

**Klasifikasi Emosi pada Lirik Lagu menggunakan Metode
Bidirectional LSTM dengan Pembobotan GloVe Word
Representation**

Tugas Akhir

diajukan untuk memenuhi salah satu syarat

memperoleh gelar sarjana

dari Program Studi Informatika

Fakultas Informatika

Universitas Telkom

1301162765

Jiddy Abdillah



Program Studi Sarjana Informatika

Fakultas Informatika

Universitas Telkom

Bandung

2020

LEMBAR PENGESAHAN

**Klasifikasi Emosi pada Lirik Lagu menggunakan Metode Bidirectional LSTM dengan
Pembobotan GloVe Word Representation**

**Emotion Classification of Song Lyrics using Bidirectional LSTM Method with GloVe
Word Representation Weighting**

NIM :1301162765

Jiddy Abdillah

Tugas akhir ini telah diterima dan disahkan untuk memenuhi sebagian syarat memperoleh
gelar pada Program Studi Sarjana Informatika

Fakultas Informatika

Universitas Telkom

Bandung, <Tanggal/Bulan/Tahun>

Menyetujui

Pembimbing I,



Ibnu Asror, S.T., M.T.

NIP: 06840031

Pembimbing II,

Yanuar Firdaus Arie Wibowo, S.T., M.T.

NIP: 05790003

Ketua Program Studi
Sarjana Informatika,

Niken Dwi Wahyu Cahyani, S.T., M.Kom., Ph.D.


NIP: 00750052

LEMBAR PERNYATAAN

Dengan ini saya, Jiddy Abdillah menyatakan sesungguhnya bahwa Tugas Akhir saya dengan judul Klasifikasi Emosi pada Lirik Lagu menggunakan Metode Bidirectional LSTM dengan Pembobotan GloVe Word Representation beserta dengan seluruh isinya adalah merupakan hasil karya sendiri, dan saya tidak melakukan penjiplakan yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan. Saya siap menanggung resiko/sanksi yang diberikan jika di kemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam buku TA atau jika ada klaim dari pihak lain terhadap keaslian karya,

Bandung, 25 Agustus 2020

Yang Menyatakan



Jiddy Abdillah

Klasifikasi Emosi pada Lirik Lagu menggunakan Metode Bidirectional LSTM dengan Pembobotan GloVe Word Representation

Jiddy Abdillah¹, Ibnu Asror², Yanuar Firdaus Arie Wibowo³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹jiddyabdillah@students.telkomuniversity.ac.id, ²iasror@telkomuniversity.ac.id,

³yanuar@telkomuniversity.ac.id

Abstrak

Perubahan pasar musik yang cukup cepat dari analog ke digital menyebabkan bertambahnya jumlah musik yang tersebar di dunia secara cepat juga karena musik lebih mudah untuk dibuat dan dijual. Banyaknya musik yang tersedia menyebabkan berubahnya cara orang menemukan musik, salah satunya yaitu berdasarkan emosi lagu. Adanya *music emotion recognition and recommendation* membantu pendengar musik menemukan lagu sesuai dengan emosi mereka. Oleh karena itu, klasifikasi emosi dibutuhkan untuk menentukan emosi sebuah lagu. Klasifikasi emosi pada sebuah lagu sebagian besar didasarkan pada ekstraksi fitur dan learning dari set data yang tersedia. Berbagai algoritma learning telah digunakan untuk mengklasifikasikan emosi lagu dan menghasilkan akurasi yang berbeda. Dalam penelitian ini, metode *deep learning Bidirectional Long-Short Term Memory* (Bi-LSTM) dengan pembobotan kata menggunakan GloVe digunakan untuk mengklasifikasikan emosi lagu menggunakan lirik dari lagu tersebut. Pada penelitian ini, hasilnya menunjukkan bahwa model Bi-LSTM dengan *dropout layer* dan *activity regularization* dapat menghasilkan akurasi sebesar 91.08%. Parameter *dropout*, *activity regularization* dan *learning rate decay* dapat mengurangi selisih *training loss* dan *validation loss* sebesar 0.15.

Kata kunci: klasifikasi emosi, BiLSTM, deep learning, GloVe

Abstract

The rapid change of the music market from analog to digital has caused a rapid increase in the amount of music that is spread throughout the world as well because music is easier to make and sell. The amount of music available has changed the way people find music, one of which is based on the emotion of the song. The existence of music emotion recognition and recommendation helps music listeners find songs in accordance with their emotions. Therefore, the classification of emotions is needed to determine the emotions of a song. The emotional classification of a song is largely based on feature extraction and learning from the available data sets. Various learning algorithms have been used to classify song emotions and produce different accuracy. In this study, the Bidirectional Long-short Term Memory (Bi-LSTM) deep learning method with weighting words using GloVe is used to classify the song's emotions using the lyrics of the song. In this study, the result shows that the Bi-LSTM model with dropout layer and activity regularization can produce an accuracy of 91.08%. Dropout, activity regularization and learning rate decay parameters can reduce the difference between training loss and validation loss by 0.15.

Keywords: emotion classification, BiLSTM, deep learning, GloVe

1. Pendahuluan

Latar Belakang

Music Emotion Recognition (MER) menggunakan fitur musik untuk mengidentifikasi emosi dalam musik. Perkembangan minat dalam mengevaluasi sistem MER dapat memberikan fitur emosi dari sumber musik yang ada. Alasan perkembangan minat ini yaitu perubahan pasar musik yang cukup cepat dari analog ke digital. Menurut survei 2016 pada pasar musik global, konsep kepemilikan seperti pendapatan fisik dan pendapatan unduhan menurun sebesar -7.6% dan -20.5%. Di sisi lain, pada tahun 2016, konsep sharing bertumbuh dengan cepat, seperti pendapatan *global digital share*, pendapatan digital dan pendapatan streaming tumbuh sebesar 50%, 17.7% dan 60.4% dan tren musik berubah menjadi era elektronik [1]. Salah satu fitur musik yang digunakan untuk mengidentifikasi emosi dalam musik adalah lirik. Lirik mempunyai nilai semantik yang kaya dan ekspresif dan memiliki dampak mendalam pada persepsi manusia tentang musik [2]. Mengklasifikasikan emosi musik menggunakan fitur lirik dapat dilakukan dengan metode klasifikasi teks. Algoritma seperti algoritma pembelajaran mesin sering digunakan untuk klasifikasi teks.

Salah satu model yang dapat digunakan untuk mengklasifikasi emosi yaitu model *deep learning*. Penelitian yang menggunakan model *deep learning* untuk menentukan emosi lagu biasanya menggunakan fitur lirik dan audio dari lagu tersebut. Pada penelitian ini, penulis mencoba menentukan emosi lagu berdasarkan fitur lirik saja. Model *deep learning* yang digunakan pada tugas akhir ini yaitu *Bidirectional Long-Short Term Memory* (Bi-LSTM).

Alasan penulis menggunakan metode ini yaitu model Bi-LSTM mempertimbangkan konteks dari informasi teks dan bisa mendapatkan representasi teks lebih baik [3]. Hasil akurasi dari model BiLSTM untuk permasalahan klasifikasi emosi berdasarkan lirik lagu pun mempunyai hasil yang cukup baik, yaitu 69.01% [4]. Untuk pembobotan kata, penulis menggunakan *pre-trained word embeddings* GloVe yang mempunyai akurasi lebih baik disbanding model *word embeddings* lain seperti CBOW dan *skip-grams*. Secara keseluruhan, GloVe mengungguli model lain dalam hal analogi kata, kemiripan kata dan tugas *named entity recognition* [5].

Topik dan Batasannya

Berdasarkan latar belakang, rumusan masalah pada tugas akhir ini yaitu bagaimana performa model dan perbandingan performa parameter klasifikasi emosi berdasarkan lirik lagu menggunakan model BiLSTM dengan pembobotan kata *GloVe Word Representation*.

Adapun batasan masalah untuk penelitian ini agar permasalahan yang diambil dapat disesuaikan dengan kebutuhan dan kemampuan penulis, yaitu:

- a) Jumlah kelas yang digunakan dalam klasifikasi emosi sebanyak empat kelas, yaitu *Happy* (senang), *Relaxed* (santai), *Sad* (sedih) dan *Angry* (marah).
- b) Lirik lagu yang digunakan yaitu menggunakan bahasa inggris.
- c) Daftar lagu dan emosi dari dataset *MoodyLyrics* [6] yang berjumlah 2189 baris dan lirik lagu diambil dari *Genius* [7] dan *SongLyrics* [8].
- d) Menggunakan model BiLSTM dan pembobotan kata *GloVe Word Representation*.
- e) *Hyperparameter* yang digunakan yaitu *epoch* dan *learning rate*.
- f) Parameter yang akan dibandingkan performanya yaitu *learning rate decay*, *activity regularization* dan *dropout*.

Tujuan

Tujuan yang ingin dicapai berdasarkan rumusan masalah yang ada, yaitu menganalisis performa model BiLSTM dan membandingkan performa parameter pada kasus klasifikasi emosi berdasarkan lirik lagu.

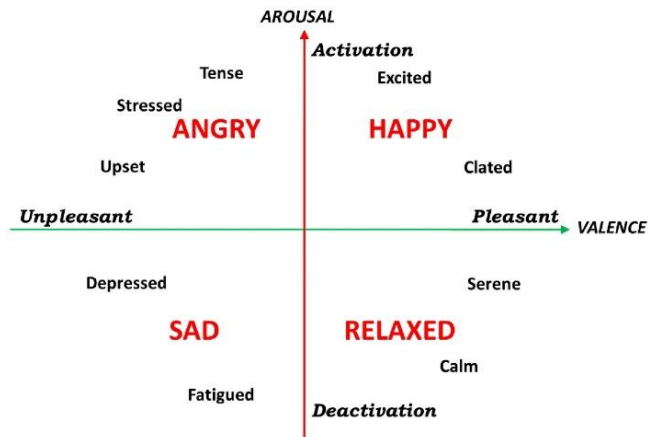
Organisasi Tulisan

Penulisan tugas akhir ini tersusun dari beberapa bagian, yaitu Bagian 1 yang menjelaskan pendahuluan dan latar belakang penelitian ini, Bagian 2 menjelaskan studi terkait dari penelitian yang dilakukan, Bagian 3 menjelaskan sistem yang dibangun untuk melaksanakan penelitian, Bagian 4 yang menjelaskan evaluasi dan analisis dari sistem yang dibangun dan terakhir Bagian 5 menjelaskan kesimpulan dan saran hasil penelitian.

2. Studi Terkait

Pada penelitian klasifikasi emosi berdasarkan lirik lagu sebelumnya, model yang digunakan untuk menentukan kelas emosi yaitu model kategorikal dan dimensi. Teks lirik berbeda dari teks biasa dalam penggunaan gaya bahasa seperti sajak, bentuk puisi, dan bahasa kiasan. Lirik lagu membantu memusatkan perhatian pendengar pada emosi tertentu. Sejak tahun 1930-an, psikolog telah menafsirkan nilai afektif dari kata-kata, berdasarkan survei empiris dan penilaian ahli. Skala pengukuran emosi dibuat untuk mengukur laporan verbal dari keadaan psikologis pendengar berdasarkan banyak emosi dan dimensinya (misal *intensity*, *valence* dan *dominance*) [9]. Salah satu model dimensi yang ada untuk menentukan sebuah emosi yaitu *circumplex model* pada Gambar 1 yang dikembangkan oleh James Russel. Model ini menunjukkan bahwa emosi didistribusikan dalam ruang dua dimensi, yaitu *valence* dan *arousal* [10].

Penelitian terdahulu terkait dengan klasifikasi emosi berdasarkan lirik lagu bahasa Inggris telah beberapa kali dilakukan [4, 6, 11]. An et al. melakukan penelitian menggunakan *Naïve Bayes Classifiers* dengan empat kelas emosi dan mendapat akurasi sebesar 68% [12]. Pendekatan menggunakan *unsupervised learning* juga telah dilakukan sebelumnya, seperti klasifikasi menggunakan leksikon dan *clustering*. Erion Çano dan Maurizio Morisio menggunakan ANEW dan *WordNet Lexicon* untuk representasi kata dan pengelompokan untuk memasukkan nilai-nilai emosi dari semua kalimat dan emosi untuk seluruh lagu. Keakuratan leksikon dan metode dibandingkan dengan dataset lirik yang dianotasi oleh tag pengguna dan subjek manusia adalah 74,25%, dibandingkan dengan dataset yang dibuat oleh pendengar dan label beranotasi menggunakan tag pengguna [6]. Kelemahan dari metode ini adalah leksikon memiliki kata yang terbatas dan ada beberapa kata dalam lirik lagu yang tidak digunakan dan ditinggalkan dalam proses pelatihan.

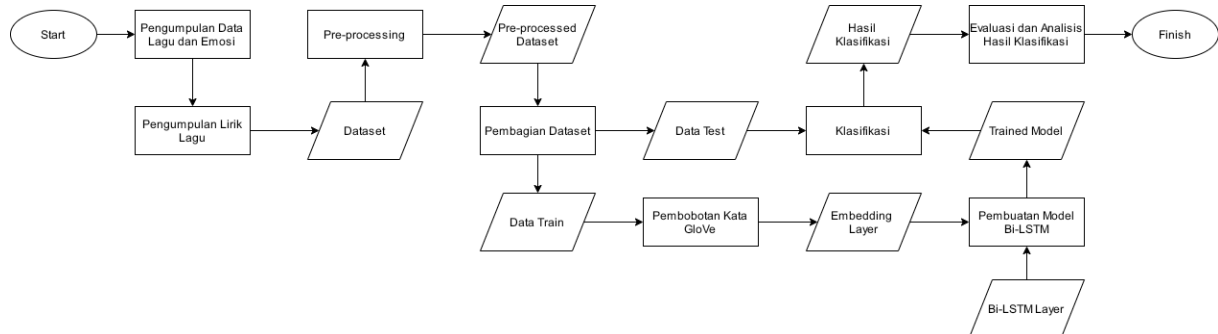


Gambar 1. Circumplex Model

Pada penelitian yang dilakukan oleh Revanth Akella dan Teng-Sheng Moh [4], beberapa model *deep learning* digunakan, seperti *Convolutional Neural Network* (CNN), BiLSTM dan *Convolutional Recurrent Neural Network* (CRNN). Pada penelitian tersebut, model CNN menghasilkan akurasi 71%, BiLSTM 69.01% dan CRNN 67.04%. Meskipun klasifikasi emosi berdasarkan lirik menggunakan metode BiLSTM sudah pernah dilakukan, tetapi model yang dibuat Akella et al. tidak menggunakan *regularization layer* seperti *dropout layer* dan *activity regularization layer*. Pada model *deep learning*, *hyperparameter* seperti *learning rate* dan *decay scheduling* juga mempengaruhi performa dari model. Maka dari itu, penulis akan mencoba mengimplementasikan *regularization layer* dan *hyperparameter tuning* ke model BiLSTM yang dibuat.

3. Sistem yang Dibangun

Sistem klasifikasi emosi berdasarkan lirik lagu yang dibuat pada penelitian ini dibangun untuk membuat model yang dapat mempelajari struktur pattern lirik lagu secara menyeluruh dari depan ke belakang berdasarkan konteks dari lirik lagu tersebut. Gambaran sistem yang dibangun dapat dilihat dari *flowchart* Gambar 2.



Gambar 2. Flowchart Sistem

3.1 Pengumpulan Data

Data lagu dan emosi yang digunakan ini diambil dari dataset MoodyLyrics. Kolom pada dataset tersebut terdiri dari empat kolom, yaitu index lagu, artist, judul lagu dan emosi. Pada dataset tersebut belum terdapat lirik dari lagu. Maka dari itu, kami menggunakan data artist dan judul lagu untuk mencari lirik pada website Genius dan SongLyrics dengan cara *web crawling*. Lirik yang sudah diambil digabungkan ke dataset MoodyLyrics. Contoh dataset yang dihasilkan dapat dilihat di Tabel 1.

Tabel 1. Dataset

Artist	Title	Lyrics	Emotion
Tyler, the Creator	A BOY IS A GUN*	No, don't shoot me down (yeah) No, don't shoot me down (okay) No, don't shoot me down ...	Angry

Dataset yang telah dikumpulkan dan diproses selanjutnya dibagi menjadi dua bagian, *data train* dan *data test*. Dataset berjumlah 2189 data kemudian dipecah menjadi 80% *data train* dan 20% *data test*. *Data train* digunakan untuk input data model yang dibuat dan *data test* digunakan untuk input klasifikasi dan analisis akurasi dari model yang dibuat.

3.2 Preprocessing

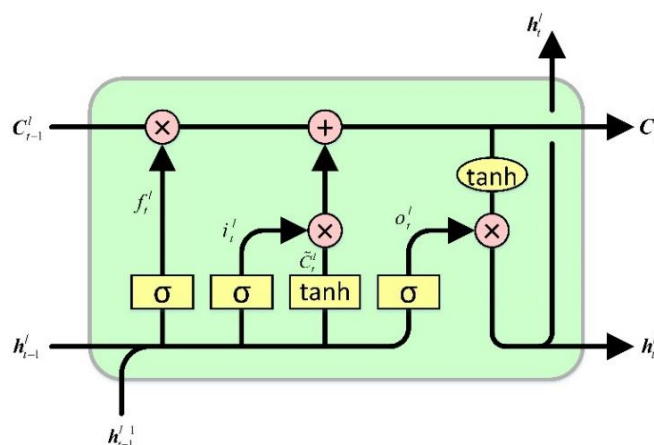
Preprocessing adalah proses perubahan bentuk data yang belum terstruktur menjadi data yang terstruktur. Tahap preprocessing mengubah data tekstual menjadi data yang siap dijadikan model text mining [13]. Lirik dari dataset yang sudah dibuat dilakukan *preprocessing* untuk memperbaiki struktur dan menghindari data yang tidak sempurna. Ada beberapa tahapan yang biasa dilakukan pada tahap ini, yaitu *tokenization*, *stop-word removal*, *lowercase conversion*, dan *lemmatization* [14]. Penjelasan *preprocessing* pada lirik lagu dapat dilihat pada Tabel 2.

Tabel 2. Tahap *Preprocessing*

Tahapan	Proses	Input	Output
<i>Case Folding</i>	Setiap kata pada lirik yang ada di data training akan diubah menjadi lowercase.	No, don't shoot me down (yeah) No, don't shoot me down (okay) No, don't shoot me down ...	no, don't shoot me down (yeah) no, don't shoot me down (okay) no, don't shoot me down ...
<i>Stopword Removal</i>	Kata yang tidak memiliki ketergantungan suatu topik atau tanda baca akan dihapus karena tidak relevan pada perhitungan model klasifikasi.	leave, stay right here, yeah, i want you right near (you started with a mere hello)	leave stay right here yeah right near started mere hello
<i>Tokenization</i>	Proses tokenization dilakukan untuk memisahkan sebuah kalimat menjadi kata-kata atau disebut token.	leave stay right here yeah right near started mere hello	['leave', 'stay', 'right', 'here', 'yeah', 'near', 'started', 'mere', 'hello']
<i>Lemmatization</i>	Setiap kata harus diubah ke akar kata tersebut atau menjadi kata dasar. Imbuhan yang ada pada kata dihapus sehingga kata tersebut berubah menjadi kata dasar.	['leave', 'stay', 'right', 'here', 'yeah', 'near', 'started', 'mere', 'hello']	['leave', 'stay', 'right', 'here', 'yeah', 'near', 'start', 'mere', 'hello']

3.3 Bidirectional Long-Short Term Memory

Long-Short Term Memory (LSTM) adalah arsitektur *recurrent neural network* (RNN) yang dirancang untuk mengatasi masalah fluktuasi gradien pada RNN konvensional [15]. LSTM mempunyai tiga jenis gerbang (*gate*) yaitu *forget gate*, *input gate* dan *output gate*. Struktur dari model LSTM dapat dilihat di Gambar 3



Gambar 3. Struktur unit LSTM

Elemen dalam input sekuens dilacak oleh *cell memory* yang berfungsi untuk melacak dependensi antar elemen. *Input gate* menangani nilai baru yang masuk ke *cell state*. Untuk memilih nilai yang tersisa di *cell state*, unit LSTM menggunakan *forget gate*. Nilai yang tersisa di *cell state* akan pergi ke *output gate*, di mana perhitungan dimulai dengan fungsi aktivasi LSTM, yang sering disebut fungsi sigmoid logistik. Alur LSTM diatur dengan persamaan sebagai berikut.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

$$C_t = f_t \times C_{t-1} + i_t * \tilde{C}_t \quad (4)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \times \tanh C_t \quad (6)$$

dimana:

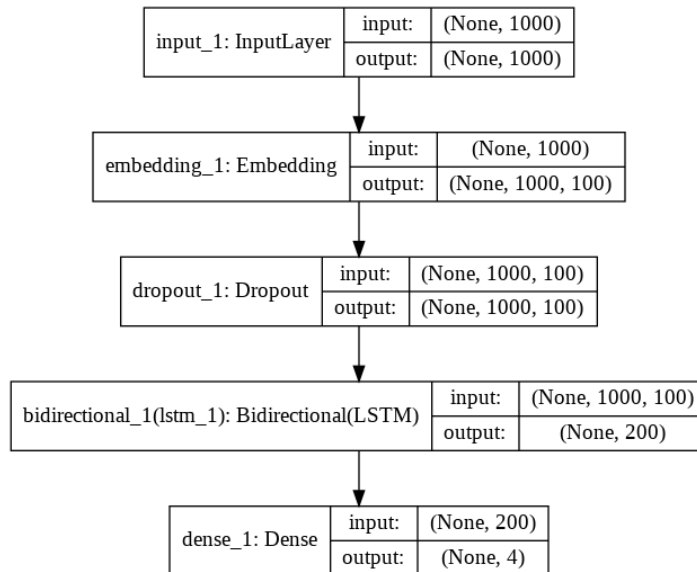
- W_i, W_f, W_C, W_o : *trained weights*.
- b_i, b_f, b_C, b_o : *trained biases*.
- σ : fungsi sigmoid.
- x_t : input pada waktu t .
- C_t : *cell state* pada waktu t .
- h_t : output pada waktu t .
- f_t : *forget gate* pada waktu t .
- i_t : *input gate* pada waktu t .
- o_t : *output gate* pada waktu t .

LSTM satu arah hanya dapat menggunakan informasi kontekstual lampau. *Bidirectional LSTM* bisa menggunakan informasi secara kontekstual, sehingga dapat membuat dua sekuens independen dari vektor output LSTM [16]. Kedua parameter LSTM *network* pada model Bi-LSTM mempunyai *word embeddings* yang sama dari kalimat dan kedua *network* tersebut independen. Output dari setiap *time step* adalah gabungan dari dua *output vector* dari kedua arah, yang diformulasikan seperti berikut dimana h_t adalah *forward* atau *backward state*.

$$h_t = \vec{h}_t \oplus \overleftarrow{h}_t \quad (7)$$

3.4 Arsitektur Model

Pada penelitian ini, model Bi-LSTM yang dibuat terdapat beberapa layer yang digunakan yang bisa dilihat pada Gambar 4, yaitu *embedding layer*, *dropout layer*, Bi-LSTM layer dan *output layer*. Semua pembuatan model Bi-LSTM dan eksperimen dikembangkan menggunakan *Keras deep learning* [17]. Konfigurasi parameter *default* yang digunakan pada penelitian ini dapat dilihat pada Tabel 3.



Gambar 4. Arsitektur Model

Untuk memproses kata pada model, diperlukan representasi vektor dari input kata sehingga kata dapat diproses. Pada penelitian ini, GloVe 6B pre-trained model word embeddings untuk memberi nilai ke input kata. Vektor kata GloVe yang penulis gunakan yaitu GloVe 100d yang mempunyai 100 dimensi dan berjumlah 400000 kata. GloVe digunakan karena lebih mudah menangkap kombinasi kata daripada pendekatan dengan n-gram [5].

Tabel 3. Parameter Model

Parameter	Value
Dimensi	100
Jumlah <i>Hidden Unit</i> BiLSTM	100
Sekuens Maksimum	1000
<i>Batch Size</i>	64
Fungsi Aktivasi	<i>Softmax</i>
<i>Optimizer</i>	<i>Adaptive Moment Estimation</i>

3.5 Evaluasi Performa

Untuk mengevaluasi model yang dibentuk, penulis menggunakan metrik *loss*, *accuracy*, *precision*, *recall* dan *f1-score* yang didapat dari tabel *confusion matrix*.

Tabel 4. Confusion Matrix

	<i>Relevant</i>	<i>Not-Relevant</i>
<i>Retrieved</i>	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>
<i>Not-Retrieved</i>	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>

Pada Tabel 4, *True Positive* (TP) merupakan nilai kategori hasil klasifikasi dan nilai kategori yang sesungguhnya sama positif. *False Positive* (FP) merupakan nilai hasil kategori hasil klasifikasi positif dan nilai kategori yang sesungguhnya negatif. *True Negative* (TN) merupakan nilai kategori hasil klasifikasi dan nilai kategori yang sesungguhnya sama negatif. Terakhir, *False Negative* (FN) merupakan nilai kategori hasil klasifikasi negatif dan nilai kategori yang sesungguhnya positif [18]. Untuk menentukan *accuracy*, *precision*, *recall* dan *f1-score*, dibutuhkan nilai TP, TN, FP dan FN. Formulasi dari metrik tersebut yaitu.

$$Accuracy = \frac{TP + TN}{(TP + FP + TN + FN)} \quad (8)$$

$$Precision = \frac{TP_i}{(TP_i + FP_i)} \quad (9)$$

$$Recall = \frac{TP_i}{(TP_i + FN_i)} \quad (10)$$

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

Untuk menghitung *loss* dari model, penulis menggunakan fungsi *categorical crossentropy*. Fungsi ini digunakan pada saat terdapat 2 kelas atau lebih. Pada penelitian ini, terdapat 4 kelas yaitu *Happy*, *Angry*, *Sad* dan *Relaxed*. Formulasi dari *categorical crossentropy* yaitu.

$$loss(p, t) = - \sum_{c=1}^c t_{o,c} \log(p_{o,c}) \quad (12)$$

dimana:

- p : vektor prediksi
- t : vektor target
- C : kelas

4. Evaluasi dan Analisis

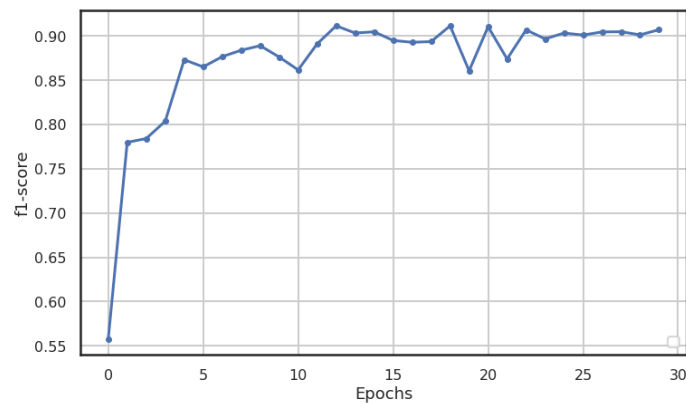
Pada bagian ini, terdapat tiga sub-bagian yang menjelaskan hasil evaluasi dan analisis dari penelitian yang telah dilakukan. Sub-bagian itu adalah konfigurasi *Hyperparameter*, Perbandingan Performansi Parameter, dan Analisis Performa Model.

4.1 Konfigurasi Hyperparameter

Parameter yang diuji performa pada penelitian ini yaitu *epoch*, *learning rate*, *dropout*, *activity regularization* dan *learning rate decay*. *Hyperparameter* yang digunakan yaitu jumlah *epoch* dan nilai *learning rate*. Nilai yang mempunyai efek terbaik ke model akan digunakan sebagai basis parameter yang nantinya akan dibandingkan dan dianalisis,

Epoch

Epoch adalah satu presentasi lengkap dari kumpulan data yang akan dipelajari model *learning*. Semakin banyak *epoch* yang digunakan, kemampuan model mengeneralisasi *pattern* data meningkat. Tetapi, jika jumlah *epoch* terlalu banyak, akan ada masalah *overfitting* pada model dan kemampuan dari model menurun.

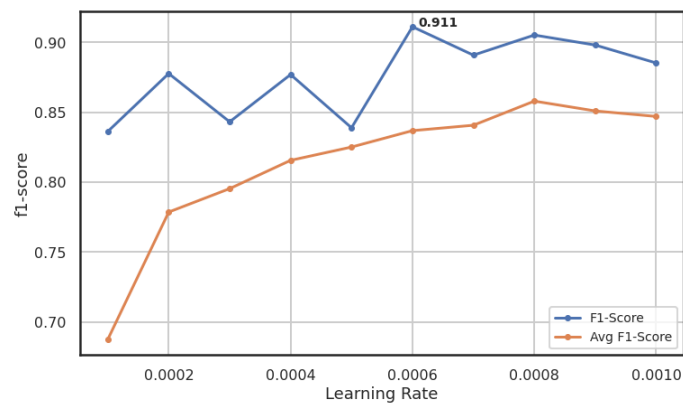


Gambar 5. Relasi antara *Epochs* dan *F1-Score*

Dari Gambar 5, dapat dilihat bahwa semakin besar *epochs*, maka performansi model dilihat dari *f1-score* semakin meningkat. Nilai *f1-score* maksimum terjadi pada saat *epoch* ke-18 dan performansi model mulai terlihat stabil pada saat *epoch* ke 22. Maka dari itu, jumlah *epochs* yang digunakan pada model yaitu 20.

Learning Rate

Pemilihan *learning rate* yang tepat penting untuk optimisasi dari *weights* dan *offset*. Jika nilai *learning rate* terlalu besar, model akan lebih mudah melebihi titik ekstrim, sehingga akan membuat model tidak stabil. Jika nilai *learning rate* terlalu kecil, maka waktu latih model akan terlalu lama.



Gambar 6. Relasi antara *Learning Rate* dan *F1-Score*

Dari Gambar 6, rata-rata *f1-score* semakin bertambah bersama dengan naiknya nilai *learning rate* dan mulai turun saat nilai *learning rate* 0.0008. *F1-score* tertinggi yaitu 0.911 pada saat *learning rate* 0.0006. Maka dari itu, nilai *learning rate* yang akan digunakan pada model yaitu 0.0006

4.2 Perbandingan Performa Parameter

Parameter *learning rate decay*, *activity regularization* dan *dropout* dibandingkan menggunakan *hyperparameter* yang telah didapat, yaitu *learning rate* = 0.0006 dan *epochs* = 20. Hasil perbandingan performa parameter dapat dilihat pada Tabel 5.

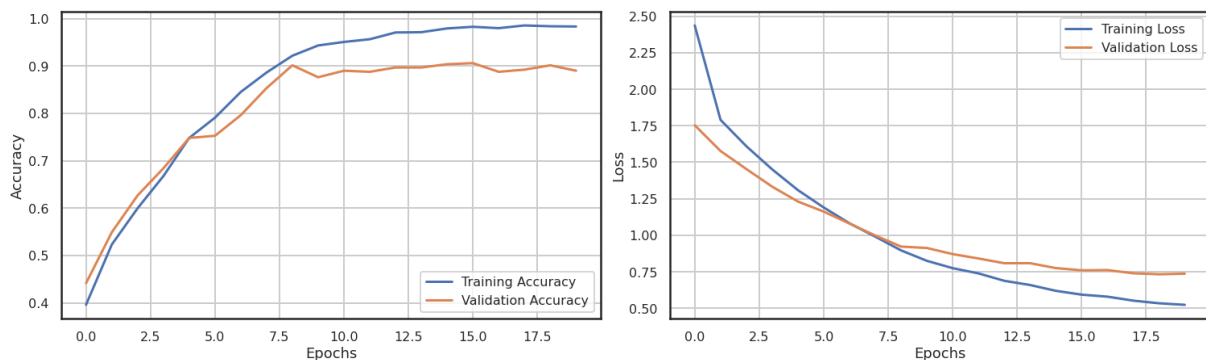
Berdasarkan Tabel 5, terbukti bahwa parameter yang digunakan dapat meningkatkan dan menurunkan performansi dari model. Kombinasi *dropout* dan *learning rate decay* mempunyai performa terburuk dengan *f1-score* 79.86%, menurun 12.08%. Diantara ketiga parameter yang digunakan, *learning rate decay* mempunyai pengaruh penurunan performa paling besar yaitu 3.36% dan *dropout* mempunyai pengaruh penurunan paling kecil yaitu 0.44%. Parameter dengan akurasi tertinggi yaitu *dropout + activity regularization* yaitu 91.08%, meningkat 0.46%. Hasil prediksi dari model dengan parameter *dropout + activity regularization* terdapat pada lampiran 1.

Tabel 5. Performa Parameter menggunakan *Learning Rate* = 0.0006 dan *Epochs* = 20 (**note. LR = Learning Rate, AR = Activity Regularization**)

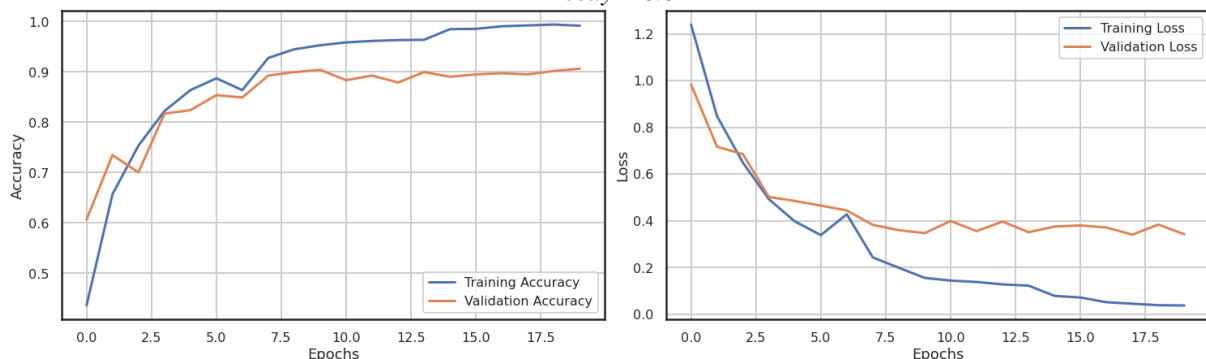
Konfigurasi Parameter	Precision	Recall	F1-Score	Accuracy
-	91.86	90.39	91.12	90.62
Dropout(0.2)	91.20	90.16	90.68	90.39
LR Decay(0.8)	89.10	87.87	88.48	88.10
AR(0.001)	91.59	87.18	89.33	89.02
Dropout(0.2) + LR Decay(0.8)	81.77	78.03	79.86	80.78
Dropout(0.2) + AR(0.001)	91.83	87.41	89.57	91.08
AR(0.001)+ LR Decay(0.8)	91.89	85.58	88.63	89.24
Dropout(0.2) + AR(0.001)+ LR Decay(0.8)	90.24	86.73	88.45	89.02

4.3 Analisis Performa Model

Performa model tanpa parameter dan performa model menggunakan parameter dibandingkan untuk mendapatkan gambaran performa keseluruhan. *Hyperparameter* yang digunakan yaitu nilai *learning rate* 0.0006 dan jumlah *epochs* 20.



Gambar 7. Performa Model dengan Parameter *Dropout* = 0.2, *Activity Regularization* = 0.001, *Learning Rate Decay* = 0.8



Gambar 8. Performa Model tanpa Parameter

Dari Gambar 7 dan Gambar 8, perbedaan signifikan dapat dilihat dari grafik *loss* masing-masing model. Model yang menggunakan parameter mempunyai perbandingan *training loss* dan *validation loss* lebih kecil yaitu 0.25, dibandingkan dengan model tanpa parameter yang mempunyai selisih 0.4, berkurang 0.15. Perbedaan dari *training loss* dan *validation loss* yang terlalu besar dapat menyebabkan model *overfitting*. Dalam kasus ini, parameter-

parameter yang digunakan dapat meminimalisir *overfitting* dari model Bi-LSTM untuk klasifikasi emosi lirik lagu, meskipun hasil akhir akurasi dari model. Model dengan parameter juga mengurangi jumlah fluktuasi nilai akurasi dan *loss* pada model.

5. Kesimpulan

Berdasarkan hasil evaluasi dan analisis yang dilakukan dalam penelitian tugas akhir ini, dapat ditarik kesimpulan bahwa:

1. Model Bi-LSTM menggunakan parameter *dropout* dan *activity regularization* menghasilkan akurasi tertinggi sebesar 91.08% pada kasus klasifikasi emosi berdasarkan lirik lagu.
2. Parameter *dropout*, *learning rate decay* dan *activity regularization* mengurangi kasus *overfitting* pada kasus klasifikasi emosi berdasarkan lirik lagu. Perbedaan *training loss* dengan *validation loss* dengan selisih 0.25, lebih kecil dibandingkan dengan model tanpa parameter dengan selisih 0.4.

Untuk penelitian berikutnya, penambahan jumlah dataset dapat berguna untuk mengurangi masalah *overfitting* pada performa model. Parameter lain seperti *attention layer*, *convolution layer* dan *pooling layer* dapat digunakan untuk meningkatkan akurasi dan performa model. Penambahan jumlah dataset dan layer akan membutuhkan waktu latih model lebih panjang.

Daftar Pustaka

- [1] DEEZER, "Global Music Report".
- [2] S. O. Ali and Z. F. Peynircioğlu, "Songs and emotions: are lyrics and melodies equal partners.," *Psychology of Music*, vol. 34, no. 4, p. 511–534, 2006.
- [3] G. Xu, Y. Meng, X. Qiu, Z. Yu and X. Wu, "Sentiment Analysis of Comment Texts Based on BiLSTM," *IEEE Access*, vol. 7, pp. 51522-51532, 2019.
- [4] R. Akella and T. S. Moh, "Mood Classification with Lyrics and ConvNets," *Proceedings - 18th IEEE International Conference on Machine Learning and Applications, ICMLA 2019*, pp. 511-514, 2019.
- [5] J. Pennington, R. Socher and C. D. Manning, "GloVe: Global Vectors for Word Representation," *Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532-1543, 2014.
- [6] E. Çano and M. Morisio, "MoodyLyrics: A sentiment annotated lyrics dataset," *ACM International Conference Proceeding Series*, vol. Part F1278, no. 11, pp. 118-124, 2017.
- [7] T. Lehman, "Genius," August 2009. [Online]. Available: <http://www.genius.com>.
- [8] "SongLyrics," 28 September 1996. [Online].
- [9] D. Yang and W. S. Lee, "Music Emotion Identification from Lyrics," *ISM 2009 - 11th IEEE International Symposium on Multimedia*, pp. 624-629, 2009.
- [10] J. Russel, "A Circumplex Model of Affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161-1178, 1980.
- [11] R. Malheiro, R. Panda, P. Gomes and R. P. Paiva, "Emotionally-relevant features for classification and regression of music lyrics," *IEEE Transactions on Affective Computing*, vol. 9, no. 2, pp. 240-254, 2018.
- [12] Y. An, S. Sun and S. Wang, "Naive Bayes Classifiers for Music Emotion Classification Based on Lyrics," *Proceedings - 16th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2017*, no. 1, pp. 635-638, 2017.
- [13] V. Srividhya and R. Anitha, "Evaluating Preprocessing Techniques in Text Categorization," *International Journal of Computer Science and Application*, pp. 49-51, 2010.
- [14] A. K. Uysal and S. Gunal, "The Impact of Preprocessing on Text Classification," *Information Processing and Management*, vol. 50, no. 1, pp. 104-112, 2014.
- [15] S. Hochreiter and J. Schmidhuber, "Long Short-term Memory," *Neural Computation*, vol. 9, pp. 1735-80, 1997.
- [16] M. Schuster and K. K. Paliwal, "Bidirectional Recurrent Neural Networks," *IEEE Transactions on Signal Processing*, vol. 45, p. 2673–2681, 1997.
- [17] F. Chollet, "Keras," 2015. [Online]. Available: <https://keras.io>.
- [18] J. Han, M. Kamber and J. Pei, *Data Mining Concepts and Techniques Third Edition*, Burlington: Morgan Kauffman Publishers, 2012.
- [19] C. W. Chen, S. P. Tseng, T. W. Kuan and J. F. Wang, "Outpatient Text Classification using Attention-based Bidirectional LSTM for Robot-assisted Servicing in Hospital," *Information (Switzerland)*, vol. 11, no. 2, p. 106, 2020.

Lampiran

Lampiran 1: Hasil Prediksi Model Bi-LSTM dengan Parameter *Dropout* dan *Activity Regularization*

Artist	Title	Actual Emotion	Predicted Emotion
Jet	Black Hearts	Angry	Angry
Rev Theory	Broken Bones	Angry	Angry
Cold	Suffocate	Sad	Angry
Nick Lowe	Without Love	Happy	Happy
Queen	Play the Game	Happy	Happy
Weekend	End Times	Sad	Sad
Exene Cervenka	Beyond You	Happy	Happy
The Autumn Defense	Step Easy	Relaxed	Relaxed
The Beatles	Free as a Bird	Relaxed	Relaxed
Irma Thomas	Same Old Blues	Relaxed	Sad
Jimi Hendrix	Who Knows	Happy	Relaxed
Vivian Girls	All the Time	Sad	Sad
The Avalanches	Flight Tonight	Angry	Angry
Melody Gardot	Love Me Like a River Does	Happy	Happy
Ruben Blades	El Padre Antonio y el Monaguillo Andres	Sad	Sad
3 Inches of Blood	Trial of Champions	Angry	Angry
Forbidden	Forsaken at the Gates	Sad	Sad
Serena Ryder	Little Bit of Red	Relaxed	Relaxed
Patty Griffin	You'll Remember	Sad	Sad
Bobby Darin	Blue Skies	Relaxed	Relaxed
Kirk Whalum	Inside	Happy	Relaxed
Fear Factory	Replica	Angry	Angry
Dolly Parton	It's Sure Gonna Hurt	Angry	Angry
Arch Enemy	I Will Live Again	Sad	Angry
The Clark Sisters	Jesus is a Love Song	Happy	Happy
Black Dub	Canaan	Happy	Happy
Joan Armatrading	A Woman in Love	Happy	Happy
The Innocence Mission	Gentle the Rain at Home	Relaxed	Relaxed