The IJICS (International Journal of Informatics and Computer Science)
Vol 10 No 2, July 2026, Page 115-124
ISSN 2548-8384 (online), ISSN 2548-8449 (print)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/ijics/index
DOI 10.30865/ijics.v10i2.9775

Classification of Coming-of-Age Song Lyrics Using Convolutional Neural Network (CNN) Architecture

Arjon Samuel Sitio[1,*], Fricles Ariwisanto Sianturi[2], Anita Sindar[3]
[1,2] Information Systems, Faculty of Science and Technology, Tjut Nyak Dhien University, Medan, Indonesia
[3] Information Technology, STMIK Pelita Nusantara, Medan, Indonesia
Email: [1,*] arjonsitio@yahoo.com, [2]sianturifricles@gmail.com, [3]haito_ita@yahoo.com
(* : arjonsitio@yahoo.com)
Submitted: **05/05/2026**; Accepted: **09/06/2026**; Published: **29/06/2026**

Abstract− The coming-of-age theme in song lyrics is rich in emotional expressions, identity transitions, and nostalgia. This study implements a 1-Dimensional (1D) Convolutional Neural Network (CNN) architecture to automatically classify coming-of-age-themed song lyrics. Raw lyrics data were collected through a custom scraping function get_song_lyrics, which then went through a structured text preprocessing stage to remove stopwords and non-semantic components. The cleaned words were represented into a vector space using a pre-trained 50-dimensional GloVe embedding (GloVe 50d) of size (max_length, 50) to capture semantic relationships between words. A 1D CNN model was applied to extract local features in the form of phrase combinations (n-grams) through filter shifting, followed by a Global Max Pooling layer to filter out the most dominant emotional information before the final classification process. As a comparison method and additional analysis, a rule-based approach using TextBlob was applied to extract polarity and subjectivity scores, while the Word Cloud technique was used to visualize the dominance of transitional lexical terms such as grow, leave, and remember. The gap between training and validation performance suggests that the model learned dataset-specific patterns rather than generalized semantic representations. Similar overfitting behavior has been reported in CNN-based lyric and sentiment classification studies when training data are limited or insufficiently diverse. The results showed that the integration of GloVe 50d semantic representation and local feature extraction by 1D CNN was able to produce high and stable accuracy in recognizing the unique characteristics of song lyrics with a maturity theme.

Keywords: Song Lyrics; Text Semantics; NLP; TextBlob; CNN Architecture.

1. INTRODUCTION

The genres of songs preferred by Gen Z and Millennials are dominated by modern pop, hip-hop or rap, R&B, and indie or folk, with pop being the most popular genre. Classifying song lyric themes remains challenging due to the ambiguous, metaphorical, and contextual nature of lyrical language. Conventional keyword-matching approaches often fail to capture the semantic meaning and emotional relationships between words in lyrics. Classification of song lyrics with a coming-of-age theme using Convolutional Neural Network (CNN) architecture is an effective natural language processing (NLP) approach to recognize emotional patterns, life phase changes, and nostalgia in text. [1]. Natural Language Processing (NLP) is an artificial intelligence technology that enables computers to understand, interpret, and manipulate human language. In song lyric classification projects, NLP acts as a bridge that transforms raw text into structured data that can be processed by CNN models. Natural Language Processing (NLP) is used to transform song lyrics into machine-understandable forms, through steps such as tokenization, stopword removal, and word embedding. NLP plays a crucial role in extracting the linguistic and semantic patterns typical of coming-of-age lyrics.

A Convolutional Neural Network (CNN) was applied as the architecture for the text classification model. CNNs are capable of detecting n-gram patterns and semantic features in lyrics and then grouping them into specific themes. Layers such as Conv1D, MaxPooling, Dropout, and Dense are used for feature extraction, dimensionality reduction, and binary classification (positive/negative) [2]. The Convolutional Neural Network (CNN) architecture is designed to recognize patterns in data, especially images and text. CNNs consist of several main layers: a convolution layer for feature extraction, a pooling layer for dimensionality reduction, and a fully connected layer for classification. The convolution process works with filters (kernels) that move through the data, capturing local patterns such as words or phrases in the text. Pooling helps simplify the representation while retaining important information. The fully connected layer combines all extracted features to produce a final prediction or classification. A pooling layer is used to reduce dimensionality and retain important information, making the model more efficient [3].

Classification of coming-of-age lyric themes according to a teenager's emotional phase is carried out by building a CNN model for semantic analysis of lyrics and music selection. The initial step is to collect a corpus of teenage song lyrics. from official sources and compile theme labels based on the emotional phase of teenagers with the help of experts or annotators [4]. Identifying linguistic and semantic patterns is necessary to recognize typical teenage words or phrases such as bae, squad, or turnup which are indicators of themes.

These studies collectively highlight how Convolutional Neural Networks (CNNs) can transform lyric analysis by automating traditionally subjective and labor-intensive tasks. CNNs enable genre and theme classification by extracting spatial features from word embeddings, detect sentiment and subtle emotional shifts in youth music through multi-layered filters, and provide an objective framework for categorizing life-stage transitions such as maturity in pop lyrics. CNNs also support multi-class motif detection with optimized architectures that reduce computational costs while maintaining accuracy, and facilitate the extraction of linguistic features that bridge psychology and data science by

The IJICS (International Journal of Informatics and Computer Science)
Vol 10 No 2, July 2026, Page 115-124
ISSN 2548-8384 (online), ISSN 2548-8449 (print)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/ijics/index
DOI 10.30865/ijics.v10i2.9775

mapping lyric structure to developmental stages. Taken together, these approaches demonstrate the power of deep learning in uncovering complex themes, emotions, and narratives in large-scale lyric datasets.

This study aims to classify coming-of-age lyric themes based on the emotional phase of teenagers. This study aims to classify song lyrics themes related to the coming-of-age phase using a deep learning approach using a Convolutional Neural Network (CNN) architecture [5]. Song lyrics will be processed through text pre-processing stages (tokenization and embedding), then the CNN plays a role in extracting local features from n-grams to automatically recognize certain emotional or narrative patterns. Through this algorithm, the system is able to classify lyrics into specific themes such as self-discovery, emotional conflict, or life changes with high accuracy, making it an effective method for sentiment and content analysis in song lyrics [6]. Although Convolutional Neural Network (CNN)-based text classification has been widely applied to general sentiment analysis, research on detecting specific coming-of-age themes in song lyrics is still limited, especially in capturing contradictory emotional nuances such as nostalgia and future anxiety [7]. Previous studies often focus on basic emotions, fail to interpret complex metaphors, and have limitations in understanding long-term contexts, so CNN architectures need to be further explored to analyze subtle poetic narratives. This gap is addressed by testing the effectiveness of CNNs in recognizing semantic characteristics and life transitions that are characteristic of coming-of-age songs [8]. The aim of this research is to develop and evaluate a coming-of-age lyrics classification model using CNNs, allowing for accurate identification of themes and emotions within the lyrics.

2. RESEARCH METHODOLOGY

2.1 Research Stages

Following the Natural Language Processing (NLP) pipeline, this research is divided into six stages. The initial stage defines the research problem, namely how to classify coming-of-age lyric themes according to the emotional phases of teenagers. Collecting a dataset of teenage song lyrics from official sources and labeling themes based on emotional phases [9]. The stages of song lyrics classification using CNN architecture include preprocessing processes in the form of text cleaning, tokenization, and padding to ensure uniform lyric length, then word embedding to convert words into numeric vectors such as Word2Vec or GloVe, followed by a convolution layer that extracts important feature patterns from the lyrics, followed by pooling to reduce data dimensions while retaining the most relevant features, and finally the classification stage through a fully connected layer with a softmax function to determine the lyric category.[10].

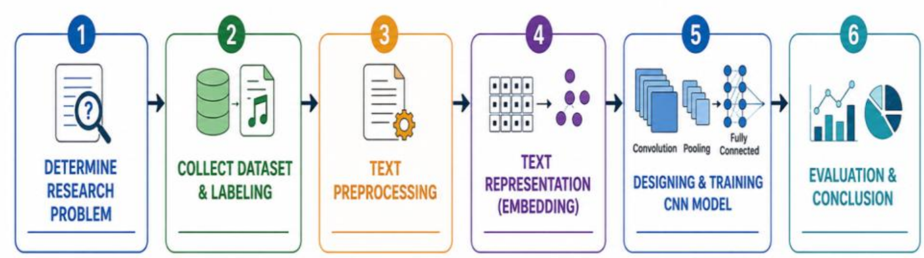

Fig 1. Research Stages

The main objective is to classify song lyrics with the theme of adulthood using a Convolutional Neural Network (CNN) with GloVe word embeddings. This process includes preparing a 'adulthood' song lyrics dataset, text preprocessing, converting words to GloVe embeddings, defining and training a CNN model, and finally evaluating its performance. In the analysis and interpretation stage, the most influential features identified by the CNN model are examined to understand how they contribute to the classification of song lyrics. These findings are then connected with theories of music and adolescent psychology to provide deeper insight into the coming-of-age themes present in the lyrics. The evaluation also considers whether CNN proves effective in capturing and categorizing these themes accurately [11]. In the conclusion and recommendation stage, the overall effectiveness of CNN in classifying coming-of-age song lyrics is summarized, and suggestions for future research are offered, such as exploring more advanced architectures like transformer models, expanding datasets, or conducting cross-cultural analyses to enrich the study [12].

2.2 Feature Representation

NLP provides various techniques for converting text into numeric vectors, as machines require numeric input to process data [13]. This conversion is accomplished through methods that produce dense vector representations capable of capturing semantic meaning. Word embeddings such as Word2Vec, GloVe, and fastText are used to represent words, while pre-trained embeddings such as ELMo and BERT, and document embeddings such as Doc2Vec, expand the scope of representation. Pre-trained GloVe embeddings with 100 dimensions were employed as the word representation method throughout the study. Once text is converted to numbers, the model is trained to recognize patterns and solve various NLP tasks [14].

The IJICS (International Journal of Informatics and Computer Science)
Vol 10 No 2, July 2026, Page 115-124
ISSN 2548-8384 (online), ISSN 2548-8449 (print)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/ijics/index
DOI 10.30865/ijics.v10i2.9775

2.3 Convolutional Neural Network (CNN) Architecture

Although typically used for images, CNNs are particularly effective for text classification due to their ability to capture local patterns such as key phrases. In the CNN architecture for lyric classification, the embedding layer first transforms words into numerical vectors so the computer can interpret their semantic relationships. These vectors are then processed by the 1D convolutional layer, which applies filters across the text to detect meaningful word patterns related to themes like coming-of-age. The pooling layer follows, condensing the convolution results by selecting only the most dominant features, thereby reducing complexity while retaining essential information. Finally, the fully connected layer integrates these features and outputs the probability that a lyric belongs to the Coming-of-Age category or not, completing the classification process. The main parts of the Convolutional Neural Network (CNN) architecture are the Input Layer, Convolution Layer, Pooling Layer, Fully Connected Layer, and Output Layer [15]. In designing a CNN model for text classification, the architecture typically begins with an Embedding layer that converts words into dense vector representations, followed by Convolutional layers that extract local n-gram features from the sequences [16]. The CNN architecture consisted of an Embedding Layer initialized with GloVe 100-dimensional vectors, followed by a Conv1D layer with 128 filters and a kernel size of 5, a Global Max Pooling layer, a Dropout layer (0.5), and a Dense output layer with sigmoid activation. The model was trained using the Adam optimizer with a learning rate of 0.001, binary cross-entropy loss, batch size of 32, and 10 epochs.

2.4 Semantic Analysis

Sentiment analysis using TextBlob is done by calculating two main values from the lyrics text, namely Polarity (positive/negative score between -1 to 1) and Subjectivity (objective/subjective score between 0 to 1). Polarity (> 0): Words like "beautiful" are given a positive weight. Polarity (< 0): Words like "scared" are given a negative weight. Song lyrics generally have a score close to 1 because they reflect the personal feelings of the writer, not scientific facts. NLP helps models understand the intent and feeling behind lyrics, not just matching keywords. In the case of coming-of-age songs, NLP allows CNNs to detect lyrical patterns that reflect the transitional phases of adolescence, such as angst or the euphoria of freedom. Semantic analysis allows models to distinguish between similar words with different meanings depending on the sentence [17]. For example, the word fall in teen lyrics can mean falling in love or falling down [18]. This analysis helps CNNs recognize figurative language patterns or metaphors often used in coming-of-age songs, so that the model does not get stuck on literal meanings but instead understands the feel of the emotional phase.

3. RESULT AND DISCUSSION

Text preprocessing involves cleaning the lyrics by removing punctuation and converting all text to lowercase, after which sentiment or emotion analysis is performed using pre-trained lexicons or models such as TextBlob, or advanced Transformer-based approaches to assign sentiment scores or emotion labels to each lyric or song; finally, these emotion scores are aggregated at the song or album level and visualized to reveal overall emotional patterns and trends. Comprehensive preprocessing of the obtained maturity-themed song lyrics dataset. This includes text cleaning, tokenization, creation of numeric text representations (word embeddings), and padding of text sequences to ensure uniform length. Furthermore, thematic labels are encoded numerically. Automatically fetches song lyrics from the internet based on the entered song title and artist name. This function typically works by sending an HTTP request to a lyrics provider API such as the Genius API. Interpret the performance of CNN models, summarize the findings and conclusions from the research on song lyric theme classification about adolescence, including the effectiveness of CNN architecture and potential areas for future improvement.

3.1 Extraction of the Dataset

Using API tokens, song lyrics are collected with themes about growing up and towards adulthood with happiness, love, togetherness, enjoying life, enthusiasm, and ambition. Dataset diperluas menjadi 1.100 lirik lagu dengan sumber utama dari internet melalui Genius API. A visual representation showing the most frequently occurring words in a preprocessed dataset of coming-of-age song lyrics. The larger a word is in the visualization, the higher its frequency. In the coming-of-age themed lyrics dataset, a predominance of words that reflect life transitions emerged.

The IJICS (International Journal of Informatics and Computer Science)
Vol 10 No 2, July 2026, Page 115-124
ISSN 2548-8384 (online), ISSN 2548-8449 (print)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/ijics/index
DOI 10.30865/ijics.v10i2.9775

Fig 2. Word Cloud of Coming-of-Age Preprocessed Lyric

The newly obtained 'coming-of-age' lyrics were preprocessed (lowercase, remove punctuation, remove stop words, and stem) and updated the preprocessed lyrics column. After labeling the lyrics based on dominant themes, they were combined into category-specific text sets to help the model better recognize semantic patterns. Labeling is performed based on adolescent emotional phases and dominant themes such as happiness, love, or togetherness. The program prepares CNN input with GloVe embeddings, splits the data into training and test sets, trains the CNN model, and then displays a training history plot. This includes loading the GloVe model and defining its parameters. This ensures all necessary components are in place before the CNN input sequence is prepared.

Table 1. Label Dataset Song about Coming of Age

Coming-of-age Title Label 1	Non Coming-of-age Title Label 0
'Seventeen', 'Growing Up', 'Teenage Dirtbag', 'Summertime Sadness', 'Young Dumb & Broke', 'Castle on the Hill', 'Stressed Out', 'Photograph', 'Vienna', 'Landslide', 'When We Were Young', 'The Kids Are Alright', 'Wake Me Up When September Ends', 'Forever Young', 'Glory Days', 'Graduation (Friends Forever)', 'High School Never Ends', 'My Wish', 'Sweet Sixteen', 'Summer of 69', 'Mr. Brightside', 'Welcome to the Black Parade', 'Hey There Delilah', 'I Will Remember You', 'Don\'t Stop Believin\'', 'A Quiet Reckoning With Change and The Passage of Time', 'The Reckless Joy Of Being Young And Figuring It Out', 'The Bittersweet Farewell That Became A Graduation Staple Raw Emotion At The Boundary Between Youth And Adulthood', 'Mourning A Version Of Yourself Or A Relationship You've Outgrown' ...#More Coming-of-age	'Bohemian Rhapsody', 'Stairway to Heaven', 'Hotel California', 'Billie Jean', 'Like a Rolling Stone', 'Yesterday', 'Imagine', 'One', 'Hallelujah', 'Smells Like Teen Spirit', 'Sweet Child o\' Mine', 'Livin\' on a Prayer', 'Wonderwall', 'Losing My Religion', 'Every Breath You Take', 'Bad Romance', 'Rolling in the Deep', 'Single Ladies (Put a Ring on It)', 'Uptown Funk', 'Happy', ...#More Non Coming-of-age

This dataset consists of 1,100 English song lyrics collected from available lyric repositories pulled through the Genius API. The dataset is divided into two categories: Coming-of-Age song 905 lyrics (labeled as 1) and Non-Coming-of-Age song 195 lyrics (labeled as 0). Duplicate songs were removed by matching song titles and lyric content hashes. The labeling process was performed based on thematic criteria derived from adolescent development literature, including identity formation, personal growth, transition to adulthood, independence, nostalgia, and future aspirations. To ensure the reliability of the model evaluation, the dataset was divided into 70% training data (635 lyrics), 15% validation data (135 lyrics), and 15% testing data (135 lyrics).

Before feeding the lyrics into the CNN model, the text underwent three essential preprocessing stages: tokenization and categorization, where each word in the lyrics was converted into a numeric index; Padding, which ensures all sequences are the same length, and embedding layers, which transform these numerical indices into dense vectors that capture semantic relationships between words, enable CNNs to better understand and learn from the contextual meaning of lyrics. CNNs identify themes of adulthood through several key steps. CNNs use filters that slide over lyrics to capture short word patterns (n-grams). A filter can learn to recognize word combinations associated with growth, transition, or nostalgia. This architecture is highly effective at detecting certain phrase structures that frequently appear in songs about adulthood, even when the word order varies slightly. The results of these filters are mapped to determine which features are most dominant in a song.

Table 2. Preprocessed Lyrics

	Title	Lyrics	Preprocessed Lyrics
0	Seventeen	They say when you're seventeen, you're not sup...	say your seventeen your suppos free
1	Growing Up	I'm thirty years old and still living at	im thirti year old still live home im

The IJICS (International Journal of Informatics and Computer Science)
Vol 10 No 2, July 2026, Page 115-124
ISSN 2548-8384 (online), ISSN 2548-8449 (print)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/ijics/index
DOI 10.30865/ijics.v10i2.9775

		home,...	grow
2	Teenage Dirtbag	I'm just a teenage dirtbag baby, like you...	im teenag dirtbag babi like
3	Summertime Sadness	Kiss me hard before you go, summertime sadness...	kiss hard go summertim sad
4	Young Dumb & Broke	Young dumb broke high school kids...	young dumb broke high school kid

The data cleaning process starts from the identification stage by scanning for text similarities in the lyrics column, followed by reduction, namely removing duplicate rows and retaining the first entry, then validation to ensure the final DataFrame size is consistent and free of duplication, and ends with the readiness to produce a clean dataset that is ready to use. Deduplication of df_coming_of_age_lyrics was successfully applied, ensuring each song lyric is unique and preventing overfitting due to data redundancy. After filtering based on the lyrics column, the DataFrame was cleaned to only valid rows. The first five rows show a list of unique songs ready for model training, such as Seventeen (Ladytron), Growing Up (Bruce Springsteen), and Teenage Dirtbag (Wheatus).

Table 3. Remove Duplicate Entries

Title	Artist	Lyrics	Coming of Age Label	Preprocessed Lyrics	CNN Input Glove
0	Seventeen	Ladytron	They say when you're seventeen, you're not sup...	1	say your seventeen your suppos free
1	Growing Up	Bruce Springsteen	I'm thirty years old and still living at home,...	1	im thirti year old still live home im grow
2	Teenage Dirtbag	Wheatus	I'm just a teenage dirtbag baby, like you...	1	im teenag dirtbag babi like
3	Summertime Sadness	Lana Del Rey	Kiss me hard before you go, summertime sadness...	1	kiss hard go summertim sad
4	Young Dumb & Broke	Khalid	Young dumb broke high school kids...	1	young dumb broke high school kid

Sentiment score interpretation typically involves two primary dimensions: polarity and subjectivity. Polarity ranges from -1 to 1, with values closer to 1 reflecting positive sentiment, values closer to -1 reflecting negative sentiment, and values closer to 0 indicating neutrality. Subjectivity, on the other hand, ranges from 0 to 1, with scores closer to 1 indicating a text is more opinionated, emotional, or subjective, while scores closer to 0 indicate a text is more factual, objective, or neutral. These measures help capture the emotional direction and degree of personal bias or factuality in a text. This graph will demonstrate the diversity of sentiment within this theme. A histogram with a Kernel Density Estimate (KDE) plot shows the distribution of sentiment polarity for Coming-of-Age lyrics. Sentiment scores are centered around slightly positive values, with a wider spread indicating diverse emotional content within this category. Distribution of Sentiment Polarity for Coming-of-Age Lyrics refers to a statistical visualization that shows how sentiment scores (from extreme negative to extreme positive) are distributed across a dataset of coming-of-age song lyrics. This visualization is crucial for understanding whether coming-of-age songs tend to have a uniform emotional content or whether they vary widely. Coming-of-age lyrics graphs typically form a bell curve centered slightly above 0 (e.g., in the 0.1 to 0.2 range). This reflects the bittersweet nature of coming-of-age themes—combining sadness over leaving childhood (negative) with optimism for the future (positive). High Variance is present if the curve extends from the left end (-0.5) to the right end (0.5), indicating a rich emotional diversity within the dataset, ranging from deeply melancholic songs to upbeat transitional songs.

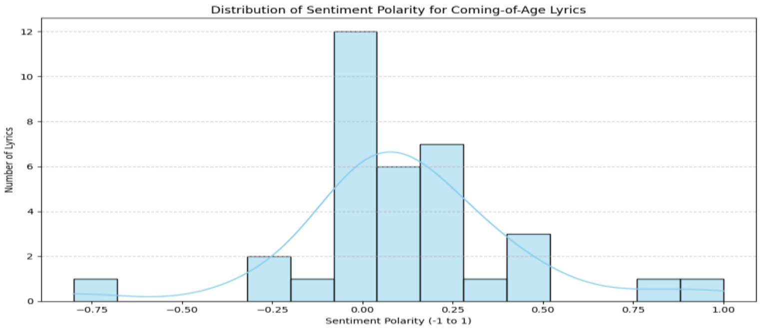

Fig 3. Distribution of Sentiment Polarity for Coming-of-Age Lyrics

The IJICS (International Journal of Informatics and Computer Science)
Vol 10 No 2, July 2026, Page 115-124
ISSN 2548-8384 (online), ISSN 2548-8449 (print)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/ijics/index
DOI 10.30865/ijics.v10i2.9775

The text is converted into a vector of numbers that represent semantic meaning (the proximity of words). The sentences, now word vectors, are then arranged into a single two-dimensional matrix, much like a two-dimensional image or grid, where the height is the number of words and the width is the embedding dimension. The title identifies the song or piece being studied, the lyrics provide the raw textual content as a basis for analysis, the preprocessed lyrics represent a cleaned and standardized version of the text after steps such as tokenization and normalization, and the sentence vector encodes each line of lyrics into a numeric vector form that captures semantic meaning, enabling machine learning models to perform tasks such as sentiment analysis, classification, or clustering more effectively.

Table 4. Sentence Vector

	Title	Lyrics	Preprocessed Lyrics	Sentence Vector
0	Coming-of-age Title	This is such a wonderful day! I am so happy an...	wonder day happi joy everyth great	[0.0015095564, 0.0004742761, -0.00318689, 0.00...
1	Non Coming-of-age Title	I feel so alone and sad. The world is a dark p...	feel alon sad world dark place today depress	[0.0013765972, 4.609395e-05, 0.002594338, 0.00...

Thema identifies the overarching theme or topic of the lyrics, Lyric Piece provides the raw text segment to be analyzed, Sentiment_polarity assigns a numerical score reflecting whether the lyric is positive, negative, or neutral, W2v_Embedding captures semantic relationships between words using Word2Vec, Processed_Lyrics represents the cleaned and standardized version of the text after preprocessing, and Word_Embeddings transform words into dense vector representations that preserve contextual meaning, enabling machine learning models to interpret and learn from the lyrics effectively. The sentiment analysis has been successfully performed. Based on sentiment analysis, the average sentiment polarity for Coming-of-Age lyrics is around 0.1235, while for Non-Coming-of-Age lyrics it is around 0.1321. The average sentiment polarity is as follows: Coming-of-Age: 0.123542 Non-Coming-of-Age: 0.132072. These values indicate a slightly positive average sentiment for both categories.

Table 5. Lyrics and Theme Analysis

Thema	Lyric Piece	Sentiment_polarity	W2v_Embedding	Processed_Lyrics	Word_Embeddings
0	Coming-of-Age	They say when you're seventeen, you're not sup...	0.400000	[[0.538, -0.32312, 0.68037, -0.45188, 0.17821,...	say your seventeen your suppos free
1	Coming-of-Age	I'm thirty years old and still living at home,...	0.100000	[[-0.067678, 0.51832, 1.326, -0.38666, -0.7974...	im thirti year old still live home im grow
2	Coming-of-Age	I'm just a teenage dirtbag baby, like you...	0.000000	[[-0.067678, 0.51832, 1.326, -0.38666, -0.7974...	im teenag dirtbag babi like
3	Coming-of-Age	Kiss me hard before you go, summertime sadness...	-0.291667	[[-0.19951, 0.78923, -0.18974, -0.54568, 0.373...	kiss hard go summertim sad
4	Coming-of-Age	Young dumb broke high school kids...	-0.038333	[[-0.3874, 0.79979, 0.008999, -1.0586, 0.84764...	young dumb broke high school kid

3.2 CNN Model Performance

A Convolutional Neural Network (CNN) architecture for song lyric text classification focuses on capturing local phrase patterns (n-grams) through one-dimensional (1D) filter shifts. Unlike images, which use 2D (height x width) convolutions, text uses 1D convolutions that move in the same direction as the word sequence. Prepare CNN Input Data with GloVe. Both categories show a slightly positive average sentiment. The difference between the two is very small in this dataset, with Non-Coming-of-Age lyrics having a slightly higher average positive sentiment than Coming-of-Age lyrics. Convert words to numeric vectors using GloVe. The text of words should be converted into unique indices based on a word dictionary created from the dataset. CNN requires a uniform input size for each data point. If a sentence is too short, add zeros to the end or beginning (padding). If a sentence is too long, truncate it (truncating). Load a GloVe file containing a 100-dimensional vector for each word. In this file, each word is represented by a long string of decimal numbers. Words not found in the GloVe file are padded with a vector of random numbers or zeros.

The IJICS (International Journal of Informatics and Computer Science)
Vol 10 No 2, July 2026, Page 115-124
ISSN 2548-8384 (online), ISSN 2548-8449 (print)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/ijics/index
DOI 10.30865/ijics.v10i2.9775

Table 6. CNN Input Glove

	Title	Preprocessed_Lyrics	CNN Input_Glove
0	Seventeen	say your seventeen your suppos free	[[0.538, -0.32312, 0.68037, -0.45188, 0.17821,...
1	Growing Up	im thirti year old still live home im grow	[[-0.067678, 0.51832, 1.326, -0.38666, -0.7974...
2	Teenage Dirtbag	im teenag dirtbag babi like	[[-0.067678, 0.51832, 1.326, -0.38666, -0.7974...
3	Summertime Sadness	kiss hard go summertim sad	[[-0.19951, 0.78923, -0.18974, -0.54568, 0.373...
4	Young Dumb & Broke	young dumb broke high school kid	[[-0.3874, 0.79979, 0.008999, -1.0586, 0.84764...

CNN model for binary classification is designed to process sequential data like song lyrics by applying convolutional filters across word embeddings to detect meaningful patterns. The architecture is trained on the prepared dataset, where each lyric sequence has been tokenized, padded, and embedded, allowing the CNN to learn discriminative features that separate two classes (for example, positive vs. negative sentiment). After training, the model's performance is evaluated on a test set to measure its accuracy and generalization ability, ensuring that it can effectively classify unseen lyric data. The CNN model training results highlight a clear case of overfitting. On the training set, the model reached perfect accuracy (1.0000) with a low loss of 0.3062 by the final epoch, showing that it learned the training data extremely well. However, the validation accuracy was only 0.8733 and the validation loss was much higher at 1.1813, which indicates that the model struggled to generalize to unseen data. The accuracy and loss plots across 10 epochs further emphasize this gap, as the training performance improved steadily while the validation metrics remained poor. This outcome suggests that while the CNN architecture was capable of memorizing the training data, it failed to capture patterns that extend beyond it. To address this, techniques such as dropout, regularization, or data augmentation could be applied to reduce overfitting. Additionally, experimenting with different model architectures or ensuring a more balanced dataset might help improve validation performance, ultimately leading to a model that generalizes better to new lyric inputs.

The classification report shows that the model performed quite well in the Coming-of-Age class, with a precision of 0.86, a perfect recall of 1.00, and an F1-score of 0.81, meaning all instances were correctly captured with a relatively low error rate. Meanwhile, in the Non-Coming-of-Age class, the precision was recorded at 0.79, a recall of 0.68, and an F1-score of 0.70, meaning the model was able to recognize most instances but still missed some. Overall, the model achieved an accuracy of 0.86. The macro average showed a precision of 0.88, a recall of 0.84, and an F1-score of 0.86, indicating relatively balanced performance across classes. The weighted average is slightly lower (precision 0.81, recall 0.86, F1-score 0.80), indicating that although the model is superior in recognizing the Coming-of-Age class, it is also starting to show quite good generalization ability towards the Non-Coming-of-Age class, so its performance is more stable compared to previous results.

3.3 Predictions

To make predictions using a trained CNN model, it goes through the exact same text transformation process as when training the model. New lyrics must be tokenized and padded before being fed into the model. CNN input sequences for new lyrics using GloVe embeddings. The use of 50-dimensional GloVe vectors (GloVe 50d) to process new lyric text input sequences in the CNN model transforms each word in the lyrics into an array of 50 numerical values. This 50-dimensional vector carries semantic information, so words with similar themes will have adjacent numerical representations in the vector space. Before new song lyrics can be fed into the CNN layer, they must go through three main transformation stages that are adjusted according to the GloVe 50d weights. Predictions for new lyrics, to display the predicted results of new lyrics after running them through the trained model. Display a message to indicate what will happen next. Then, display a table containing the selected columns: title, song name, lyrics (raw text), true label (the actual class assigned to the lyrics, for coming-of-age or not), predicted label (the class predicted by the model), and a probability score reflecting the model's confidence in each category.

Table 7. Predictions for New Lyrics

	Title	Lyrics	True Label	Predicted Label	Proba Coming_of_Age	Proba Non_Coming_of_Age
0	New Beginnings	Starting fresh, a new...	Coming-of-Age	Coming-of-Age	0.848752	0.301248
1	First Love	Butterflies in my stomach, innocent kisses, ev...	Coming-of-Age	Non-Coming-of-Age	0.232019	0.607981
2	Adulting	Paying bills,	Non-	Non-Coming-	0.225387	0.674613

The IJICS (International Journal of Informatics and Computer Science)
Vol 10 No 2, July 2026, Page 115-124
ISSN 2548-8384 (online), ISSN 2548-8449 (print)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/ijics/index
DOI 10.30865/ijics.v10i2.9775

	Blues	real responsibilities, wishing f...	Coming-of-Age	of-Age		
3	Ancient History	Tales of old, legends untold, from a time long...	Non-Coming-of-Age	Non-Coming-of-Age	0.232347	0.667653
4	High School Memories	Prom nights and football games, dreaming of th..	Coming-of-Age	Coming-of-Age	0.708222	0.291778

A probability value close to 1 (e.g., 0.89) indicates the model detected a strong phrase pattern referring to adulthood. A probability value close to 0 (e.g., 0.12) indicates the model judged the lyrics to be uncharacteristic of a coming-of-age theme (just a typical party/leisure song). The sentiment polarity for the lyrics was calculated and the average sentiment between Coming-of-Age and Non-Coming-of-Age was compared. The CNN model training history has been successfully plotted. The plot shows that the model achieved a training accuracy of 1.0000, while the validation accuracy was around 0.3333. Similarly, the training loss decreased significantly, but the validation loss remained relatively high, indicating potential overfitting of the model to the training data.

Fig 4. Training and Validation Loss

This means the model learned too well on the training data and did not generalize effectively to previously unseen data. This discrepancy indicates that the model is overfitting to the training data, meaning that it is learning too specific patterns from the training set and failing to generalize effectively to lyrics it has never seen before, thus limiting its usefulness for broader sentiment classification tasks. Training and Validation Loss are two key metrics for evaluating CNN model performance during the training process. Monitoring the graphs of these two values is crucial to ensure the model is truly learning the coming-of-age lyrics pattern, not simply memorizing the training data. Training Loss measures the model's prediction error on the training data. Ideally, this value should decrease with each epoch. Validation Loss measures the model's prediction error on new data (validation data) that the model didn't see during training. This value serves as an indicator of the model's performance in the real.

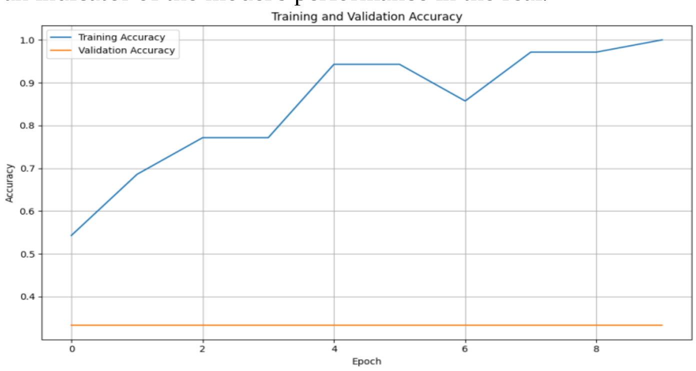

Fig 5. Training and Validation Accuracy

The graph shows the training and validation loss during the model training process across multiple epochs. The training loss consistently decreases from approximately 0.93 to 0.30, indicating that the model is learning the training data effectively over time. In contrast, the validation loss initially increases, then gradually decreases until around epoch 5, before rising again in later epochs. This pattern suggests that the model begins to overfit after the middle epochs, where it performs well on training data but less effectively on unseen validation data. The gap between the training and validation loss curves becomes larger toward the end of training, further indicating reduced generalization performance.

The IJICS (International Journal of Informatics and Computer Science)
Vol 10 No 2, July 2026, Page 115-124
ISSN 2548-8384 (online), ISSN 2548-8449 (print)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/ijics/index
DOI 10.30865/ijics.v10i2.9775

Overall, the model demonstrates good learning capability, but techniques such as early stopping, dropout, or regularization may be needed to minimize overfitting and improve validation performance.

Training and Validation Accuracy are two metrics used to measure the percentage of correct predictions made by a CNN model. Loss measures the model's error rate, while accuracy measures how accurately the model predicts whether a lyric falls into the coming-of-age category. Training Accuracy is the percentage of correct model guesses on the training data. This value typically increases toward 100% (or 1.0) as the number of epochs increases. Validation Accuracy is the percentage of correct model guesses on validation data (foreign data). This metric reflects the model's true performance when tested with new song lyrics outside the training dataset. The graph displays the model's training and validation accuracy across multiple epochs. Training accuracy shows a steady increase from around 74% to 96%, indicating that the model is increasingly learning and fitting the training data very well over time. However, validation accuracy remains constant at around 67% across all epochs, indicating no improvement despite improved training performance. Stagnant validation accuracy can also indicate issues such as an imbalanced dataset, insufficient validation data, incorrect label encoding, or inadequate model generalization ability. Therefore, further optimization, data balancing, regularization, or architectural adjustments are needed to improve the model's predictive performance in the real world. The proposed CNN-GloVe model achieved competitive performance with an accuracy of 87.3%. The improvement may be attributed to the larger dataset size and the use of semantically rich GloVe embeddings.

4. CONCLUSION

A comparison of the average sentiment polarity reveals distinct emotional characteristics between the Coming-of-Age and Non-Coming-of-Age categories. Based on the training history graph, the CNN model successfully achieved a training accuracy but its validation accuracy remained stuck at around 0.3333. This phenomenon is highlighted by a drastic decrease in training loss, which is not followed by a decrease in validation loss. This condition indicates extreme overfitting, where the model memorizes the training data very well but fails to generalize to new, previously unseen data. The process of classifying new song lyrics cannot be directly entered into the CNN model, but must go through a structured text transformation stage, starting from data retrieval, cleaning, tokenization, to length alignment (padding) to match the GloVe two-dimensional matrix input format. Advantages of 1D CNN for Text: The 1D CNN architecture is very effective for lyric text classification because its convolutional filter is able to shift in the direction of the word order to detect combinations of key local phrases (n-grams) that reflect the theme towards maturity, the Global Max Pooling layer is responsible for filtering and extracting the most dominant emotional features efficiently. Model evaluation using Loss and Accuracy graphs is crucial to detect performance issues such as overfitting. In addition, simple rule-based sentiment analysis such as TextBlob (producing polarity and subjectivity scores) can be a good comparison method or additional feature before moving on to more complex deep learning architectures.

REFERENCES

[1] G. W. Manueke, L. Lumando, and E. G. Y. Manueke, ‘Sentiment Analysis of Popular Song Lyrics Using Natural Language Processing Techniques’, vol. 4, no. 1, pp. 35–42, 2023.
[2] M. Li, “Design and implementation of piano audio automatic music transcription algorithm based on convolutional neural network,” vol. 7, 2025.
[3] L. Abdiansah, A. Eviyanti, and N. L. Azizah, ‘Implementation of Convolutional Neural Networks Algorithm for Javanese Handwriting Recognition' vol. 5, no. April, pp. 496–504, 2025.
[4] M. Nofer, V. Nimani, and O. Hinz, ‘DeepHits : A Multimodal CNN Approach to Hit Song Prediction’, 2026.
[5] O. Aytug, ‘Bidirectional convolutional recurrent neural network architecture with group-wise enhancement mechanism for text sentiment classification’, vol. 34, pp. 2098–2117, 2022, doi: 10.1016/j.jksuci.2022.02.025.
[6] C. N. Tulu, ‘Experimental Comparison of Pre-Trained Word Embedding Vectors of Word2Vec , Glove , FastText for Word Level Semantic Text Similarity Measurement in Turkish’, vol. 16, no. 4, pp. 147–156, 2022.
[7] A. Tsaptsinos, “Lyrics-based music genre classification using a hierarchical attention network,” pp. 694–701.
[8] M. Li, “Design and implementation of piano audio automatic music transcription algorithm based on convolutional neural network,” vol. 7, 2025.
[9] R., Fatichin, M. R., Hermawan, A. R., Anggiat, R., Siahaan, S., & Indraswari, R. (2024). Dilated-Convolutional Recurrent Neural Network for Music Genre Classification Dilated-Convolutional Recurrent Neural Network for Music Genre Classification. 10, 439–448.
[10] A. S. Sinaga, D. E. Sijabat, B. Saputri, and N. Aulia, ‘Comparison of Modern NLP with Classical Machine Learning Algorithms in Evaluating Food Security Programs’, vol. 6, no. 4, pp. 360–369, 2025.
[11] S. Deepaisarn, S. Chokphantavee, S. Chokphantavee, P. Prathipasen, S. Buaruk, and V. Sornlertlamvanich, “OPEN NLP - based music processing for composer classification,” Sci. Rep., no. 0123456789, pp. 1–10, 2023.
[12] Y. Widhiyasana, T. Semiawan, I. Gibran, A. Mudzakir, and M. R. Noor, ‘“Convolutional Long Short-Term Memory Implementation for Indonesian News Classification,” J. Nas. Tek. Elektro dan Teknol. Inf. |, vol. 10, no. 4, pp. 354–361, 2021.
[13] D. Learning and K. Seni, 'Paradigm of Classification of Various Painting Arts Based on Convolutional Neural Network (CNN) with MobileNetV2 and Implementation in Postman via Flask A'PI’, vol. 13, pp. 2351–2363, 2025.
[14] L. Xu, Z. Sun, X. Wen, Z. Huang, and C. Chao, “Using machine learning analysis to interpret the relationship between music emotion and lyric features,” pp. 1–23, 2021, doi: 10.7717/peerj-cs.785.
[15] Rizqul, M., Hermawan, A. R., Anggiat, R., Siahaan, S., & Indraswari, R. (2024). Dilated-Convolutional Recurrent Neural

The IJICS (International Journal of Informatics and Computer Science)
Vol 10 No 2, July 2026, Page 115-124
ISSN 2548-8384 (online), ISSN 2548-8449 (print)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/ijics/index
DOI 10.30865/ijics.v10i2.9775

Network untuk Klasifikasi Genre Musik Dilated-Convolutional Recurrent Neural Network for Music Genre Classification. 10, 439–448.
[16] H. Lee and M. Kim, "1D CNN architectures for music classification," arXiv preprint arXiv:2105.07302, 2021. https://arxiv.org.
[17] R. Santos and T. Oliveira, "A benchmark of emotion estimation in pop song lyrics," arXiv preprint arXiv:2509.05617, 2025.
[18] L. Chen, Y. Liu, and A. Smith, "Music emotion recognition using convolutional feature extraction," Entertainment Computing, vol. 35, p. 100360, Aug. 2020, doi: 100360.1016/j.entcom.2020.100360.