads on social media, these hoaxes are usually made by groups or individuals who aim for personal interests and can also be caused by certain factors. Hoaxes are being spread in various media in Indonesia both from print media and online media (Hartipa Iskandar, 2021).

The purpose of making hoaxes is to persuade, manipulate, and influence readers to do the opposite or prevent correct actions. Usually using threats, misleading, or making readers believe things that are not real and cannot be confirmed (Teresia Arum, 2016).

Hoaxes turned out to be more favoured by the public. This is evidenced by the ease with which hoaxes are spread by sharing techniques, especially in social media, such as Instagram LINE, WhatsApp, Facebook, and other media. (Hartipa Iskandar, 2021) According to research by the Ministry of Communication and Information (Kominfo), the cause is because the mainstream media is no longer in favour of the community, and the mainstream media is also less sensitive in absorbing the aspirations of the outside community and tends to be a forum for media owners. So that makes people look for alternative media to satisfy their information needs by looking for sites containing these hoaxes (Kaimal et al., 2016).

Identifying hoaxes is not always easy. Hoaxes are often made to resemble real news, with catchy titles and packaged with convincing narratives. This makes it difficult for people to distinguish which news is true and which news is false (NGABITO, 2020).

Therefore, to stop the spread of hoaxes is to detect them as early as possible and block them. The detection is proposed by using deep learning methods with BERTopic and IndoBERT (Indonesia Bidirectional Encoder Representation from Transformers) models. The selection of this method is based on the BERT method which has obtained new state-of-the-art on eleven problems in NLP tasks, one of which is text classification. [4] BERTopic is a topic modelling technique based on BERT and TF-IDF in creating clusters that produce easily interpretable topics and important words that describe topics. The BERTopic model uses word context (Muna, 2022). For BERTopic document clustering, two UMAP algorithms are used to reduce the dimension of the word embedding results and the HDBSCAN algorithm for document clustering. The document clustering process in the BERTopic model is based on the class-based TF-IDF variant value (c-TF-IDF) in determining the uniqueness of a document compared to other documents. (Guridno, Azimah, & Ningsih, 2024) Based on this, BERTopic is expected to support this research to find topics automatically. IndoBERT uses a transformer mechanism, where this mechanism works to learn the relationship between each word in a sentence. IndoBERT uses two mechanisms, namely encoder to read input and a decoder to generate predictions. This research only uses a dataset from Kaggle entitled Indonesia False News (Hoax). The dataset was obtained in the 2015-2020 timeframe so that it represents the situation at that time. An example of news that could be a hoax in the present is "Anis is the Governor of DKI Jakarta", in 2017-2020 it was true news, but in the present, it could be a hoax. The truth or untruth of the news depends on the dataset used, it is necessary to update the dataset to ensure the latest information. The dataset is dynamic and can change over time. Therefore, the results of this study may not fully reflect the current situation.

In their journal (Tandijaya, Liliana, & Sugiarto, 2021) propose BERT classify online news portals, where data obtained from Indonesia news portals, namely from kompas.com and sindonews.com, using web scraping with the help of the BeautifulSoup library, where the data is 6309 data and divided into 80% training data and 20% testing data, then the data will be preprocessed and then classified using the BERT model, wherein the classification, data is converted into an input that can be read by the model,

followed by making a data loader to help speed up data retrieval. Next, enter the modelling stage. After the model is created, training and evaluation of the model will be carried out. The pre-trained model will be used to classify new news. The evaluation results obtained with the model built obtained an accuracy of 87.54% by tuning parameters with a learning rate configuration of 5e-05, dropout 0.1, and epoch 4.

Several methods for detecting fake news in Indonesia were proposed by (Isa, Asrori, & Muharini, 2022) namely TF-IDF, SVM, Naïve Bayes, and Indobert methods. TF-IDF combined with SVM where TF-IDF is used for feature extraction while SVM is used for classification obtained precision, recall, F1-Score and Accuracy results of 90.00%, TF-IDF combined with Naïve Bayes where TF-IDF is used for feature extraction while naïve bayes for classification precision, F1-Score and accuracy of 83.00% and Recall of 85.00%, for the use of the IndoBERT model obtained precision, recall, F1-score and accuracy results of 94.66%. IndoBERT's weakness is that it requires a lot of data and takes a lot of time to process its data compared to the TF-IDF + SVM and TF-IDF + Naïve Bayes models. IndoBERT takes 15 minutes for training time, which is three times longer than TF-IDF + SVM and fifteen times longer than TF-IDF + Naïve Bayes.[9]

There is also a proposal for the Naive Bayes method to detect fake news in Indonesia by combining it with TF features. After carrying out the pre-processing data stage, the dataset will calculate the frequency of word occurrences in the document. After getting the frequency of words in all training documents, calculate the probability value P(ci). Each group has a probability value calculated based on the number of documents in the category per document. Next, sample testing was carried out, followed by static testing, which was carried out on 600 news stories, which resulted in an average accuracy of 82.60%, and finally, dynamic testing, which was carried out by entering the news content into the system. Of the 60 news stories tested, 41 produced the same news classification as the manual mark, and 19 produced a different classification with the manual mark. The percentage of valid results is 68.33%, and hoaxes are 31.67%.

A comparison between the LSTM model and the IndoBERT model was carried out to detect fake news taken from Twitter, followed by a preprocessing process to clean the dataset. After that, the training process is carried out with the LSTM model using Word2Vec, while the IndoBERT model does not need to do the Word2Vec process and directly divides the data using 10 K-Fold Cross Validation. After both models have been processed, the accuracy results of the models are compared. The results show that the IndoBERT model has better performance than the LSTM model. The IndoBERT model achieved an average accuracy of 92.07%, while the LSTM model achieved an average accuracy of 87.54%.

To analyze and predict the validity of news in Indonesian (Nayoga, Adipradana, Suryadi, & Suhartono, 2021) using a supervised text classification approach. In this journal, several deep learning models are proposed, including LSTM, BI-LSTM, GRU, BI-GRU, and 1D-CNN as well as two conventional classifications, namely SVM and naïve Bayes. Of the five deep learning models and two conventional classifications, the

best model for detecting fake news in the Indonesian language is 1D-CNN with 97.90% accuracy.

Research conducted by (Hanifa, Isa, & Mohamad, 2021) conducted a comparison between the LSTM and GRU (RNN) methods for the classification of fake news in Indonesian. The data used for original news is a collection of various online news portals circulating in Indonesia such as detik.com, tribunnews.com, and liputan6.com, while fake news is taken from the turnback hoax.id portal. Thus, a total of 600 news articles were obtained, with a total of 372 original news articles and 228 fake news articles. The data is then classified using the LSTM-RNN model with an accuracy of 38% while using the GRU-RNN model with an accuracy of 38%. Because it gets low accuracy results, then hyperparameter tuning is carried out, obtaining the best accuracy results with epochs: 15, optimizer: rmsprop, and batch size: 64 obtained an accuracy result of 72.50% from the LSTM-RNN model, while GRU-RNN obtained an accuracy result of 64.20%.

Research conducted by (Ramadhan et al., 2022) using Random Forest and logistic regression methods, with the dataset used from the Kaggle dataset using 6,560 news titles, with data division of 70% training data and 30% testing data, using the Random Forest (Entropy) model obtained an accuracy result of 84%, random forest (Gini) obtained an accuracy result of 83%, Logistic Regression obtained an accuracy result of 77%.[12]

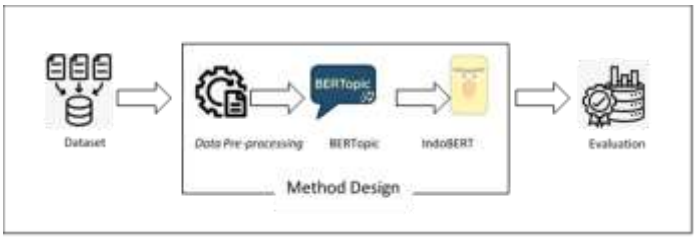
The journal made by (Kurniati et al., 2021) uses a dataset of social media news from the kumparan.com site where the data is taken using a library from Python, namely BeautifulSoap, using the naïve Bayes model, the accuracy result is 81%.[13]

Research conducted by (Haumahu, Permana, & Yaddarabullah, 2021) uses datasets taken from several online news sites, namely kompas.com, detik.com, cnnindonesia.com, liputan6.com, and turnbackhoax.id which was taken in the 2015-2020 time span of 500 news, which consists of 250 real news and 250 fake news, the data is divided into 80% training data, 20% testing data using the Extreme Gradient Boosting (XGBoost) method obtained an accuracy of 92%.

In the journal Nurhikam et al (2023) detecting fake election news using the random forest algorithm, using a dataset of 859 news where 670 are factual while 189 are fake news, the data uses data from online news platforms, namely detik.com, liputan6.com, okezone.com, kominfo, turnbackhoax.id. The data is trained using the Random Forest model with balanced training and testing data obtained balanced accuracy results obtained accuracy results of 84.55%.

Based on previous research, it is concluded that 1D-CNN has the highest accuracy result of 97.90% followed by IndoBERT with 94.66%. In this study, the author will use IndoBERT as a model used for the classification of fake news in Indonesian, where IndoBERT can learn complex relationships between words.

Research Methods



Research Stages

The research stages have been illustrated in Fig. 1 in the figure it can be seen that several stages will be carried out in this study, starting from dataset collection, and design of the proposed method, and will be evaluated in the final stage to determine whether the method design can be used or not. Each of these stages will be explained in detail in the following subsections.

Dataset

The dataset used in this research is the Indonesia False News (HOAX) Dataset, where this dataset is obtained by downloading from the web www.kaggle.com the dataset is 4,231 training data from the dataset there are 3,465 labeled 1 (fake news), and as many as 766 labeled 0 (real news).

Data Pre-Processing

From the previously obtained dataset, data preprocessing is carried out because the data obtained is not structured and consistent so it is necessary to do some data preprocessing, among others:

Fig. 2 Data Pre-Processing

1) Normalization of non-standard words: this stage is carried out to normalize the data in the previous dataset which has non-standard words converted into standard words so that they are by the Enhanced Spelling (EYD). This is done to make it easier for the system to interpret the words from the dataset so that they do not have ambiguous meanings.
2) Case folding: this stage is carried out because the data obtained is not structured and consistent in the use of uppercase and lowercase.

3) Data cleaning: this stage is carried out to remove several symbols, characters, enter, newline, excess spaces, emojis, and repetition of characters in the sentence.
4) Data splitting: used to divide training data, validation data and test data. The data division is 80% training data, 10% validation data, and 10% test data.
5) Tokenization: this stage uses the IndoBERT transformers model, where IndoBERT tokenization represents tokens that represent each word. Each token will then be represented

BERTopic

Fig. 3 Topic Categorization

After pre-processing the dataset, we continued with topic categorization using BERTopic (Fig. 3) and searched for important words in the news. BERTopic uses clustering techniques to group topics from words that have similar meanings. Where the clustering technique uses HDBSCAN which is used to group sentence embeddings based on similarity in meaning, each cluster represents the topic obtained.

To get a topic, BERTopic identifies words that represent each topic. After obtaining the TF-IDF calculation results, the words are selected based on their high TF-IDF scores. TF-IDF measures how important a word in a news story is compared to all the news stories in the data set. BERTopic outputs the important words for each topic as well as the TF-IDF score of those words.

The following is an example of topic clustering from 1 topic group using important words that have been calculated with high TF-IDF using BERTopic, as follows:
(‘indonesia’, 0.4196265), (‘ternyata’, 0.41169137), (‘jakarta’, 0.36498433), (‘kepada’, 0.3534637), (‘kalau’, 0.34819508), (‘beberapa’, 0.3468621, (‘sampai’, 0.3382529), (‘nama’, 0.32713622), (‘orang’ 0.30558452), (‘tentara’, 0.29993442)

After obtaining the results of topic grouping, the next step is to combine the important words obtained with the narrative.

The following is an example of a sentence that will be grouped to get a topic:

'Dear Mr. President of the Republic of Indonesia, I want to ask what is the difference between us and the Rohingya refugees, they give us temporary viewing assistance for indigenous Indonesians, Bali, one of the provinces that provides the largest income in Indonesia, but with the current incident it seems that you have forgotten with us, because the Rohingya are the same people as you, so special they are just what I want to ask' Keywords/Topics of the sentence are: Indonesia, apparently, Jakarta, to, if, several, until, name, person, tentara.

Combination of Topic and Narrative

Next, the merging of topic and narrative is done, for example as follows:
‘<uniq>Indonesia turns out to be Jakarta to some of the names of soldiers <UNIQ> to the honorable President of the Republic of Indonesia, I want to ask you what is the difference between us and the Rohingya refugees, they give you temporary viewing assistance for indigenous Indonesians, Bali, one of the provinces that provides the largest income in Indonesia, but with the current incident, it seems that you have forgotten with us what because the Rohingya are the same people as you, so they are so special, that's it. all I want to ask'

After combining important words and narratives, then tokenization is carried out using IndoBERT tokenization, tokenization and combining important words with the narrative is carried out on each data that has been divided, namely training data, validation data and test data. Tokenization using IndoBERT represents tokens that represent each word. Each token will then be represented by a vector.
The following is an example of tokenization from the merged data:
[2, 30521, 300, 678, 1369, 455, 1179, 599, 493, 232, 531, 899, 30521, 636, 4925, 166, 493, 5107, 5875, 469, 119, 1085, 5759, 8133, 43, 9648, 719, 857, 8292, 34, 344, 5875, 5108, 5585, 13841, 57, 1831, 1202, 92, 3], [2, 30521, 300, 678, 1369, 455, 1179, 599, 493, 232, 531, 899, 30521, 5580, 15349, 3320, 7970, 1081, 90, 1874, 1871, 19145, 3],

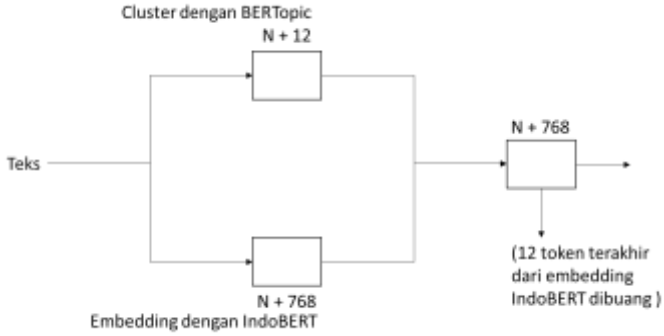


Fig. 4 Word Combinations in Narration

In fig. 4 describes the merging of important words with the narrative in summary, where the text is clustered using BERTopic so that 12 tokens are obtained, of which the tokens consist of 10 important words and 2 others are special tokens located at the beginning and end, while the narrative is embedded using IndoBERT with a total embedding of 768, after getting the results of clustering with BERTopic and embedding with IndoBERT, then the token is merged, because the merger exceeds the limit of IndoBERT embedding, then the last 12 tokens will be discarded to add the BERTopic clustering results at the beginning of the sentence.

IndoBERT

After tokenizing each data, the model can understand the data. then called the IndoBERT model. To get good results, there are several things that need to be determined before classifying fake news, namely:

Optimizer

1. Learning rate

2. Batch size
3. Epoch

A custom dataset is created to create a data loader and then the dataloader is called for training, validation and testing.

Evaluation

After the data is trained, the performance of the model is evaluated with metrics suitable for classification tasks such as accuracy, precision, recall and F1-score to see which method can be used for fake news classification with the best results.

Confusion Matrix

After conducting the evaluation, the next step is to obtain the confusion matrix to assess the performance of the classification model on the validation and test data.

Results and Discussion

Visitors are impressed when they take a gallery tour.

When taking a gallery tour, visitors can enjoy several services provided by the gallery. The services provided for the gallery tour visitors start from the presence of *a welcome drink* in the entrance area which indicates the welcome of gallery guests before starting a series of activities in the gallery tour. The gallery manager provides several kinds of drinks and snacks for visitors to enjoy before starting their activities to take a tour of the gallery on that day. An example observed by the researcher was when there was a gallery tour and workshop by artist Angky Pu which was attended by students from SD Grow Yogyakarta. The same thing was also found when researchers attended the opening of the Riono Tanggul artist exhibition in November this year at the Tirtodipuran Link Gallery in Yogyakarta. The gallery manager through its employees provides several kinds of snacks and drinks and even adds to the satisfaction of visitors, guests are welcome to order drinks at the *gallery's Ooze Bar* for free (Table 1, Figures 1.1 and 1.2).

The series of activities after visitors are given a *welcome drink* service is to follow the gallery guide to go around enjoying the artworks exhibited in the gallery. The guiding officer is the *public relations section* of the gallery because the *public relations department* is the one who knows the most about the products of the works being exhibited and can explain the intentions and objectives of the artist in his works. In the researcher's observation while participating in the gallery tour and Angky Pu's workshop, the role of *public relations* in the gallery tour is very important to give a good impression to the gallery visitors at that time were the students of SD Grow Yogyakarta (Table 1, Figure 1.3). *Public relations* is responsible for explaining to visitors about the existing works, and must also be able to attract visitors to actively interact so that the gallery tour feels warmer and more memorable for visitors who take part in the tour activities that day. The same thing was also observed when researchers participated in the opening of the Riono Embankment exhibition. In the opening ceremony, visitors who came had the opportunity to take a tour of the gallery guided directly by the artists from the works on display. Visitors can not only enjoy the works on display but also interact directly with the existing artists so that it becomes a satisfaction for visitors who come at the opening

of the exhibition (Table 1, Figure 1.4). *Public relations* also plays a role in publicizing the exhibition which makes visitors come to the exhibition from the opening to the exhibition. Publication activities carried out by *public relations*, for example, by collaborating with various parties in the context of installing exhibition posters and distributing catalogues provided at the opening of the exhibition to visitors who come. The posters and catalogues themselves are part of the advertising efforts made by the gallery management (Table 1, Figure 1.7 and Figure 1.8).

In addition to the *welcome drink* and gallery tour guided by *public relations* at Angky Pu's gallery tour and workshop, visitors who are elementary school students are invited to witness the process of printing stickers from the design results on the computer into the form of printed stickers which are then used by the students to make their work in the form of a piggy bank that is attached to stickers according to their preferences and later can be taken home as *souvenirs* (Table 1, Figure 1.5 and Figure 1.6).

From the three series of activities in the gallery tour, both regular gallery tours and during the opening of the exhibition, it can be observed that there is a deep interaction between tour participants and *public relations* and artists who guide gallery tour activities. This interaction can make a good impression on the participants of the gallery tour which is then conveyed to their closest people at home after the tour is completed. In addition to the interaction between *public relations* and artists and tour participants, the presence of *a welcome drink* when participants first arrive is also expected to give the impression of a good welcome from the gallery management to visitors.

Tabel 1
Dokumentasi Penelitian

No	Picture	Information
1.1		*Welcome drink* provided for visitors during the Angky Pu workshop
1.2		*Welcome drinks* were provided for visitors who came to the opening of the exhibition "Di Timur Matahari" by artist Riono Tanggul

1.3		*Public relations* guided the gallery tour of workshop participants from SD Grow Yogyakarta.
1.4		Artist Rofik accompanies and interacts with gallery visitors during the opening of the exhibition.
1.5		Students of SD Grow Yogyakarta see the process of printing stickers by artist Angky Pu at the workshop.

1.6		Artist Angky Pu interacted directly with the students in his workshop to make each participant's work take home as souvenirs.
1.7		An example of a publication activity carried out by *public relations* by attaching posters in public locations that have collaborated with the gallery
1.8		The catalogue provided by the gallery manager for exhibition visitors is a form of publication and advertising.

Source: Writer, 2023

Satisfaction and Good Impression of Gallery Visitors as a Word of Mouth (WOM) Strategy

Holding the principle of WOM that one customer will talk to another customer about his or her experience when using the product or service purchased or used, then even though it looks simple, this method is the most effective in marketing and even selling a product or service. The management of the Tirtodipuran Link gallery understands this very well, so it tries its best to meet the needs of gallery visitors.

In this study, in addition to direct observations, researchers also conducted interviews to complete the results of the research that had been carried out. The researcher conducted interviews with visitors who came at the opening of the exhibition, an artist who was participating in the exhibition and a gallery guard.

Gallery visitors increase in the month of the opening of exhibitions and workshops.

The researcher asked for data on the number of visitors in the gallery administration to see whether or not there was an increase in the number of visitors in a certain month, especially if there was an opening of an exhibition or workshop in that month. The data received by the researcher is visitor data from August to November because the time of the research conducted by the researcher is from October to December, so the visitor data needed is at least three months of data from the previous month, including those that include the opening of exhibitions and artist workshops.

Table 2
Gallery Visitor Recap August 2023

No	Day, date	Number of visitors
1	Tuesday, 1 August 2023	21
2	Wednesday, 2 August 2023	24
3	Thursday, 3 August 2023	24
4	Friday, 4 August 2023	31
5	Saturday, 5 August 2023	45
6	Sunday, 6 August 2023	37
7	Tuesday, 8 August 2023	33
8	Wednesday, 9 August 2023	32
9	Thursday, 10 August 2023	22
10	Friday, 11 August 2023	32
11	Saturday, 12 August 2023	59
12	Sunday, 13 August 2023	67
13	Tuesday, 15 August 2023	50
14	Wednesday, 16 August 2023	39
15	Thursday, 17 August 2023	63
16	Friday, 18 August 2023	35
17	Saturday, 19 August 2023	54
18	Sunday, 20 August 2023	59
19	Tuesday, 22 August 2023	36
20	Wednesday, 23 August 2023	37
21	Thursday, 24 August 2023	23

22	Friday, 25 August 2023	40
23	Saturday, 26 August 2023	45
24	Sunday, 27 August 2023	41
25	Tuesday, 29 August 2023	30
	Wednesday, 30 August 2023	44
26	Thursday, 31 August 2023	32
Total		**1,055**
Average		**39**

From the August table, it was recorded that the total number of visits in one month was 1,055 people with an average daily visit of 39 people. It is known from public relations that in August there will be the opening of exhibitions and workshops for artists.

Table 3
September 2023 Visitor Recap

No	Day, date	Number of visitors
1	Friday, 1 September 2023	40
2	Saturday, 2 September 2023	75
3	Sunday, 3 September 2023	70
4	Tuesday, 5 September 2023	45
5	Wednesday, 6 September 2023	42
6	Thursday, 7 September 2023	24
7	Friday, 8 September 2023	21
8	Saturday, 9 September 2023	43
9	Sunday, 10 September 2023	48
10	Tuesday, 12 September 2023	29
11	Wednesday, 13 September 2023	18
12	Thursday, 14 September 2023	16
13	Friday, 15 September 2023	17
14	Saturday, 16 September 2023	33
15	Sunday, 17 September 2023	59
16	Tuesday, 19 September 2023	17
17	Wednesday, 20 September 2023	22
18	Thursday, 21 September 2023	15
19	Friday, 22 September 2023	30
20	Saturday, 23 September 2023	30

21	Sunday, 24 September 2023	24
22	Tuesday, 26 September 2023	21
23	Wednesday, 27 September 2023	16
24	Thursday, 28 September 2023	48
25	Friday, 29 September 2023	22
26	Saturday, 30 September 2023	32
Total		**755**
Average		**33**

From the September table, it can be seen that the total number of visits in one month is 755 people with an average of 33 visits per day. In that month the exhibition is still ongoing but there is only a gallery tour if needed by visitors (by request).

Table 4
Gallery Visitor Recap October 2023

No	Day, date	Number of visitors
1	Sunday, 1 October 2023	26
2	Tuesday, 3 October 2023	22
3	Wednesday, 4 October 2023	21
4	Thursday, 5 October 2023	21
5	Friday, 6 October 2023	23
6	Saturday, 7 October 2023	36
7	Sunday, 8 October 2023	33
8	Tuesday, 10 October 2023	26
9	Wednesday, 11 October 2023	15
10	Thursday, 12 October 2023	19
11	Friday, 13 October 2023	16
12	Saturday, 14 October 2023	17
13	Sunday, 15 October 2023	15
Total		**290**
Average		**22**

From the table in October, it can be seen that the number of visitors is only 290 people with an average of 22 visits per day. That month is the end of the ongoing exhibition and half of the month the gallery is closed for work decline and renovation.

Table 4
Gallery Visitor Recap November 2023

No	Day, date	Number of visitors
1	Friday, 10 November 2023	290
2	Saturday, 11 November 2023	39
3	Sunday, 12 November 2023	42
4	Monday, 13 November 2023	24
5	Tuesday, 14 November 2023	17
6	Wednesday, 15 November 2023	22
7	Thursday, 16 November 2023	26
8	Friday, 17 November 2023	24
9	Saturday, 18 November 2023	57
10	Sunday, 19 November 2023	48
11	Tuesday, 21 November 2023	32
12	Wednesday, 22 November 2023	20
13	Thursday, 23 November 2023	28
14	Friday, 24 November 2023	23
15	Saturday, 25 November 2023	34
16	Sunday, 26 November 2023	37
17	Tuesday, 28 November 2023	22
18	Wednesday, 29 November 2023	19
19	Thursday, 30 November 2023	17
Total		**821**
Average		**43**

Finally, in the November table, it can be seen that the total number of visits is 821 people with an average daily visit of 43 people. In that month, there was the opening of the exhibition "Di Timur Matahari" by artist Riono Tanggul.

From the four tables, it can be seen that the highest number of visitors is in August, which is 1,055 people with an average of 30 visitors per day, followed by 821 people in November with an average of 43 visitors per day. This is because in August there is an artist workshop in which there is also a gallery tour and the opening of the exhibition. Similarly, in November there was the opening of the exhibition "In the East of the Sun" by artist Riono Tanggul. In other words, the existence of gallery tours given to visitors, both in artist workshops and exhibition openings, has a clear impact based on gallery visitor data.

The results obtained by the researcher through direct observation, interviews, and administrative data regarding the number of gallery visitors answered some of the

researcher's initial analysis of the effectiveness of gallery tours as a promotional strategy to increase the number of visitors, especially gallery tours contained in artist workshops and exhibition openings. First, the researcher obtained evidence that visitors were very impressed by taking part in the gallery tour both during the Angky Pu artist workshop which was attended by SD Grow students, as well as the impression of visitors who could interact and be accompanied by artists to see the works at the opening of the Riono Tanggul exhibition which was known from the researcher's interview with one of the visitors encountered. However, the role of public relations is also very much needed in this case, both to guide gallery tours if needed, maintain good communication with loyal visitors, and make publications related to ongoing exhibitions in the gallery.

Second, based on the results of interviews with Rofik artists and Vivi visitors, the researcher found that visitor satisfaction if they get gallery tour services both accompanied by direct artists and gallery guards will be a good impression that has the potential to become a Word of Mouth (WoM) promotional strategy. This is due to the theory that truly loyal customers have the potential to become word-of-mouth advertisers, loyal to the company's product portfolio for a long period (Hasan, 2009:81). This means that visitors with a good impression have the potential to revisit the gallery and tell their good impressions to friends or family so that they can become potential visitors to the next gallery. These results are by customer satisfaction indicators according to Hawkins and Lonney (Ngabito, 2020:11) including (1) the suitability of expectations felt by customers, (2) interest in revisiting, and (3) willingness to recommend to friends or family.

Third, based on visitor data obtained by the researcher, in certain months (August and November) that include artist workshops or exhibition openings, including gallery tours in a series of activities, gallery visitors become more crowded than usual days. This is proof of conformity with the definition of word-of-mouth communication according to Ali Hasan customers who have loyalty to the product will be willing to tell good things about the company and its products to others, friends and family which is much more persuasive than advertising (Hasan, 2009:81). So in this case, visitors as loyal customers will tell the good impression of the gallery tour service and other supporting facilities that they receive to the people around them to come and visit the gallery. This is also by the opinion of Gremler and Brown (Hasan, 2009:83) that customer loyalty is a customer who not only re-buys a good or service but also has a commitment and a positive attitude towards the service company, for example by recommending others to buy.

Conclusion

So some of the researchers' findings have proven that a gallery tour can be a form of promotional strategy that prioritizes visitor satisfaction and loyalty by hoping that the good impression obtained through the gallery tour can be a source of positive information in word-of-mouth communication. The results of the study prove that there is a direct comparison between visitor satisfaction through gallery tour services that create a good impression can increase the number of exhibition visitors in the gallery because the

formation of WOM promotion which contains this good impression and synergizes with other forms of promotion, namely advertising and public relations can increase the number of exhibition visitors who come to the gallery. The role of public relations is also needed to create this WOM in addition to carrying out publication activities and helping promotions in the form of advertising from the manager so that visitors come to the exhibitions held, public relations is also tasked with maintaining the loyalty of gallery visitors in the form of good communication when guiding gallery tours.

This study is useful in finding out the importance of gallery tour services that can also be applied to national museums, but with the limitations of researchers in the search for literacy, additional literacy is needed to do so. The researcher proposes to conduct further research by comparing the effectiveness of this direct gallery tour service as a form of promotion with the existence of a virtual tour gallery as a technological development for digital promotion so that it can be known the difference in the impression obtained by gallery and museum visitors between physical and digital gallery tour services. Finally, the researcher can conclude that the success of the promotion strategy carried out is inseparable from several interrelated promotional mixes, both forms of advertising, public relations, word-of-mouth information and other forms of promotion that go hand in hand and do not stand alone to achieve goals, especially marketing in the field of art which requires an emotional bond with the public and art connoisseurs.

Bibliography

Chien, Ting Fang. (2017). Enhancing young adult learning through interpretive skills training: A case study of student tour-guide interns at a university photography center. *International Journal of Education & the Arts*, *18*(34).

Guridno, Ciptoningaji, Azimah, Ariana, & Ningsih, Sari. (2024). ANALISIS HYBRID METODE CNN DAN LSTM DALAM MEDIA BERITA ONLINE INDONESIA. *Jurnal Sistem Informasi Bisnis (JUNSIBI)*, *5*(1), 86–101.

Hanifa, Rafizah Mohd, Isa, Khalid, & Mohamad, Shamsul. (2021). A review on speaker recognition: Technology and challenges. *Computers & Electrical Engineering*, *90*, 107005.

Hartipa Iskandar, Arnia. (2021). *Strategi Marketing Public Relation pada Sanggar Gong Sitimang dalam Mempublikasikan Musik Tradisional Melayu Jambi*. ISI Yogyakarta.

Haumahu, J. P., Permana, S. D. H., & Yaddarabullah, Y. (2021). Fake news classification for Indonesian news using Extreme Gradient Boosting (XGBoost). *IOP Conference Series: Materials Science and Engineering*, *1098*(5), 52081. IOP Publishing.

Isa, Isa, Asrori, Muhammad, & Muharini, Rini. (2022). Peran kepala sekolah dalam implementasi kurikulum merdeka di sekolah dasar. *Jurnal Basicedu*, *6*(6), 9947–9957.

Kaimal, Girija, Drescher, Jon, Fairbank, Holly, Gonzaga, Adele M. L., Junkin, Janelle S., & White, George P. (2016). Learning about leadership from a visit to the art museum. *International Journal of Education & the Arts*, *17*(6).

Kurniati, Tuti, Yusup, Iwan Ridwan, Hermawati, Asni Sri, Kusumahwardani, Devi, Wijayanti, Dewi, & Irhamudzikri, Irhamudzikri. (2021). Respon guru terhadap kendala proses pembelajaran biologi di masa pandemi covid-19. *Jurnal Educatio FKIP UNMA*, *7*(1), 40–46.

Muna, Cut Nailil. (2022). *Manajemen Pameran Oppo Art Jakarta Virtual 2020*. Institut Seni Indonesia Yogyakarta.

Nayoga, Bagas Pradipabista, Adipradana, Ryan, Suryadi, Ryan, & Suhartono, Derwin. (2021). Hoax analyzer for Indonesian news using deep learning models. *Procedia Computer Science*, *179*, 704–712.

NGABITO, Olivia Febrianty. (2020). *Pengaruh e-wom Terhadap Niat Berkunjung Ulang Dengan Kepuasan Pengunjung Sebagai Pemediasi Studi Pada Taman Bunga Gumuk Pasir*. ISI Yogyakarta.

Ramadhan, Iwan, Manisah, Ayu, Angraini, Dini Agra, Maulida, Diah, Sana, Sana, & Hafiza, Nurul. (2022). Proses perubahan pembelajaran siswa dari daring ke luring

pada saat pandemi covid-19 di madrasah tsanawiyah. *EDUKATIF: Jurnal Ilmu Pendidikan*, *4*(2), 1783–1792.

Tandijaya, Jehezkiel Hardwin, Liliana, Liliana, & Sugiarto, Indar. (2021). Klasifikasi dalam Pembuatan Portal Berita Online dengan Menggunakan Metode BERT. *Jurnal Infra*, *9*(2), 320–325.

Teresia Arum, S. K. (2016). *Strategi Promotion Mix Untuk Peningkatan Daya Tarik Seni Pertunjukan Tradisional Di Kabupaten Purworejo: Studi Pada Pertunjukan Tari Dolalak*. Pascasarjana ISI Yogyakarta.