Journal of Theoretical and Applied Information Technology
31st December 2024. Vol.102. No. 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

COMPARISON OF THE ENSEMBLE XGBOOST AND TRANSFORMER MODELS WITH MACHINE LEARNING FOR CLASSIFICATION OF INDONESIAN MUSIC MOODS OF THE 70'S AND 80'S ERA

SARTIKA LINA MULANI SITIO [1], THOYYIBAH.T [2], MARYANI [3], MAULANA ARDHIANSYAH [4], NENY ROSMAWARNI [5], YAN MITHA DJAKSANA [6], NUNIK DESTRIA ARIANTI [7]

[1,4,6]Faculty of Computer Science, Jl. Raya Puspitek No. 46, Serpong, Banten, Indonesia, 15316
[2]Information Systems Management Department, BINUS Graduate Program-Master of Information System Management Bina Nusantara University, Jakarta, Indonesia 11480
[3]Information Systems Department School of Information System Bina Nusantara University Jakarta, Indonesia 11480
[5]Information Technology, University Pembangunan Nasional Veteran, Jakarta, Indonesia
[7]Information System, University Nusa Putra, Sukabumi, Indonesia
E-mail: dosen00847@unpam.ac.id[1], thoyyibah.t@binus.ac.id[2], yanie@binus.edu[3], dosen00374@unpam.ac.id[4], nenyrosmawarni@upnvj.ac.id[5], dosen01994@unpam.ac.id[6], nunik@nusaputra.ac.id[7]

ABSTRACT

The aim of this research is to compare the Xgboost and transmormer machine learning models. In the field of psychology, research on musical mood aims to find out why humans have an emotional response to music. In the field of Music Information Retrieval (MIR), research on the emotion and mood of music aims to create music metadata to make it easier to manage and retrieve music as an entity. The method used in this research is the CRISP-DM method with the stages of business understanding, data understanding, data preparation, modeling, evaluation and deployment. This research has only reached the evaluation stage. This research used 1160 data using lyrics and audio data and used 31923 different words. This research uses a dataset of Indonesian songs from the 70's and 80's. The labels used for mood classification are happy, sad and neutral. The lyric representation used is word embedding and audio consisting of Chromagram, Mel-Spectrogram. This research uses the Xgboost model and transformer machine learning. With several analyzes used, the transformer model accuracy achieved is 98%.
Keywords : *Indonesian Music, Xgboost, Transformer, Crisp-DM*

1. INTRODUCTION

Music is an object of research in the field of machine learning, especially on the topic of audio signal processing, which has attracted many researchers with broad potential applications, including in the fields of education, advertising and entertainment. On the other hand, there is a research problem, if given a number of music samples, how to select the dominant musical features extracted from the music as a representation of musical "content" that is not related to a particular musical genre to be used in developing a robust musical mood classification model. In the field of psychology, research on musical mood aims to find out why humans have an emotional response to music. In the field of Music Information Retrieval (MIR), research on musical mood aims to create music metadata to make it easier to manage and retrieve music as an entity. So it is necessary to study further about what musical features are most optimal for representing the musical mood in several musical genres. After that, further analysis is which machine learning model has the most optimal performance for classifying musical mood. The final stage is measuring the performance of the classification model using metrics. Several studies in the field of music use the Xgboost and transformer machine learning method consist of credit scoring using xgboost [1]. Second research for analyzing sentiment on social media [2]. The third study analyzed the classification of syn attacks

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[3]. fourth research, Xgboost analysis of Indonesian smart cards for college students [4]. Fifth research on the use of Xgboost for traffic predictions in Banjarmasin [5]. Sixth research on the use of Xgboost social media analysis internet

services [6]. Seventh research on xgboost on cyberbullying on social media [7]. Eighth research about xgboost on Playstore application ratings [8]. Ninth research about xgboost on plagiarism detection [9]. Tenth research on xgboost on customer classification [10]. Eleventh research on transformer architecture [11]. Twelfth research on transformers in political models [12]. Thirteenth research on transformers in NLP (Natural Language Processing) [13]. Fourteenth research on transformers, Bert, GPT [14]. Fifteenth research on transformers in attention [15]. Sixteenth research on transformers on BERT [16].

Even though there have been many publications of previous research results that propose various methods for modeling and analyzing music, research in the field of music still faces a number of challenges. The main challenges in the field of music representation. Music has various attributes (features), so choosing a musical representation is very important. Current musical representations cannot be used to represent musical "content" (as is the case with semantics in linguistic sentences). Currently there is no representation of music across music genres so that we can measure the similarity of musical "content", for example: mood.The two main challenges in the field of music research are:

1) Availability of publicly labeled data for research purposes, although currently many researchers have published the results of their research in the field of music modeling, but not all of the datasets used are available to other music researchers.

2) Music representation, currently there is no agreement among music researchers regarding the best music features or universal features to represent musical "content" such as semantic representation (word embeddings or document embeddings) which are cross-lingual in linguistics. Music representations that represent musical "content" and are cross-genre music are very important for analyzing music such as: musical similarity, automatic grouping of music, music retrieval, and detecting plagiarism in musical compositions from various musical genres.

3) Optimal music model for music analysis purposes, although a number of researchers in the field of music have explored and reported several statistical models for several computational problems, for example classification and clustering, but there is no model that can achieve the best performance for each music dataset used as input

Transformer is a concept in Natural language Processing (NLP) in Deep Learning [3].

Transformer is part of NLP which is an open source library [3]. Natural Language process uses many words that are processed through transformers [3]. Several studies using transformers, namely the application of machine learning using the transformer method in the health sector [5]. Transformers used to examine people's views through politics on social media apply BERT, BERT, LSTM, Support Vector machine, Decision Trees, Naïve Bayes, Electra. In this study, the highest accuracy value used the Electra model with a value of 70% [6].

The use of the TurnGPT transformer model used in spoken dialogue [7]. Research that applies transformers to musical chord recognition using the bi-directional Transformer for chord recognition (BTC) method [8]. Transformer using block model [9]. Transformer imaging using 3D on the plane [10]. The transformer used for sentiment analysis uses the BERT model [11]. Machine translation uses the concept of transformers [12]. Transformers were used in the best food review using BERT [13]. Transformers are used to assess tourist visits [14]. Sentiment analysis on film reviews using transformers [15]. Transformer introduces BERT and GPT with encoder and decoder [4]. Transformers are used for text compression in readings [16].

Transformers are used to recognize chemical images with an accuracy rate of 96% [17]. Transformers are used to detect a person's emotions [18]. Transformer to detect emotions in social media conversations, for example on Facebook with categories of happy, sad and angry [18]. Machine learning that uses a transformer model with 1000+ transformers to analyze health problems [19]. Transformers were used to predict the next 10 weeks with influenza case subjects [20].

Journal of Theoretical and Applied Information Technology
31st December 2024. Vol.102. No. 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. LITERATURE REVIEW

The transformer

This study uses some literature reviews consist of:

A. Spectogram

Music Spectrum sequences that can be described with two-dimensional images, namely time and frequency, are called spectrograms like in Figure 1. The color brightness in the image shows the frequency of the music spectrum, and by incorporating music sounds into the image, a complexity neural network based model can be used for further music recognition [17]. Sonographs, voiceprints, or voicegrams are other terms for a spectogram in music that shows the signal strength of several sound wave frequencies over a period of time.

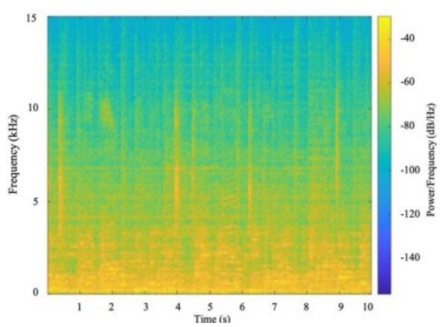


Fig 1. Spectogram [18]

B. Chromagram

Chromagaram in Figure 2 is a musical feature that describes the pitch profile (pitch class) of music. In music, the term "pitch" indicates how high or low a sound is, depending on frequency. Musical pitches are grouped into twelve categories in a chromagram, namely C, C#, D, D#, E, F, F#, G, G, A, A, B. These categories are usually used to encode harmonies with variations in octave height, loudness, and timbre. In addition, the key and chord distribution of the pitch is described in the pitch category [19]. The chroma_stft function from the Python librosa library is used to extract the resulting music samples.

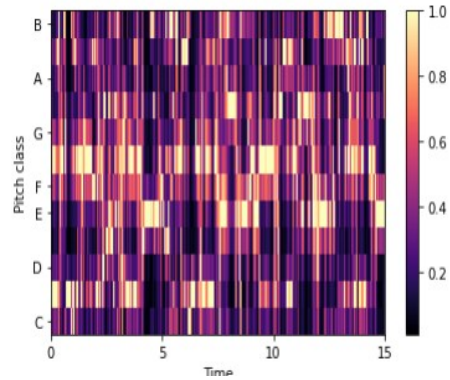


Fig 2. Chromagram

C. Hyperparameter

The term "hyperparameter" refers to the number of model parameters that come into play before or during the learning process of the model. Hyperparameters help in controlling the training data. Since hyperparameters are very important for choosing an optimal model, many methods for selecting hyperparameter values have been proposed, such as random search, grid search, and Bayesian optimization. However, some researchers suggest

using search for machine learning algorithms for optimizing the computational complexity [20]. Besides that, according to several studies, the best way to find the hyperparameter value, for example in neural network of the model greatly affects machine learning performance [21].

D. BERT

BERT (Bidirectional Encoder Representations from Transformers) is an embedding layer designed to train deep bidirectional representations of unlabeled text by jointly conditioning on left and right contexts in all layers [16]. BERT is the latest innovation that represents learning contextually [22].

BERT in Figure 3 is a reading review, extracted aspects and classification of sentiment aspects [23]. INDOBERT is a variant transformer with an Indonesian language corpus. Indobert is a transformer using BERT [24]. Several studies use IndoBert, namely an Indonesian BERT-based model trained on the Indo4B set [25]. INDOBERT which is used in INDOLEM covers various morpho-syntactic, semantic and discourse analysis competencies for Indonesian, to be a benchmark for progress in Indonesian NLP [26]. IndoBert which focuses on the context and input sentences for Indonesian fake news [27].

Journal of Theoretical and Applied Information Technology
31st December 2024. Vol.102. No. 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

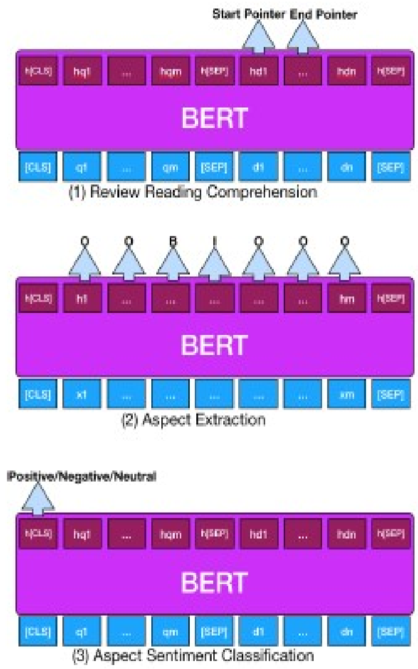


Fig 3. BERT [23]

3. RESEARCH METHODOLOGY

CRISP-DM (Cross Industry Standard Process for Data Mining) was designed in 1996 [28] as a research design framework. CRISP-DM has six phases (see Figure 4), namely business understanding, data understanding, data preparation, modeling, evaluation, and implementation [29][30]. CRISP-DM was originally a framework commonly used in the field of data mining [31].

This framework is then applied in various data processing with data mining methods [32]. The advantages of CRIP-DM are mainly its ability to handle big data, including in the field of Data Science. In addition, CRISP-DM is widely used in various scientific fields due to its well-defined sequence of steps [33].

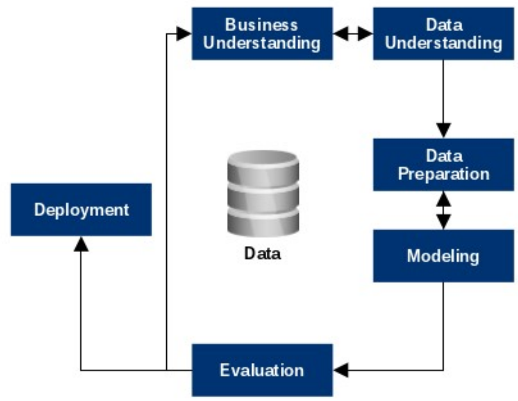


Fig 4. CRISP-DM [30]

A. Business Understanding and Data Understanding

Business understanding and Data understanding is done by preparing a dataset in the form of lyrics and audio data. Audio data is stored in .csv format and lyrics data is stored in excel. Business understanding is carried out with the aim of developing a music mood classification model through public music data sources. Lyric data processing using Bert. Where BertTokenizer is a tokenizer based on the BERT base model. This research uses IndoBertTokenizer. While the audio data uses the Spectogram, Chromagram Convert audio data into waveform using the librosa library. The description of the data is in the form of a data set containing song lyrics from the 70s and 80s in csv format. The song lyrics used as input are the chorus. Music audio datasets in .wav format. The dataset has been labeled like in Figure 5 (3 mood categories): happy (0), sad (1), and neutral (2)[34].

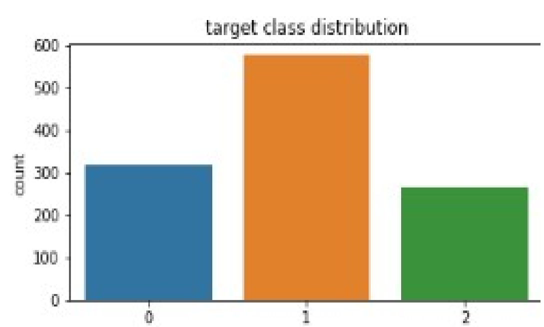


Fig 5. Target Class Distributiom

B. Data Preparation

The number of samples is 1,160. After matching the description of the data into data for training, namely 1.053 samples (lyrics and

Journal of Theoretical and Applied Information Technology
31st December 2024. Vol.102. No. 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

audio). The happy category is 376 samples. The sad category is 357 samples. The neutral category is 320 samples. cross- validation method which hold out 80 % data training and validation 20 % data testing.

C. Modelling

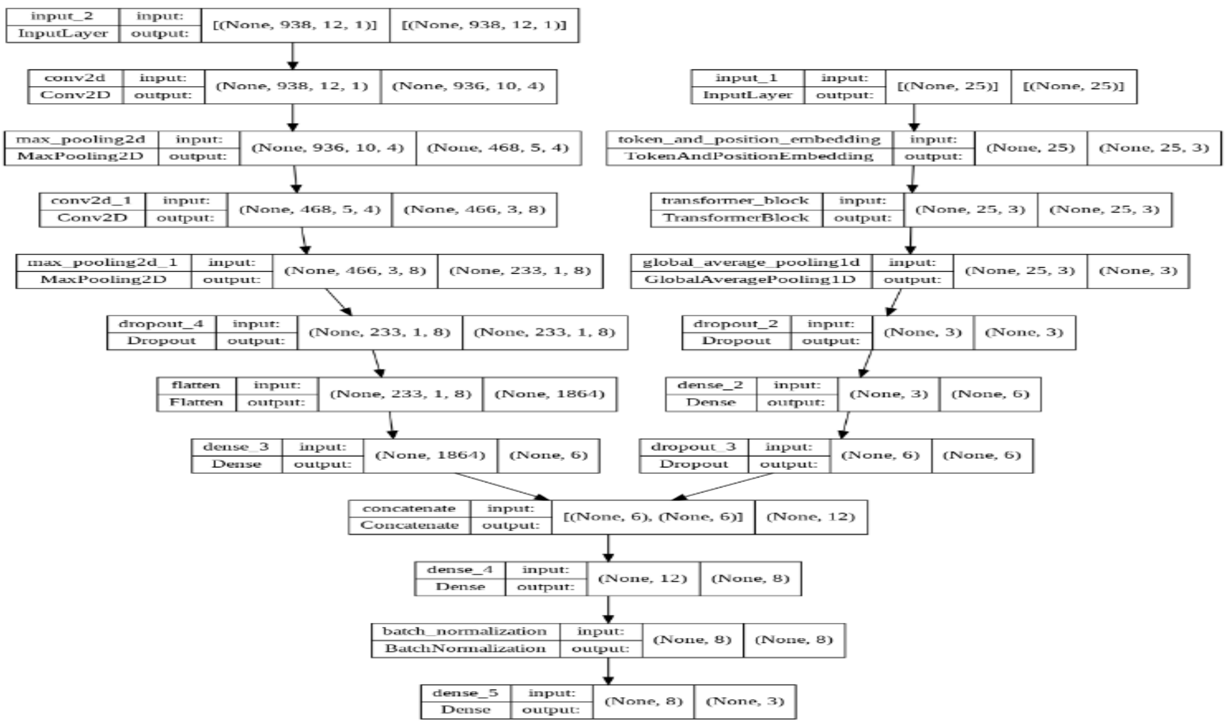


Fig 6. Model Transformer Decision Fusion

Fig 6 The modeling data used consists of several stages, The modeling data used consists of several stages, where the initial stage in Figure 6 is decision fusion, this stage is a type of multimodal. In Figure 6 it is divided into two, namely text and audio processing. The stage in text data is embedding which is a vector. The embedding used is IndoBert, where IndoBert is a variant of transformer which is an embedding token. From words, text is transformed or raw words are converted into tensors which receive word embeddings as input and then process the word embeddings. Indobert tried to understand the text and then calculated the average.

Besides that, the audio provided is only a .wav signal. The next step is converted into a spectrogram or chromagram. After that, the data is processed in the Gated Recurrent Unit (GRU) to understand the audio which is the variance of the RNN. GRU is used to study data that has and understand sequence. To understand the modality provided by GRU and processed into a Fully Connected layer, an average calculation is carried out. After the dimensions are the same between the text and audio, the merge is carried out using concatenate. After that, a new fully connected layer is carried out to change the shape, for example linear transformation. The final stage is mapping to probability distribution or softmax.

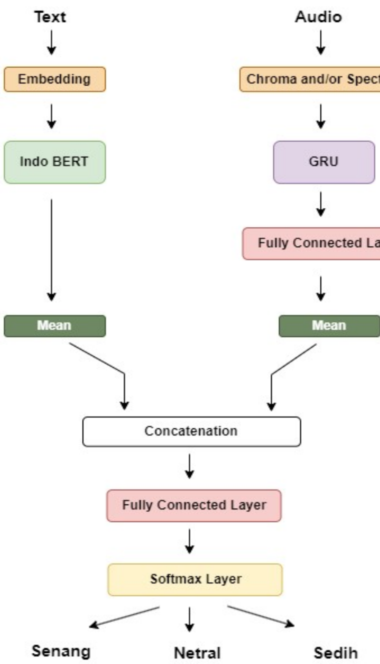


Fig 7. Model Processing Transformer

The libraries used include OS for the operating system, Pandas for processing text data and processing dataframes, Librosa for processing audio data and many more libraries used as shown in attachment part 1. The next step is to process the lyrics or text data first. What is done is loading the dataset, handling text data, and building the model. After loading the dataset, text processing or cleaning the text data begins with several steps. This step consists of first lowering the text, namely deleting uppercase letters. The second step is to delete new lines, namely punctuation marks, numbers, words that consist of only 1 or 2 letters. The third step is to delete non-ASCII characters or letters outside the alphabet such as Japanese letters, delete emote icons, delete double or more than 2 spaces, delete starting and ending spaces.

Initialization is carried out by collecting stop words in Indonesian. After that, clean the data stored in reff_cleaned. Then df.shape will check the data size, namely (1160, 7). 1160 is the number of rows and 7 is the number of columns.

The next search matching is by checking the song titles between the text dataset folders. Before the matching process, it is necessary to first check the amount of data. The initial data consists of 272 happy, 579 sad, 321 neutral. So the process of deleting data that doesn't match is done with df_filter. After filtering or matching the data then there were 981 with the happy category being 160 and label 0, the sad category being 357 and label 1, and the neutral being 188 and label 2. Splitting the dataset into training data and testing data was carried out, 90% of the training and validation data , 10% testing data.

The distribution of each label is still unbalanced, labels 2 and 0, namely 50% less than label 1, so it could be a problem if the data is directly used for the training process. This is because the data is not balanced and there is a possibility of Bayes or the model considering model 1 as the strongest label and in the end it will make the model predict whatever data is labeled 1. So that the machine learning model is smarter, augmentation needs to be done for labels 0 and 2. Augmentation is carried out by adding data using the back translation method.

The process of changing data which was initially in the form of Indonesian text data into English and then translated back into Indonesian. Usually when you have translated data in it, for example Google Translate, it will be slightly different but still refers to the same words. This technique can be used because there are changes in grammar but in terms of meaning it is still the same. After augmenting the data, it becomes 1053 samples so that the data distribution is not too far apart, where the data is now label 1, namely sad, with a total of 357, label 0, namely happy, with a total of 320, label 2, namely neutral, with a total of 376. Thus, the processing of the text data is complete. After that, the process of changing the text into numbers so that it can be recognized by the computer at the stage after processing the audio data.

The next processing is audio data. For audio data it is almost the same. The first is to extract the feature chromagram with a sampling rate of 16000 and an audio duration of 30 or 30 seconds. When there is data whose duration is less, it can be added with the number 0. However, if the duration is more than 30 seconds, it will be cut to 30 seconds. Augmentation of audio data is different from augmentation of text data. Augmentation of audio data is done by adding noise, pitch and stretch. The process of adding noise is not too big, it just blurs the previous audio with a slight difference. noise creates a function to create random noise for audio data using numpy random of 0.048 with a seed of 42. After that using numpy random uniform, that is, the data will be distributed uniformly over half the open interval so it is between high and low. Then the data is added with noise which is multiplied by the numpy random return data. So the data is added with noise which is multiplied by numpy random return data.

Journal of Theoretical and Applied Information Technology
31st December 2024. Vol.102. No. 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The addition of stretch with a rate of 0.6 for the stretch function stretches the time of the audio series with a predetermined rate, via librosa.effects.time_stretch(data, rate). Stretch is used if the rate is more than 1 then the signal is accelerated but if the rate is less than 1 then the signal is slowed down.

Added pitch with sampling_rate and pitch_factor of 0.6. The addition of Noise, Pitch and Stretch is carried out on data that is not unbalanced on happy and neutral data. The augmented data is saved in the chromagram folder in .pkl or pickle format. After having the dataset_chroma folder, the feature extraction stage is no longer carried out, because you already have feature data that has been saved and saves time in the data training process.

Then the data was augmented to become label 1, namely sad, with a total of 357, label 0, namely happy, with a total of 320, label 2, namely neutral, with a total of 376, the same as the text data. After ensuring that the label order is the same between the text data, namely lyrics and audio data, with train data, test data and validation. So that 1053,1053 are produced between text data and audio data.

The next processing of text data is tokens or tokenization because computers can only understand data in the form of numbers or sequences of numbers with the get_dataset command. Get this dataset with dictionary format. Converting sentences to numbers with tokenization. In this research, the transformer model uses IndoBert. IndoBert is a rule model with an Indonesian corpus. This model is considered good because it can process millions of Indonesian words from websites, Wikipedia or news. Then determine the sequence of length. This sequence of length is the same as audio data, namely set duration. The sequence length used in the lyrics data is 25. So if the lyrics are more than 25 it will be truncated but if the lyrics are less than 25 it will be added 0. After that the data is compared or distributed, the text data and audio data with the train data on the text totaling 1053 and matrix 25. The total number of audio data is 1053, with a matrix size of 981 x 12.

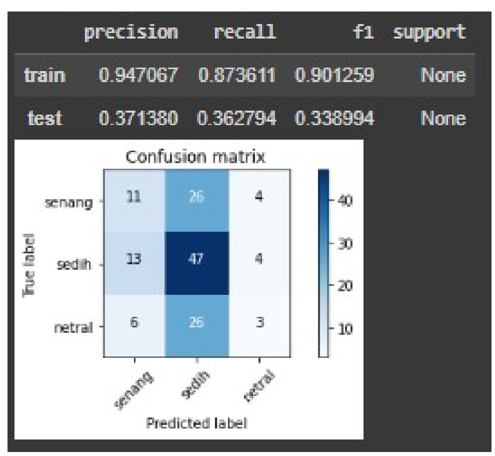
Next, the block model defines a transformer with the num class parameter, namely the number of classes with sad, happy and neutral. And look for the number of different words using vocab_size, namely 31923. Then use hyperparameter tuning for the transformer model, namely EMBED_DIM=3, FF_DIM=4 for the output layer, NUM_HEAD=2 for processing transformers and pattern recognition from the data held.

After that, build a text and audio model like Figure 7 which consists of variable X which is lyrics or text and variable Y which is audio. To unify all the steps, define the final model by adjusting the text data and audio data. The left side is audio data. To the right of the text data. There are 2 architectures which are combined using concatenate, then processed so that the output comes out using softmax.

After that, training is carried out using IF ELSE logic. The if condition is if the folder called multimodal_v4 is already owned then no training is carried out. Just load the model so the multimodal_v4 model will only be called. But if the folder doesn't exist then the model will run and the output will be a series of epochs, so there is progress. By paying attention to the model training process and monitoring the increase and increase in performance from epoch 1 to the last epoch. However, when the increase always remains at a certain epoch, early stopping is carried out, namely the epoch condition stops at the beginning when there is no further increase.

D. Evaluation

The evaluation in Figure 8 is the result of the confusion matrix on text and audio data. The evaluation resulted in Xgboost Clasifier when in data train precision 94%, recall 87% f1 90%. The results of the Confusion Matrix in Figure 9 precision 98%, recall 98% f1 98% are the results of multimodal audio and lyric data training using a transformer model.



	precision	recall	f1	support
train	0.947067	0.873611	0.901259	None
test	0.371380	0.362794	0.338994	None

Fig 8. Evaluation Xgboost Clasifier

	precision	recall	f1-score	support
senang	0.98	1.00	0.99	40
sedih	0.98	0.99	0.98	90
netral	1.00	0.96	0.98	47
accuracy			0.98	177
macro avg	0.98	0.98	0.98	177
weighted avg	0.98	0.98	0.98	177

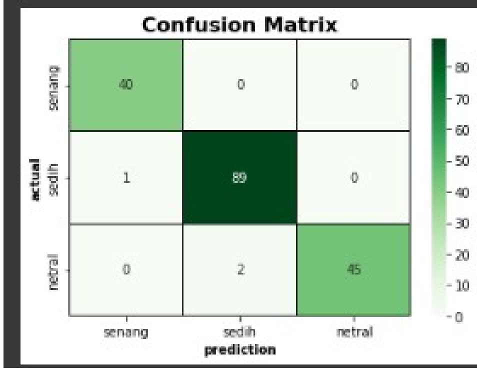


Fig 9. Evalution Transformer Model

5. CONCLUSION

The conclusions in this study consist of:

a. This study uses audio datasets and Indonesian song lyrics in the 70s and 80s.

b. This study uses the CRISP-DM method

c. This study uses parameters, namely 80% training data, 20% test data, wav sequential and text data in the form of the Bert model.

d. This study uses a number of classes that are used, namely through sad, happy and neutral moods. The number of different words used is 31923 words.

e. The evaluation resulted in Xgboost Clasifier when in data train precision 94%, recall 87% f1 90%.

f. The results of the Confusion Matrix in Transformer precision 98%, recall 98% f1 98% so that the best accuracy is produced by the transformer model.

g. This research only uses lyrical data on the chorus, for further research it uses the intro and the end of the song by way of comparison.

h. This research only uses high level features and no experiments have been carried out on low level audio

REFERENCES :

[1]. R. G. Yaqin A, “Penilaian Kredit Menggunakan Algoritma XGBoost dan Logistic Regression,” vol. 8, no. 1, pp. 4–10, 2023.

[2]. R. Rismala, L. Novamizanti, K. N. Ramadhani, Y. S. Rohmah, and S. Parjuangan, “Kajian Ilmiah dan Deteksi Adiksi Internet dan Media Sosial di Indonesia Menggunakan XGBoost,” vol. 7, no. 1, pp. 1–11, 2021.

[3]. R. Gunawan et al., “Jurnal Computer Science and Information Technology (CoSciTech),” vol. 3, no. 3, pp. 453–463, 2022.

[4]. I. R. Amiruddin, “Penerapan Algoritma XGBoost untuk Seleksi Atribut pada K-Means dalam Clustering Penerima,” vol. 5, pp. 192–196, 2023.

[5]. M. Alkaff, A. R. Baskara, T. Informasi, F. Teknik, and U. L. Mangkurat, “Penerapan Metode XGBoost Untuk Memprediksi Jumlah Kejadian Kecelakaan Lalu Lintas di Kota Banjarmasin,” vol. 7, no. 1, pp. 61–69, 2023.

[6]. J. Iskandar, V. C. Mawardi, and J. Hendryli, “Analisis Media Sosial Penyedia Layanan Internet Menggunakan Algoritma XGBOOST”.

[7]. M. R. Kurniawanda, F. Adline, T. Tobing, and P. S. Informatika, “Analysis Sentiment Cyberbullying in Instagram Comments with XGBoost Method,” vol. 9, no. 1, 2022.

[8]. D. Ananda, A. Pertiwi, U. N. Semarang, M. A. Muslim, and U. N. Semarang, “Prediksi Rating Aplikasi Playstore Menggunakan Xgboost Prediksi Rating Aplikasi Playstore Menggunakan Xgboost,” no. October 2020, 2022.

[9]. T. Widjaja and A. Gunawan, “Deteksi Plagiarisme pada Kode Bahasa Pemrograman Java menggunakan XGBoost”.

[10]. S. Elina, H. Yulianti, O. Soesanto, and Y. Sukmawaty, “Penerapan Metode Extreme Gradient Boosting (XGBOOST) pada Klasifikasi Nasabah Kartu Kredit,” vol. 4, no. 1, pp. 21–26, 2022.

[11]. R. Reinauer, M. Caorsi, and N. Berkouk, “Persformer : A Transformer Architecture for Topological Machine Learning arXiv : 2112 . 15210v2 [cs . LG] 26 Sep 2022,” pp. 1–27, 2022.

[12]. D. Luitse, “The great Transformer : Examining the role of large language models in the political economy of AI,” 2021, doi: 10.1177/20539517211047734.

[13]. S. Singla and N. Ramachandra, “Comparative Analysis of Transformer Based Pre-Trained NLP Models International Journal of Computer Sciences and Engineering Open Access Comparative Analysis of Transformer Based Pre-Trained NLP Models,” no.

December, 2020, doi: 10.26438/ijcse/v8i11.4044.
[14]. B. Ghojogh, A. L. I. Ghodsi, and U. Ca, "Attention Mechanism , Transformers , BERT , and GPT : Tutorial and Survey," 2017.
[15]. A. Vaswani, "Attention Is All You Need," no. Nips, 2017.
[16]. M. Munikar, S. Shakya, and A. Shrestha, "Fine-grained Sentiment Classification using BERT," Int. Conf. Artif. Intell. Transform. Bus. Soc. AITB 2019, pp. 2–5, 2019, doi: 10.1109/AITB48515.2019.8947435.
[17]. Turetzky, T. Michelson, Y. Adi, and S. Peleg, "Deep Audio Waveform Prior".
[18]. M. Koseoglu and H. Uyanik, "Effect of Spectrogram Parameters and Noise Types on The Performance of Spectro-temporal Peaks Based Audio Search Method," Gazi Univ. J. Sci., vol. 36, no. 2, pp. 624–643, 2023, doi: 10.35378/gujs.1000594.
[19]. Z. Zhang et al., "Deep Audio Priors Emerge From Harmonic Convolutional Networks," 8th Int. Conf. Learn. Represent. ICLR 2020, no. 2019, pp. 1–12, 2020.
[20]. Y. A. Ali, E. M. Awwad, M. Al-Razgan, and A. Maarouf, "Hyperparameter Search for Machine Learning Algorithms for Optimizing the Computational Complexity," Processes, vol. 11, no. 2, pp. 1–22, 2023, doi: 10.3390/pr11020349.
[21]. A. Koutsoukas, K. J. Monaghan, X. Li, and J. Huan, "Deep-learning: Investigating deep neural networks hyper-parameters and comparison of performance to shallow methods for modeling bioactivity data," J. Cheminform., vol. 9, no. 1, pp. 1–13, 2017, doi: 10.1186/s13321-017-0226-y.
[22]. P. Knees, E. Pampalk, and G. Widmer, "Artist Classification with Web-Based Data . ARTIST CLASSIFICATION WITH WEB-BASED DATA," no. January 2004, 2014.
[23]. H. Xu, B. Liu, L. Shu, and P. S. Yu, "BERT post-training for review reading comprehension and aspect-based sentiment analysis," NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf., vol. 1, pp. 2324–2335, 2019.
[24]. M. C. Kenton, L. Kristina, and J. Devlin, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," no. Mlm, 1953.
[25]. B. Wilie et al., "IndoNLU : Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," pp. 843–857, 2020.
[26]. B. T. Koto F, Rahimi A, Lau JH, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," 2020.
[27]. S. M. Isa, G. Nico, and M. Permana, "Indobert for indonesian fake news detection," vol. 16, no. 3, pp. 289–297, 2022, doi: 10.24507/icicel.16.03.289.
[28]. R. Wirth and J. Hipp, "CRISP-DM: towards a standard process model for data mining. Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining, 29-39," Proc. Fourth Int. Conf. Pract. Appl. Knowl. Discov. Data Min., no. 24959, pp. 29–39, 2000, [Online]. Available: https://www.researchgate.net/publication/239585378_CRISP-DM_Towards_a_standard_process_model_for_data_mining
[29]. W. Y. Ayele, "Adapting CRISP-DM for idea mining a data mining process for generating ideas using a textual dataset," Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 6, pp. 20–32, 2020, doi: 10.14569/IJACSA.2020.0110603.
[30]. T. Thoyyibah, E. Abdurachman, Y. Heryadi, and A. Zahra, "CRISP-DM Method for Mood Classification in Indonesian Music 70 and 80 era," Int. J. Appl. Eng. Technol., vol. 4, no. 1, pp. 45–49, 2022.
[31]. T. Darmawan, "Credit Classification Using CRISP-DM Method On Bank ABC Customers," Int. J. Emerg. Trends Eng. Res., vol. 8, no. 6, pp. 2375–2380, 2020, doi: 10.30534/ijeter/2020/28862020.
[32]. S. C, "The CRISP-DM Model: The New Blueprint for Data Mining," J. Data Warehous., vol. 26, no. 1, pp. 39–48, 1981, doi: 10.1016/0022-4073(81)90019-4.
[33]. M. Cazacu1 and E. Titan, "BRAIN. Broad Research in Artificial Intelligence and Neuroscience Adapting CRISP-DM for Social Sciences," vol. 11, pp. 99–106, 2020, [Online]. Available: https://doi.org/10.18662/brain/11.2Sup1/97
[34]. P. J. Upadhyay, S. Shetty, V. Murari, and J. Trinidade, "Mood Based Music Playlist Generator Using Convolutional Neural Network," Int. J. Res. Appl. Sci. Eng. Technol., vol. 10, no. 3, pp. 531–540, 2022, doi: 10.22214/ijraset.2022.40668.