Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024
DOI : https://doi.org/10.33395/sinkron.v9i1.13005
e-ISSN : 2541-2019
p-ISSN : 2541-044X

News Classification Using Bidirectional Long Short-Term Memory And GloVe

Elisabet Margaret Sirait[1], Raynaldo Silalahi[2], Annessa Aprilly Tambunan[3], Junita Amalia[4]*
[1)2)3)4)] Information System, Institut Teknologi Del, Sitoluama, Laguboti, Sumatera Utara, Indonesia
[1)]iss19044@students.del.ac.id, [2)] iss19001@students.del.ac.id, [3)] iss19017@students.del.ac.id,
[4)]junita.amalia@del.ac.id

Submitted : Sep 12, 2023 | **Accepted** : Oct 13, 2023 | **Published** : Jan 1, 2024

Abstract: The dissemination of information and news via online media encompasses not only established news platforms but also contributions from internet users, lacking oversight. News constitutes fact-grounded insights into ongoing occurrences. This research employed Bidirectional Long- and Short-Term Memory with Hyperparameter tuning on GloVe for news classification. This research aims to optimize news categorization through hyperparameter tuning on GloVe. GloVe facilitated the transformation of words into vector matrices, exploring its efficacy in news classification with hyperparameter tuning and Bi-LSTM for text analysis. Experiments encompassed untuned and hyperparameter-tuned approaches, employing GloVe's hyperparameters using Gridsearch and manual methods. The purpose is to achieve the best accuracy performance in news classification through hyperparameter tuning of Global Vectors. GloVe's hyperparameter tuning reveals the potential for enhancing word vector representations. Surprisingly, non-hyperparameter tuned news classification yielded superior evaluation results compared to the hyperparameter approach. The untuned experiment achieved an accuracy of 0.98, while the gridsearch method yielded 0.85 accuracy, and hyperparameter tuning generated a 0.88 precision in the -11 model. These findings underscore the nuanced interplay of hyperparameters in optimizing text classification models like GloVe.

Keywords: Classification text, Word embedding, Hyperparameter Tuning, GloVe, Bi-LSTM

INTRODUCTION

The dissemination of information and news is out of the control of all online media users. Information or news can be disseminated through online media not only by well-known news sites but also by any Internet user without undergoing any inspection. However, some of the information distributed individually or in groups cannot be held responsible for being true or being indicated as fake news because it reaches thousands of users. Twitter has become one of the most actively used social media platforms, with more than 372.9 million users by April 2023. On Twitter, there are many tweets that contain fake news that has a negative impact on both individuals and groups. Fake news is information or news that contains things that are not known to be true or uncertain (Juditha, n.d.) (Tambunan, Nataliani, & Lestari, 2021).

In the previous study, the classification of the research classified the news in Indonesia into three classes using LSTM and GRU with fasttext and GloVe. In this study, combining Bi-LSTM with GloVe resulted in an accuracy of 86.767%. The accuracy of the combination of GRU and GloVe yielded 88.967%. In the study (Nayoga, Adipradana, Suryadi, & Suhartono, 2021) it used seven classification models: LSTM, Bi-LSTM, GRU, Bi-GRU, Nave Bayes, SVM, and ID-CNN. In this study, the accuracy

*name of corresponding author

value on Bi-LSTM achieved a higher precision compared to Bi-GRU. The accuracy value of Bi-LSTM is 96.2%, while Bi-GRU is 95.2%. (Adipradana, Nayoga, Suryadi, & Suhartono, 2021) (Nayoga, Adipradana, Suryadi, & Suhartono, 2021)

In the study using LSTM and GloVe methods as word representations to produce Word embeddings In this study, the GloVe model was divided using Xmax, Alpha, and iteration parameters. The parameter values used in this study refer to papers that use the English Wikipedia dataset with values of Xmax 100 and Alpha 34. The iteration value used in this study is 50; iterative values on GloVe can be used varyingly, and the larger iterational values will result in better performance (Pardede & Ibrahim, 2020).

In this study, researchers will find how does hyperparameter tuning affect GloVe in classifying news to find best accuracy performance in news classification through hyperparameter tuning of GloVe used the Bi-LSTM and GloVe algorithms. GloVe shows how to involve global statistical information contained in documents. The semantic meaning of a word is not only influenced by the surrounding words but also by the global statistical information of the document. GloVe uses the ratio of co-occurrence probability between words. To optimize the value of GloVe, researchers use several GloVe parameters, namely alpha, Xmax, and iteration, in classifying fake news. Bidirectional LSTM (BiLSTM) is a neural network consisting of two interconnected LSTM layers. BiLSTM processes information from both directions in the data sequence by utilizing two separate layers, namely the forward LSTM and the backward LSTM. The forward LSTM aims to understand the preceding context in the data sequence, while the backward LSTM aims to comprehend the succeeding context. The choice of the Bi-LSTM method is due to its ability to understand the context of text well. Besides, Bi-LSTM selection is also due to the ability of the Bi-LSTM algorithm to remember values at variable time intervals.

LITERATURE REVIEW

There have been many studies that discuss forecasting using deep learning. Especially deep learning that is supervised learning. The literature review in this study reviews research using the ARIMA and LSTM models. the objectives, conclusions and suggestions of these studies are presented in Table 1, in addition, a comparison between previous research and the research to be carried out is also presented.

Table 1. Comparison of Previous Research

Reference	Title	Conclusion	Suggestion	Comparison
(Adipradana et al., 2021)	*Hoax analyzer for Indonesian news using RNNs with fasttext and GloVe embeddings*	GRU has better performance compared to LSTM when the dataset has less frequent occurrence and is widely spread. Both bidirectional models of LSTM and GRU yield better metrics score than their unidirectional counterparts; fastText is better than GloVe in performance as fastText can handle out of vocabulary (OOV) words and rare words better than GloVe; fastText and bidirectional GRU combined yielded the	To obtain better result, it is better to use dataset that contain Indonesia language news without being translated from English.	This research uses LSTM and GRU with fasttext and GloVe to classify news while this research to be carried out uses BILSTM with hyperparameter tuning in GloVe to classify news.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v9i1.13005 p-ISSN : 2541-044X

		highest result in this experiment, mainly because the dataset is widely spread and has shorter text length.		
(Pardede & Ibrahim, 2020)	*Implementasi Long Short-Term Memory* Untuk Identifikasi Berita Hoax Berbahasa Inggris Pada Media Sosial	In this research, the LSTM method has been implemented to identify hoax news in the English language on social media. The proposed method can identify hoax news with an average precision, recall, accuracy, and F-measure of 0.94, 0.96, 0.94, and 0.95, respectively. The results of the study demonstrate that LSTM outperforms the Support Vector Classifier, Logistic Regression, and Multinomial Naive Bayes methods in terms of performance.	In future research, it is hoped that performance can be improved by building a more diverse vocabulary through the selection of a more varied corpus. This would result in a more varied vocabulary.	In the research, GloVe word embeddings and the LSTM method were utilized through various stages, including preprocessing, training, testing, and evaluation. GloVe was configured with parameters such as Xmax, alpha, and iteration. The default values for each parameter were set as follows: Xmax = 100, Alpha = ¾, and the number of iterations could vary.
(Nayoga et al., 2021)	*Hoax Analyzer for Indonesian News Using Deep Learning Models*	One Dimensional Convolutional Neural Network with Batch Normalization is a viable option as a feature extractor for Natural Language Processing. Dropout technique is best used for One Directional Recurrent Neural Network and Bidirectional Neural. Network as it may improve performance for both models and is not suitable for Gated Recurrent Unit as it may cause NaN loss. One	The outcome of this research can be used for real-world applications such as Hoax Analyzer. This concludes our research regarding the educational and explorative purposes	In this study, seven classification models were employed, namely LSTM, Bi-LSTM, GRU, Bi-GRU, Naïve Bayes, SVM, and ID-CNN.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024
DOI : https://doi.org/10.33395/sinkron.v9i1.13005
e-ISSN : 2541-2019
p-ISSN : 2541-044X

		Dimensional Gated Recurrent Unit has higher performance and efficiency ratio compared to One Dimensional Long Short-Term Memory. Deep Neural Networks is a better option for Natural Language Processing compared to Conventional Classifiers		
(Susanty & Sukardi, 2021)	Perbandingan Pre-trained *Word embedding* dan Embedding *Layer* untuk Named-Entity Recognition Bahasa Indonesia	In this research, a comparison of various word embedding techniques was conducted for Named Entity Recognition (NER) tasks in the Indonesian language using a BiLSTM model architecture. The unsupervised approach using pre-trained embeddings demonstrated better performance compared to the supervised approach using a trainable embedding layer. This unsupervised approach proved to be more resistant to overfitting. Hyperparameter optimization improved the performance of the trainable embedding layer, but the performance of pre-trained embeddings remained superior. Among the various pre-trained embeddings compared in this study, the embedding technique that achieved the best average micro F1 score was GloVe.	For future research, it is advisable to use a larger dataset to reduce model overfitting to the training data.	In this research, two types of embedding methods were tested: trainable embedding layer and pre-trained word embedding. Both of these methods serve the same purpose, which is to convert textual data (a collection of words or tokens) into numerical representations that capture semantic relationships between words. In this study, three pre-trained word-level embeddings were compared: Word2Vec, GloVe, and FastText.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024
DOI : https://doi.org/10.33395/sinkron.v9i1.13005
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Data Preprocessing

Data in general mostly consists of raw data that has a lot of noise, is large, and comes from various sources. Data preprocessing is an important and necessary step to obtain better data quality (Pradnyana & Agustini, n.d.).

Data Integration

Data integration is the process of combining two or more data sets into one unified set. This process integrates data in one place that comes from a variety of different sources. Data Integration in this study is needed to combine research data from two different datasets, namely Fake.csv and true.csv. csv . The data integration stage is carried out with the aim of combining two datasets into one with the same attributes. Two different datasets are put together and have Fake and True attributes. Before going through the integration process, each dataset did not yet have a label. Then labeling is done on the dataset, namely 0 = fake and 1 = True. The example of data integration process is shown in Table (Han, Kamber, & Pei, n.d.).

Table 2 Data Integration

Title	Text	Subject	Date	Label
Alabama to certify Democrat Jones winner of Senate election	Jones will be the first Democrat sent to the Senate from Republican stronghold Alabama in a quarter century.	politicsNews	22/12/2017	1
Bad News For Trump â€” Mitch McConnell Says No To Repealing Obamacare In 2018	Republicans have had seven years to come up with a viable replacement for Obamacare but they failed miserably.	News	21/12/2017	0

Data Cleaning

Data cleaning is one of the stages of data mining. Data cleaning is done because, in general, data is not found in a complete state but in a state of error (noisy data) and is also inconsistent. Data cleaning needs to be done on the dataset that has been integrated because the dataset to be used still has duplication and unused attributes so that there are no errors in producing text classification. In data cleaning, which involves dropping columns, removing null data, and removing duplicates, the application of data cleaning results in data changes in the first and second datasets in terms of quantity. The example of data cleaning process is shown in 3 (Han et al., n.d.).

Table 3 Data Cleaning

Before data cleaning		After data cleaning	
Attribute	Number of Datasets	Attribute	Number of Datasets
Title, text, subject, date, and label	44898 rows × 5 columns	Text, dan Label	3987 rows × 2 columns

Text Preprocessing

Text Preprocessing is an important part of Natural Language Processing (NLP) systems, as the characters, words, and sentences identified at this stage are the basic units that are passed on to all subsequent processing stages, from analysis and tagging components, such as morphological analysis and part-of-speech taggers, to applications such as information retrieval systems and machine translation (Han et al., n.d.).

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Case Folding

This operation is used to convert all letters to lowercase. It aims to prevent the model from recognizing that a word is different because it uses lowercase and uppercase letters in the text. Case Folding needs to be done to facilitate the data classification process, such as reducing variations in the same word. After case folding is implemented, all letters in the text will be converted to lowercase. The example of cae folding is shown in 4 (Işik & Dağ, 2020).

Table 4 Case Folding

	News text
Input	A lawyer for the Trump transition team is claiming that the emails had been illegally turned over by the General Services Administration because the account owners never received notification of the request and he s claiming that they were privileged communications.
Output	a lawyer for the trump transition team is claiming that the emails had been illegally turned over by the general services administration because the account owners never received notification of the request and he s claiming that they were privileged communications

Remove the URL and Remove Username

The process of removing URLs or hyperlinks from text in a dataset URL stands for Uniform Resource Locator, which is a web address or website. URL removal needs to be done to improve data quality and facilitate text classification because URLs in the dataset are considered noise or irrelevant information in data classification. The process of removing usernames from the text in the dataset Usernames can appear in various forms in text, such as @username, on social media platforms like Twitter or Instagram. Removing usernames needs to be done to improve data quality and facilitate text classification because usernames are considered noise or irrelevant information in data analysis (Aggarwal, 2015). After remove URL is implemented, all URL links contained in the text will be removed. And in remove username, after being implemented, all usernames contained in the text will be deleted. The example of remove URL and remove username is shown in Table (Aggarwal, 2015).

Table 5 Remove URL and Remove Username

	News text
Input	Especially if you continue to lie. Months after decision not to charge Clarke, email search warrant filed https://t.co/zcbyc4Wp5b KeithLeBlanc (@KeithLeBlanc63)
Output	especially if you continue to lie. months after decision not to charge clarke, email search warrant filed keithleblanc

Remove Punctuation and Remove Number

The process of removing punctuation from the text in the dataset Punctuation can appear in various forms in the text, such as periods, commas, question marks, exclamation marks, or dashes. Removal of punctuation is necessary because punctuation in the dataset can be considered noise or irrelevant information (Aggarwal, 2015). The process of removing numbers or figures from text in a dataset Numbers can appear in various forms in text, such as whole numbers, fractions, or percentages. The removal of numbers is necessary to improve the quality of the dataset and facilitate text classification. After removing punctuation is implemented, all punctuation marks contained in the text will be removed. And in remove numbers, after being implemented, all numbers contained in the text will be removed. The example of Remove punctuation and remove number is shown in Table (Aggarwal, 2015).

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 6 Remove Punctuation and Remove Number

	News text
Input	A victory by Simonds would shift Republicansâ€™ slim control of the 100-member House of Delegates to an even 50-50 split with the Democrats, forcing the two parties into a rare power-sharing arrangement.
Output	a victory by simonds would shift republicans slim control of the member house of deegates to an even split with the democrats forcing the two parties into a rare power sharing arrangement

Remove the stopword.

Stopwords are words that appear frequently in a language and have little meaning in text mining. Remove stopwords because stopwords in the dataset can cause differences in the classification of text that should be the same. After remove stopword is implemented, all words that have no meaning or are irrelevant in the sentence will be removed. The example of remove stopword is shown in Table (Işik & Dağ, 2020).

Table 7 Remove The Stopword

	News text
Input	In Michigan, the upcoming election in 2018 presents voters with an interesting option: They could, for the first time, have an all-female ticket for every major office being contested.
Output	michigan coming election presents voters interesting option they could first time have all female ticket every major office contested

Tokenization

Tokenization is the process of breaking down text into smaller units, usually in the form of words or pieces of meaning referred to as tokens. By Tokenizing the dataset, we can speed up the data classification process because there is no need to break the text manually. After tokenization is implemented, all words in the data will be converted into tokens. The example of tokenization is shown in Table (Işik & Dağ, 2020).

Table 8 Tokeization

	News text
Input	make changes left
Output	‘make’ ‘changes’ ‘left’

Bidirectional Long Short Term Memory (Bi-LSTM)

Bidirectional LSTM (BiLSTM) is a neural network architecture composed of two interconnected LSTM layers. BiLSTM processes information bidirectionally within a sequence of data by employing two separate layers: the forward LSTM and the backward LSTM. The forward LSTM aims to comprehend the preceding context within the data sequence, while the backward LSTM focuses on understanding the subsequent context. Each LSTM layer in BiLSTM utilizes cell states and gates, similar to those found in a regular LSTM.

Global Vector (GloVe)

GloVe (Global Vectors for Word Representation) is an algorithm for constructing word vector representations, proposed by Jeffrey Pennington, Richard Socher, and Christopher Manning in 2014. GloVe is a model for word representation that directly captures global corpus statistics. GloVe's vector representations are based on a co-occurrence matrix that counts the frequency of co-appearances between pairs of words in a text corpus. The GloVe algorithm generates word vector representations by minimizing the error between the dot product of word vector representations and the logarithm of the co-occurrence frequency in the co-occurrence matrix.

*name of corresponding author

METHOD

In this sub-chapter, it describes the methods to be used in the research, namely the stages of work to be carried out to meet the purposes of the research. An explanation of the stages of the research conducted is presented in the following information and shown in Figure 1:

1. The data used in this research is sourced from Kaggle, consisting of two datasets named True.csv and Fake.csv. The data used in this research is in English and comprises 21,417 rows for True.csv and 23,481 rows for Fake.csv. The attributes available in the datasets are title, text, subject, date, and label. However, due to memory limitations, only 2000 rows from each dataset were used for analysis.
2. In the next stage, data preprocessing is carried out, namely, data integration and data cleaning.
3. At the next stage, text preprocessing is carried out, namely: case folding, removing URL, removing username, removing number, removing punctuation, removing the special word, removing stopwords, and tokenization.
4. The preprocessing results will then be input to the word embedding stage with GloVe. The parameters used to get the most optimal value from GloVe are, among others, Xmax, alpha, and iteration. The input obtained from the preprocessing stage is used to calculate the co-occurrence matrix value. Hyperparameter tuning is done to get the optimal value of the parameters used in the co-occurrence matrix.
5. Next, the classification stage is carried out using the Bi-LSTM method, which studies the semantics of words (words before and words after) and relationships between words that are influential in determining the final result (fake or real) of the sentence under consideration. The first step is to take data from each word vector. Then classification is carried out using the existing layers in Bi-LSTM. Then compile the model; the results of the compiled model will be fitted to the model; and then calculate the accuracy and cross values. The news will be classified into fake and real classes.
6. The final stage is the evaluation of the news classification model. Evaluation results are obtained by calculating the accuracy, precision, recall, and f1-score values. The calculation results will be in the range of 0 to 1, where if the results of these values are close to 1, then the classification is said to be good, and if otherwise, it is a bad classification.

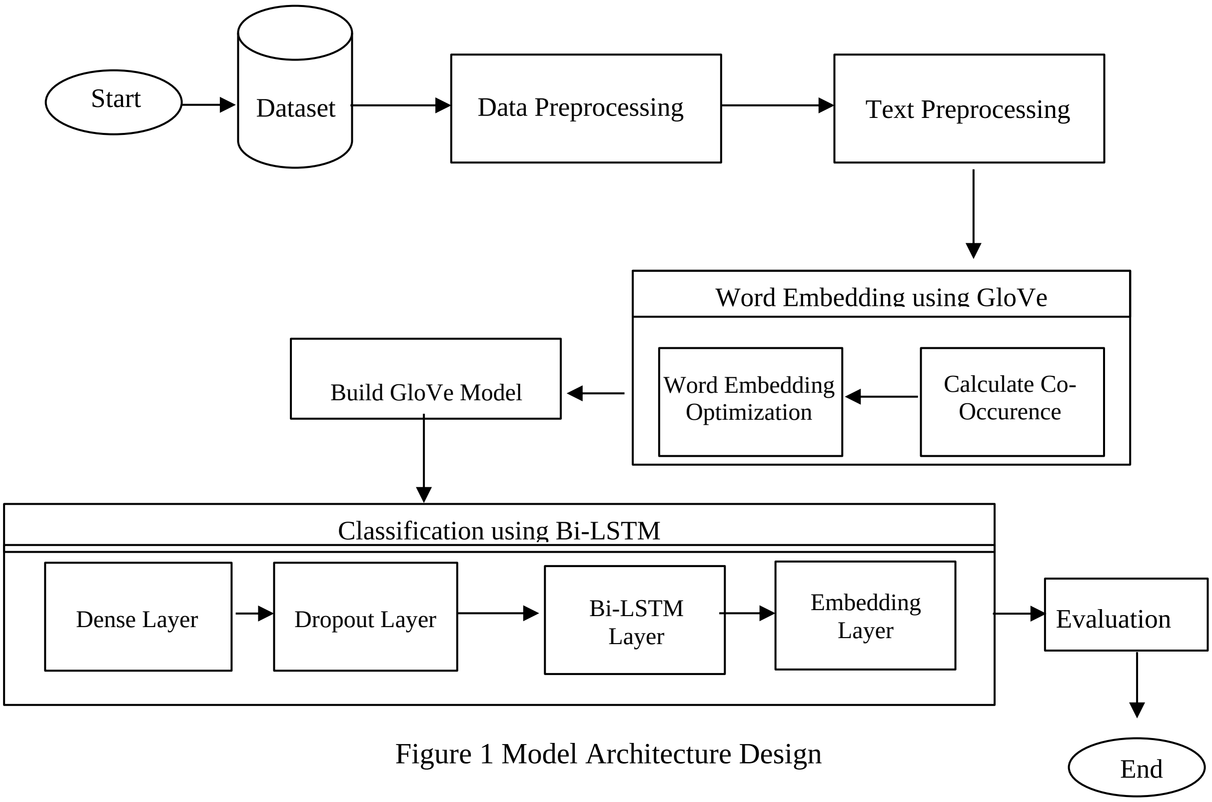

Figure 1 Model Architecture Design

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024
DOI : https://doi.org/10.33395/sinkron.v9i1.13005
e-ISSN : 2541-2019
p-ISSN : 2541-044X

RESULT

In this chapter, the implementation results obtained from each stage of the experiment that has been carried out on the research are outlined. The experiments were conducted without using hyperparameters, using gridsearch, or using manual.

Experimental Results Without Hyperparameter Tuning

Based on the implementation that has been carried out, the following are the results of experiments carried out without tuning hyperparameters, shown Table .

Table 9 Without Using Parameters

	Precision	Recall	F1-Score
0	0.96	0.99	0.98
1	0.99	0.97	0.98
accuracy			0.98
macro avg	0.98	0.98	0.98
weighted avg	0.98	0.98	0.98

The accuracy result of the experiment without tuning hyperparameter tuning on GloVe yielded an accuracy score of 0.98.

Experimental Results With Hyperparameter Tuning

Based on the implementation, use gridsearch hyperparameter tuning with values Xmax = 100, 150, and 200; Alpha = 0.75, 1, and 1.25; and iteration = 100, 150, and 200. The best parameters of the experiment using gridsearch are Alpha = 1.0, Xmax = 150, and iteration = 200. The results of the news classification using the gridsearch is shown in Table 10.

Table 10 Using Gridsearch

	Precision	Recall	F1-Score
0	0.95	0.73	0.83
1	0.79	0.96	0.87
accuracy			0.85
macro avg	0.87	0.85	0.85
weighted avg	0.87	0.85	0.85

3.3 Experimental Results With Using Manual Hyperparameter Tuning

The accuracy result of the experiment with hyperparameter gridsearch on GloVe yielded an accuracy score of 0.85.

The implementation use manual hyperparameter tuning with values Xmax = 100, 150, and 200; Alpha = 0.75, 1, and 1.25; and iteration = 100, 150, and 200. The best parameters of the experiment using the hyperparameter manual are Alpha = 1, Xmax = 100, and Iteration = 150. The results is shown in Table 1.

Table 1 Using Manual Hyperparameter Tuning

	Precision	Recall	F1-Score
0	0.94	0.80	0.87
1	0.83	0.93	0.88
accuracy			0.88
macro avg	0.89	0.88	0.88
weighted avg	0.89	0.88	0.88

The accuracy result of the experiment with the hyperparameter manual on GloVe yielded an accuracy score of 0.88.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024
DOI : https://doi.org/10.33395/sinkron.v9i1.13005
e-ISSN : 2541-2019
p-ISSN : 2541-044X

DISCUSSIONS

According to the results of experiments without hyperparameter tuning, the performance of the model is better than with hyperparameter tuning. Hyperparameter tuning is done using two methods: gridsearch and manual. In both methods, there are three parameters that were experimented with in this study: alpha, Xmax, and iteration. The values of each parameter used are 0.75; 1; and 1.25 for the alpha value, 100; 150; and 200 for the Xmax value, and 100; 150; and 200 for the iteration value. The comparison of accuracy values with GloVe without a tuning hyperparameter and using a tuning hyperparameter is shown in Figure 2.

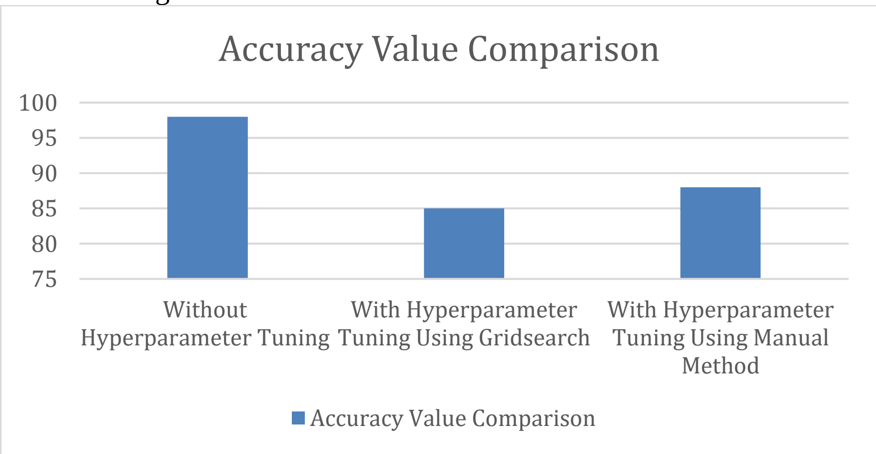

Figure 2 Comparison of accuracy values

In this study, by conducting experiments using gridsearch and using hyperparameters, the accuracy result was not optimal compared to the result without hyperparameter tuning. The accuracy result of the experiment without performing the GloVe hyperparameter tuning yielded an accurate value of 0.98. Hyperparameter experiments using gridsearch yield accuracy of 0.85 with best alpha parameters of 1.25, Xmax = 100, and iterations of 150: Alpha = 1.0, Xmax = 150, and iteration = 200. Meanwhile, the hyperparameter experiment using manual methods produced an accuracy of 0.88 on the experimental model 11 with a parameter value of alpha = 1, Xmax = 100, and Iteration = 150.

The comparison of recall values with GloVe without a tuning hyperparameter and using a tuning hyperparameter is shown in Figure 3.

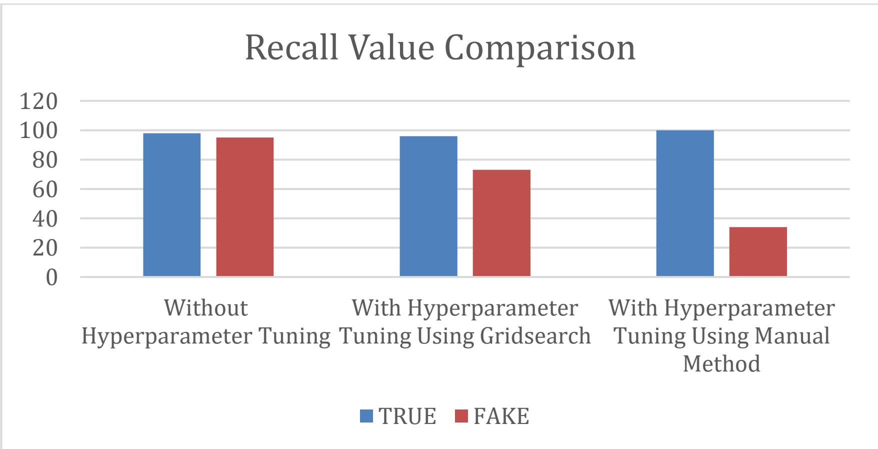

Figure 3 Comparison of recall values

In this study, by conducting experiments using gridsearch and using hyperparameters, the recall result was not optimal compared to the result without hyperparameter tuning. The recall results from the experiment without performing the GloVe hyperparameter tuning resulted in a recall value of 0.98 on the true label and 0.95 on the fake label. The hyperparameter experiment using gridsearch yielded a

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

recall of 0.96 on the true label and 0.73 on the fake label. Meanwhile, the hyperparameter experiment using the manual method yielded a recall of 1 on the true label and 0.34 on the fake label in model 5, with the alpha parameter value of 0.75, Xmax = 150, and Iteration = 150.

The comparison of precision values with GloVe without a tuning hyperparameter and using a tuning hyperparameter is shown in Figure 4.

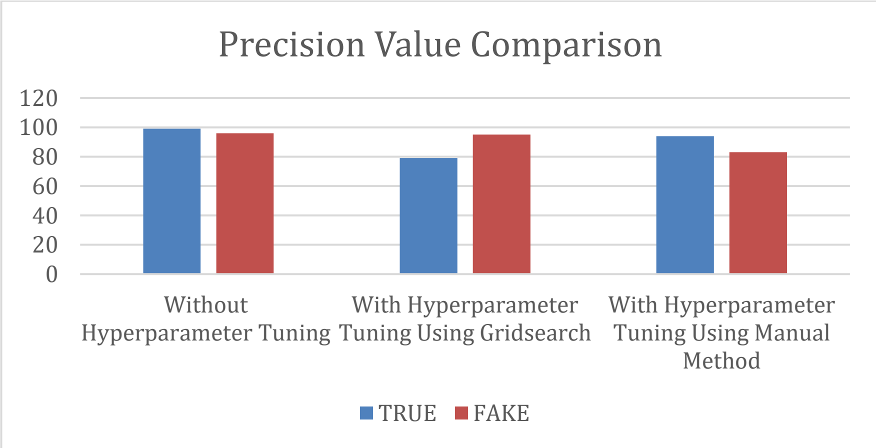

Figure 4 Comparison of Precision values

In this study, by conducting experiments using gridsearch and using hyperparameters, the precision result is not optimal compared to the result without hyperparameter tuning. The precision result of the experiment without performing the GloVe hyperparameter tuning yielded a precision value of 0.99 on the true label and 0.96 on the fake label. Hyperparameter experiments using gridsearch yielded a precision of 0.79 on the true label and 0.95 on the fake label. Meanwhile, the hyperparameter experiment using manual methods yielded precision results of 0.94 on the true label and 0.83 on model 11, with alpha parameter values of 1; Xmax = 100; Iteration = 150.

The comparison of F1-Score values with GloVe without a tuning hyperparameter and using a tuning hyperparameter is shown in Figure 5.

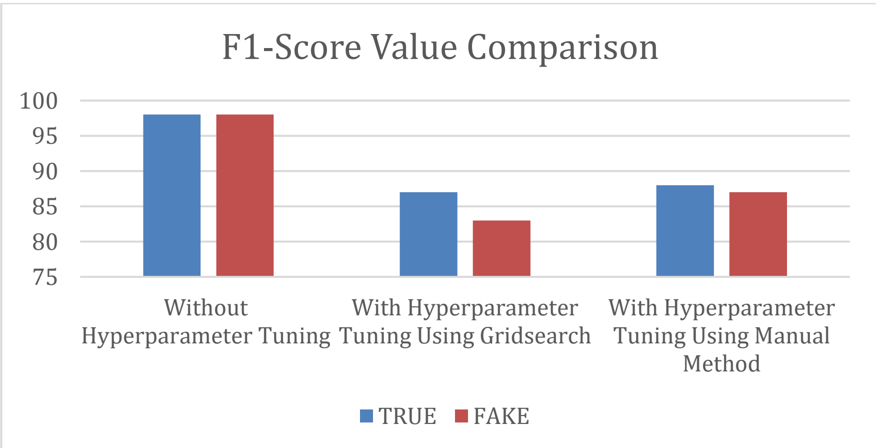

Figure 5 Comparison of F-1 Score values

In this study, by conducting experiments using gridsearch and using hyperparameters, the F1-Score result is not optimal compared to the result without hyperparameter tuning. F1-Score results from experiments without performing the GloVe hyperparameter tuning resulted in F1-Score values of 0.98 on the true label and 0.98 on the fake label. Hyperparameter experiments using gridsearch yielded an F1-Score of 0.87 on the true label and 0.83 on the fake label. Meanwhile, the hyperparameter experiment

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024
DOI : https://doi.org/10.33395/sinkron.v9i1.13005
e-ISSN : 2541-2019
p-ISSN : 2541-044X

using the manual method of F1 yielded a score of 0.88 on the true label and 0.87 on the fake label in model 11 with the alpha parameter value of 1; Xmax = 100; Iteration = 150.

Based on the experiments that have been carried out, experiments with hyperparameters yield lower results than experiments without the use of hyperparameters. There are several factors that affect accuracy, recall, precision, and lower F1-Score values. Here are some factors that influence

1. The size of the data set used
   In the experiment, using a data set of 4,000 data points divided into 2,000 true data points and 2,000 fake data points,The use of this data set is due to the software used in the experiment having limited memory, i.e., 16.2 GB. It's written in sub-chapter 5.2 of the implementation limitations.
2. Unoptimal parameter values
   The values used in this experiment are based on the default values, and range selection still uses a small multiple. At the selection of this range, according to the researchers, the hyperparameter value is inaccurate, so the value of the evaluation matrix is not optimal. It can be seen from experiments performed by researchers using hyperparameters that accuracy, recall, precision, and f1-score values are lower than in experiments with GloVe without hyperparameter.

CONCLUSION

To compare the performance of each experiments can be seen in result chapter. The data used in this research is sourced from Kaggle, consisting of two datasets named True.csv and Fake.csv. The data used in this research is in English and comprises 21,417 rows for True.csv and 23,481 rows for Fake.csv. The attributes available in the datasets are title, text, subject, date, and label. The allocation of training, testing and validation data that produce the best accuracy values is with a dataset allocation of 70% for training data, 20% for testing data and 10% for validation data.

In the result of this study, experiments without using the hyperparameter tuning for accuracy 0.98, recall 0.98 for true data and 0.95 for false data, precision ; 0.99 for true and 0.96 for fake data, and F1-Score 0.98 to true and 0.98 for falsified data. Experiments using the automated method of gridsearch obtain the best parameter Alpha = 1.0; Xmax = 150; iteration = 200. Receives an accuracy value of 0.85, recall of 0.73 for true data and 0.96 for false data, precision of 0.79 for true and 0.95 for fake data, and F1-Score of 0.87 for True and 0.83 for False data. Experiment using manual methods with hyperparameters alpha = 1 ; Xmax = 100 ; Iteration = 150. Receives an accuracy value of 0.88, recall of 0.80 for true data and 0.93 for fake data, precision of 0.94 for true Data and 0.83 for false data, and F1-Score of 0.87 for true and 0.88 for False data. This study has limitation, due to memory limitations, only 2000 rows from each dataset were used for analysis.

REFERENCES

Adipradana, R., Nayoga, B. P., Suryadi, R., & Suhartono, D. (2021). Hoax analyzer for Indonesian news using RNNs with fasttext and GloVe embeddings. *Bulletin of Electrical Engineering and Informatics*, *10*(4), 2130–2136. https://doi.org/10.11591/eei.v10i4.2956

Aggarwal, C. C. (2015). *Data Mining: The Textbook*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-14142-8

Augustyniak, Ł., Kajdanowicz, T., & Kazienko, P. (2019). Aspect Detection using Word and Char Embeddings with (Bi)LSTM and CRF. *2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE),* 43–50. https://doi.org/10.1109/AIKE.2019.00016

Han, J., Kamber, M., & Pei, J. (n.d.). *Data Mining Concept and Techniques third edition* (3rd ed.). Morgan Kaufmann.

Işik, M., & Dağ, H. (2020). The impact of text preprocessing on the prediction of review ratings. *TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES*, *28*(3), 1405–1421. https://doi.org/10.3906/elk-1907-46

Juditha, C. (n.d.). *Interaksi Komunikasi Hoax di Media Sosial serta Antisipasinya Hoax Communication Interactivity in Social Media and Anticipation*. *3*(1).

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v9i1.13005 p-ISSN : 2541-044X

Li, H., Caragea, D., Li, X., & Caragea, C. (2018). Comparison of Word Embeddings and Sentence Encodings as Generalized Representations for Crisis Tweet Classification Tasks. *New Zealand*.

Nayoga, B. P., Adipradana, R., Suryadi, R., & Suhartono, D. (2021). Hoax Analyzer for Indonesian News Using Deep Learning Models. *Procedia Computer Science*, *179*, 704–712. https://doi.org/10.1016/j.procs.2021.01.059

Pardede, J., & Ibrahim, R. G. (2020). Implementasi Long Short-Term Memory untuk Identifikasi Berita Hoax Berbahasa Inggris pada Media Sosial. *Journal of Computer Science and Informatics Engineering (J-Cosine)*, *4*(2), 179–187. https://doi.org/10.29303/jcosine.v4i2.361

Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543. Doha, Qatar: Association for Computational Linguistics. https://doi.org/10.3115/v1/D14-1162

Pradnyana, G. A., & Agustini, Dr. K. (n.d.). *Konsep Dasar Data Mining*.

Reimers, N., & Gurevych, I. (2017, August 16). *Optimal Hyperparameters for Deep LSTM-Networks for Sequence Labeling Tasks*. arXiv. Retrieved from http://arxiv.org/abs/1707.06799

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (n.d.). *Dropout: A Simple Way to Prevent Neural Networks from Overfitting*.

Susanty, M., & Sukardi, S. (2021). Perbandingan Pre-trained Word Embedding dan Embedding Layer untuk Named-Entity Recognition Bahasa Indonesia. *PETIR*, *14*(2), 247–257. https://doi.org/10.33322/petir.v14i2.1164

Tambunan, S. M., Nataliani, Y., & Lestari, E. S. (2021). Perbandingan Klasifikasi dengan Pendekatan Pembelajaran Mesin untuk Mengidentifikasi Tweet Hoaks di Media Sosial Twitter. *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, *7*(2), 112. https://doi.org/10.26418/jp.v7i2.47232

*name of corresponding author