

Pencegahan Dini Terhadap Kasus Perundungan  
Melalui *Tool* Deteksi Cyberbullying di Media Digital



Disusun Oleh

TEC2021008

Fendi Irfan Amorokhman

Mukhammad Fahlevi Ali Rafsanjani

UNIVERSITAS TELKOM

*TADULAKO ESSAY COMPETITION (TIC)*

TADULAKO DATA CHALLENGE

UNIVERSITAS TADULAKO

2021



## ABSTRAK

Pada era kemajuan teknologi, literasi masyarakat yang relatif rendah pada dunia digital akan mengakibatkan banyaknya perilaku-perilaku yang justru memberikan dampak negatif seperti, perundungan secara verbal di media digital. Hal ini disebabkan maraknya pengguna media digital di Indonesia dan kurangnya pencegahan yang memadai serta aturan - aturan berlaku masih banyak yang belum diikuti oleh semua pengguna media digital. Sehingga mengakibatkan banyaknya kasus *Cyberbullying* yang dilakukan oleh individu/sekelompok orang terhadap orang lain. Juga terbatasnya kemampuan manusia dalam melakukan pengawasan pada dunia digital mengakibatkan perlunya sebuah detektor otomatis. Untuk mengatasi permasalahan tersebut, sebuah detektor akan mengklasifikasikan kalimat menjadi dua kategori yaitu kalimat perundungan dan kalimat normal. Detektor yang dibangun dengan model *Support vector machines(SVM)*, *Naive bayes*, dan *Long Short-Term Memory(LSTM)*. dengan berbagai macam kombinasi sumber data ujaran digital. Hal ini dilakukan mengingat kalimat perundungan seringkali sukar untuk dibedakan dengan kalimat normal. Hasil akhir yang didapatkan dalam penelitian ini sangat baik dengan model terbaik adalah model LSTM dengan akurasi 1.00 pada *training set* dan 0.97 pada *test set* yang menggunakan *word embedding word2vec*. Hasil klasifikasi model diharapkan dapat ditindaklanjuti oleh pihak terkait untuk dilakukannya langkah pencegahan terhadap pelaku perundungan atau *Cyberbullying* sehingga kasus perundungan di Indonesia dapat menurun.

Kata Kunci: *Cyberbullying*, Neural Network, *Machine learning*, *Word2vec*.



## 1. PENDAHULUAN

### 1.1. Latar Belakang

Perkembangan teknologi dan informasi seperti layanan jejaring sosial dan internet berkembang pesat setiap tahunnya[1]. Berdasarkan laporan dari Penyelenggara Jaringan Internet (APJI), setiap tahunnya di Indonesia terdapat peningkatan pengguna internet yang didominasi oleh remaja berusia 13-18 tahun sebesar 75,50 persen dengan mayoritas penggunaan internet per hari selama 1-3 jam[2]. Peningkatan jumlah pengguna internet khususnya tingkat remaja menyebabkan banyak permasalahan di dunia internet, misalnya *cyberbullying*. *Bullying* adalah kekerasan, ancaman, atau paksaan dari seseorang untuk *abuse* atau mengintimidasi orang lain[3]. Trauma yang mendalam yang diterima oleh korban bisa menyebabkan kerugian secara psikologis, rasa sakit, penderitaan, dan terbukti memiliki dampak traumatis pada korban[4]. Hal ini menunjukkan bahwa tingkat kesadaran tertib bermedia sosial amatlah buruk. Keadaan ini diperburuk dengan minimnya pengawasan dan pendeteksian dini oleh pihak terkait yang masih minim dikarenakan sulitnya untuk menjangkau luasnya layanan media digital.

Berdasarkan laporan Badan Pusat Statistik kelompok usia remaja menduduki peringkat tertinggi dari total penduduk Indonesia[5]. Sebanyak 60 persen remaja mengaku pernah mengalami tindakan *bullying* dan 87 persen lainnya pernah mengalami perundungan secara online[6]. Hal ini tidak boleh dibiarkan karena remaja merupakan penerus generasi bangsa Indonesia. *Survey* dari Microsoft menunjukkan peringkat kesopanan warga Indonesia saat bertransaksi elektronik menempati urutan ke-29 dari 32 negara yang sebagai responden *survey*[7]. Dari data yang telah dipaparkan menunjukkan bahwa pentingnya dan sangat mendesak untuk menjaga generasi penerus bangsa dari kasus *cyberbullying* yang marak terjadi.

Untuk memecahkan masalah di atas, dilakukanlah penelitian membuat *tool* untuk mendeteksi kalimat/*tweet* yang diunggah terklasifikasi sebagai *bullying* atau tidak. Model *Machine learning* dan model *deep learning* akan dikembangkan pada *tool* tersebut untuk dapat mengklasifikasikan apakah suatu kalimat



berpotensi menjadi kasus *bullying* atau tidak.

## 1.2. Tujuan

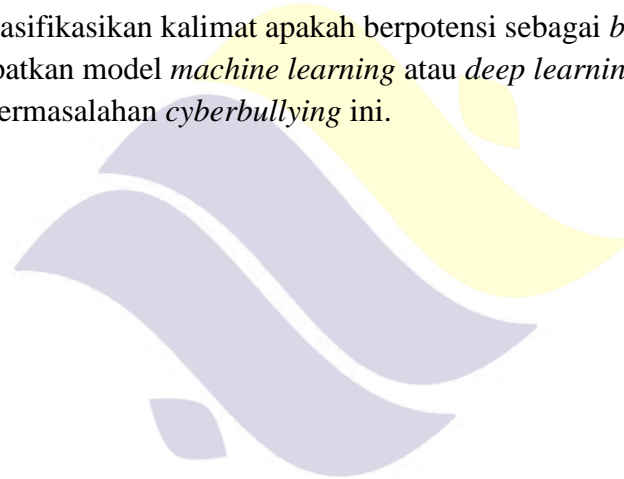
Tujuan dari penelitian ini adalah:

- Membangun tool berbasis model *machine learning* dan *deep learning* untuk pengklasifikasian kalimat dapat berkategori sebagai *bullying*.
- Membandingkan beberapa model *machine learning* dan *deep learning* yang paling baik untuk permasalahan pengklasifikasian *cyberbullying* ini.

## 1.3. Manfaat

Manfaat Dari penelitian ini adalah:

- Mengklasifikasikan kalimat apakah berpotensi sebagai *bullying* atau tidak.
- Mendapatkan model *machine learning* atau *deep learning* paling baik untuk permasalahan *cyberbullying* ini.



TADULAKO  
DATA CHALLENGE



## 2. STUDI LITERATUR

### 2.1. Cyberbullying

Perkembangan teknologi informasi memberikan dampak terhadap kehidupan sosial seperti *cyberbullying*. *Cyberbullying* merupakan suatu tindakan tidak menyenangkan yang dilakukan secara sengaja dan terus menerus melalui teks elektronik[8]. Kasus *cyberbullying* pada umumnya berjenis serangan verbal, seperti menghujat, menyerang suku, agama, ras, dan antargolongan (SARA).

### 2.2. TF-IDF

*Term Frequency-Inverse Document Frequency* (TF-IDF) adalah metode yang digunakan untuk menghitung bobot setiap kata yang telah diekstrak[9]. Dalam penelitian ini TF-IDF akan digunakan untuk model *machine learning* yaitu *naive bayes* dan *support vector machine*. Ada dua konsep penting dalam TF-IDF yaitu *term frequency*(TF) yaitu penghitungan jumlah kemunculan suatu *term* dalam kumpulan dokumen. Konsep kedua yaitu *inverse document frequency* yaitu proses penghitungan *term* yang dianggap tidak penting pada kumpulan dokumen.

### 2.3. Word2Vec

*Word2vec* adalah metode untuk merepresentasikan setiap kata dalam dokumen sebagai sebuah vektor dengan N dimensi. Cara kerja dari *Word2vec* dalam mempresentasikan kata dengan menghitung *contextual* dan *semantic similarity* dari setiap kata yang berbentuk *one-hot encoded vectors*. Hasil dari *word2vec* adalah representasi relasi setiap kata dengan kata lainnya.

### 2.1. Machine learning

*Machine learning* adalah salah satu bidang ilmu komputer yang berfokus pada pemberian kemampuan belajar tanpa harus memprogram secara eksplisit[10]. Pada penerapannya *Machine learning* dapat diterapkan pada segala bidang misalnya, memprediksi cuaca, *sentiment analysis*, *text classification* dll. Penelitian ini menggunakan *machine learning* dikarenakan kebutuhan suatu tool yang bisa belajar dengan sendiri dan dapat bekerja secara *massive* dalam pencegahan dini kasus *cyberbullying*. Model *machine learning* yang digunakan adalah *naive bayes* dan *support vector machine*.

#### 2.1.1. Naive bayes

*Naive bayes Classifier* adalah konsep probabilitas penentuan kelompok kelas dokumen[11] dari pengalaman sebelumnya dengan ciri-ciri adalah asumsi yang sangat kuat akan ketergantungan dari masing-masing kondisi. Model *naive bayes* mengasumsikan bahwa setiap kata tidak memiliki keterkaitan satu sama lain oleh karena itu disebut *naive*.

#### 2.1.2. Support vector machine.

*Support vector machine* (SVM) merupakan salah satu metode dalam *supervised learning* yang biasanya digunakan untuk klasifikasi (seperti *Support*



*Vector Classification*) dan regresi (*Support Vector Regression*)[12]. Konsep kerja dari SVM sendiri yaitu untuk menemukan hyperplane yang memisahkan himpunan data ke dalam dua kelas secara linier. Hyperplan adalah istilah yang dibuat general untuk semua dimensi. SVM berusaha untuk menemukan hyperplane yang paling optimum atau terbaik[13]. Dalam pembuatan model ada dua proses penting yaitu proses training dan proses testing. Proses

## **2.2. Neural Network**

*Neural Network* adalah model pemrograman yang mensimulasikan otak manusia dengan mengandalkan *weights* pada setiap *neuron*. Perilaku jaringan saraf buatan bergantung dari pada *weights* dan fungsi input output yang didefinisikan pada units *neuron*[7]. Hal ini lah yang dijadikan basis pembelajaran pada jaringan saraf buatan. *Neural Network* sendiri memiliki banyak bentuk (arsitektur) dengan kompleksitas yang berbeda-beda.

### **2.2.1. Recurrent Neural Network**

*Recurrent Neural Network* adalah model koneksionis yang menangkap dinamika urutan melalui siklus node jaringan [14]. RNN tidak seperti model *Neural Network*, RNN dapat merepresentasikan state agar informasi penting tidak hilang dari panjangnya sebuah kalimat.

### **2.2.2. LSTM**

Arsitektur RNN yang di desain untuk modelling kalimat lampau dan panjang kalimatnya dengan akurat dari pada RNN Tradisional[15]. Berdasarkan statement tersebut LSTM dapat memberikan performansi yang jauh lebih baik dikarenakan LSTM mempertimbangkan dependensi pada kalimat kalimat yang dijadikan inputan. Pentingnya dependensi pada kalimat adalah hal yang amat penting pada studi kasus NLP. Hal yang membedakan LSTM dengan RNN tradisional adalah LSTM memiliki *memory cellblock* sedangkan RNN tidak, hal inilah yang membuat LSTM memahami dependensi pada kalimat lebih baik .





### 3. METODE

#### 3.1. Pengumpulan data

Dalam pengumpulan data yang digunakan, dilakukan pelabelan dari beberapa dataset cuitan media sosial Twitter dan instagram. Adapun beberapa dataset tersebut, diantaranya:

1. *Cyber Bullying Indonesia Dataset* [16]
2. *Indonlu Sentiment Prosa Dataset* [17]
3. *The Dataset for Abusive Language Detection in Indonesian Social Media* [18]
4. *Indonesian Abusive & Hate Speech Twitter Text* [19]

Semua dataset diatas dikompilasi, dirapikan, dan dibersihkan menjadi satu dataset tunggal. Dataset ini dapat diakses di [Cyberbullying Dataset](#).

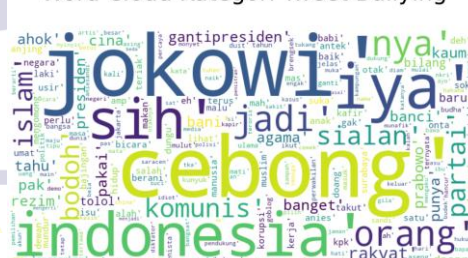
##### 3.1.1 Kategori Label Dataset

*Dataset* yang telah dikumpulkan akan diklasifikasikan menjadi dua kategori yaitu kalimat dengan label 0 atau normal dan kalimat dengan label 1 atau *bullying*. Berikut visualisasi *word cloud* dari masing-masing label.

Word Cloud Kategori Tweet Normal



Word Cloud Kategori Tweet Bullying



3.1. Word cloud kategori dengan label normal

3.2. Word cloud kategori dengan label *bullying*

Word Cloud merupakan visualisasi data yang menggambarkan kata yang paling sering muncul dari suatu kumpulan kalimat. Gambar 3.1 merupakan word cloud dari label normal, dapat dilihat kata-kata yang ada merupakan kata-kata yang tidak menunjukkan adanya *bullying* seperti kata enak, suka dll. Semnetara pada gambar 3.2 yang merupakan word cloud kategori label *bullying* menampilkan kata-kata yang kasar atau menyerang orang seperti kata cebong, sialan, tolol dll



### 3.2. Pra pemrosesan Data

Pra pemrosesan data dilakukan agar intisari input data teks dapat dengan mudah di ekstrak oleh model, dengan demikian model yang dihasilkan dapat generalisasi data dengan baik. berikut tahapan - tahapan *preprocessing* yang dilakukan, sebagai berikut:

#### 1. Mengubah seluruh alfabet menjadi huruf kecil

Keberagaman alfabet dapat mengakibatkan *preprocessing* dan pembelajaran *model* akan menurun.

#### 2. Melakukan *Mapping* singkatan kepada kata yang sesuai

Terdapat banyak permasalahan kosakata di dataset dimulai dari ketidaklengkapan kata, kata kata gaul dan typo. Hal ini perlu diperbaiki, kata - kata tersebut akan di *mapping* dengan kamus yang tersedia dari sumber [Kamus1](#) [20] dan [Kamus2](#). Sehingga, dapat meningkatkan performansi pada sistem yang dibuat.

Kalimat

‘sy BENCI kamu. g mau kenal :(’

Hasil *Mapping*

‘saya BENCI kamu. tidak mau kenal :(’

#### 3. Melakukan pelabelan data

Pelabelan didasarkan pada model *pipeline* yang dibentuk, parameter dilakukan pelabelan untuk setiap dataset berbeda, untuk dataset berita bohong, telah didasarkan pada pelabelan [turnbackhoax.id](https://turnbackhoax.id), dataset ujaran kebencian dan *A ofensif* dilakukan pelabelan secara manual dengan mendiskusikan pelabelan yang tepat bagi observasi data.

Kalimat

‘Dasar artis MURAHAN ga tau diri’

Hasil Pelabelan Data

1 atau *Bullying*

#### 4. Melakukan penghapusan karakter selain alfabet

Teknik ini bertujuan untuk karakter selain alfabet dan diharapkan model dapat belajar dari kata dan tidak menurunkan performansi dalam pembelajarannya sebab ada karakter yang tidak umum digunakan.





Kalimat

‘saya BENCI kamu. tidak mau  
kenal :’(

Hasil Penghapusan Selain Alfabet

‘saya BENCI kamu tidak mau  
kenal’

#### 5. Melakukan penghapusan *StopWord*

*Stopword Removal* adalah proses pembuangan kata yang tidak memiliki arti atau tidak relevan[21]. Hal ini juga bertujuan agar model dapat belajar pada kata yang esensial. Adapun list *stopword* yang digunakan pada penelitian ini berasal dari [Stop Word](#)[22].

Kalimat

‘saya benci kamu tidak mau kenal  
saya’

Hasil Penghapusan *StopWord*

‘benci kamu mau kenal’

#### 6. Melakukan penghapusan data dengan panjang kalimat kurang dari 4 kata

Panjangnya kalimat dapat menentukan dependensi di sebuah kalimat, hal ini tentu penting dalam mengembangkan model *deep learning*, kalimat dengan minimal 4 kata memungkinkan model mendapatkan hasil yang tepat.

#### 6. *Tokenizing*

Dikarenakan komputer hanya dapat memproses angka, *tokenizing* berfungsi untuk *mapping* urutan kata menjadi urutan angka.

### 3.3. Desain Sistem

#### 3.3.1. Arsitektur Model

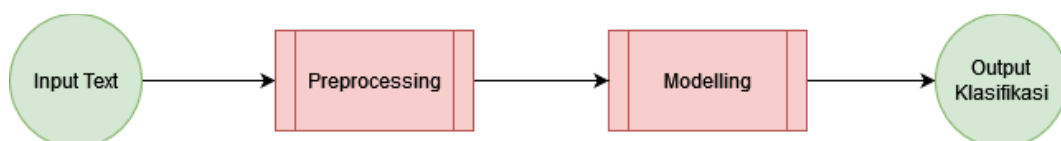
Dalam proses pembuatan dan training, ada dua jenis model yang akan dibangun. Pertama *training* dilakukan dengan menggunakan model *Machine learning* sebagai berikut:

1. *Naive bayes*
2. *Support vector machine*

Menggunakan model RNN sebagai Berikut:

1. LSTM ( *Long Short Term Memory* )

#### 3.3.2. Model Pipeline



Gambar 3.3 Model Pipeline

Gambar 3.3 merupakan gambaran alur model, pertama *text* akan melewati



proses preprocessing yang berguna untuk menyiapkan data yang lebih baik dan dapat dimengerti oleh model yang selanjutnya model akan mengeluarkan output klasifikasi apakah text terklasifikasi 0 (normal) atau 1 (*bullying*).

### 3.4. Evaluasi Model

#### 3.4.1. Metrics

Dalam penelitian ini ada beberapa *metrics* yang digunakan diantaranya sebagai berikut:

1. *Accuracy* : Akurasi memberikan informasi seberapa akurat model mengklasifikasikan data dengan benar.
2. *Precision* : Biasa disebut *specificity*, memberikan informasi terkait presentasi luaran prediksi yang relevan, dengan artian sesuai dengan labelnya dan tidak melakukan misklasifikasi FP.
3. *Recall* : Biasa disebut *sensitivity*, Memberikan informasi terkait presentasi luaran prediksi yang relevan, dengan artian sesuai dengan labelnya dan tidak melakukan misklasifikasi FN.

#### 3.4.2. Stratified Cross-validation

Stratified Cross-validation merupakan metode statistika yang bertujuan untuk mengevaluasi model agar tidak terjadi bias saat dilakukan pengukuran menggunakan *metrics* seperti akurasi. Konsep kerja dari *Stratified Cross-validation* dengan mengambil data dengan proposisi yang sesuai dan membaginya menjadi dua yaitu data latih dan data validasi. pada penelitian ini akan dilakukan

### 3.5 Studi Kasus

Terdapat macam jenis model *Machine learning* tradisional dan *Recurrent Neural Network* yang memiliki tujuan dan konsep kerja masing-masing. Pada penelitian ini akan digunakan tiga model dengan struktur dan parameter sebagai berikut:

#### 1. Machine learning

Pada penelitian ini menggunakan model seperti *support vector machine*(SVM) dan *Naive bayes* dengan parameter sebagai berikut

Parameter	Nilai
Word Embedding	TF-IDF
Maximum Features	14

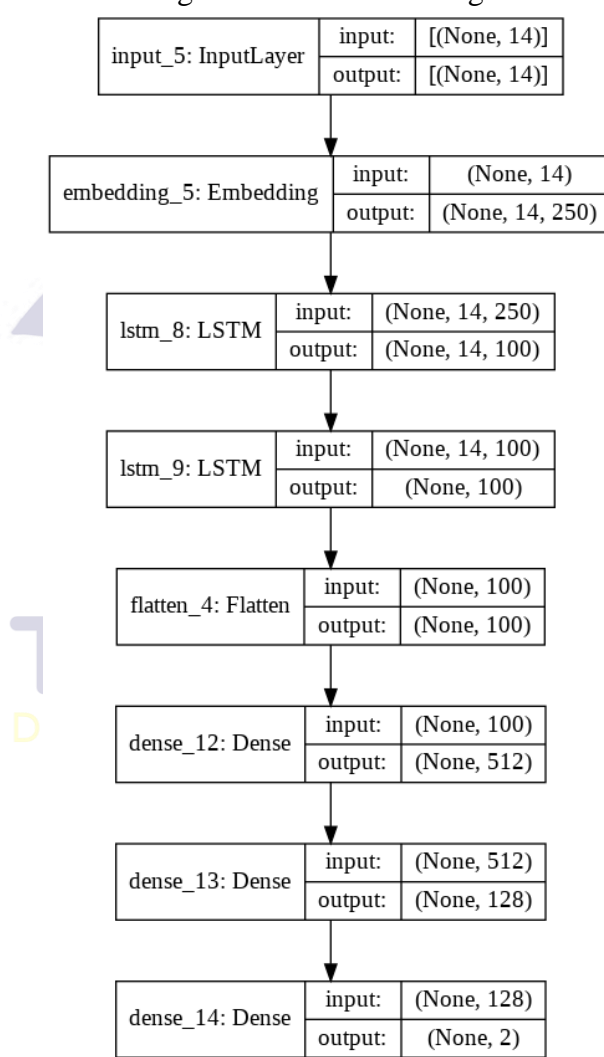


Maximum Length	250
C (SVM)	1.0
Gamma (SVM)	1.0
Var_smoothing ( <i>Naive bayes</i> )	1e-9

Tabel 3.1 Parameter Model SVM dan *Naive bayes*

## 2. Recurrent Neural Network

Model *Recurrent Neural Network* yang digunakan adalah LSTM dengan arsitektur dan parameter sebagai berikut model sebagai berikut:



3.4. Gambar arsitektur model LSTM

Parameter	Nilai
Word Embedding	Word2vec



<i>Maximum Features</i>	14
<i>Maximum Length</i>	250
<i>Epoch</i>	5
<i>Batch Size</i>	64
<i>LSTM</i>	activation=Tanh
<i>Dense</i>	activation=ReLu
<i>Output Layer</i>	activation=Softmax
<i>Optimizer</i>	Adam
<i>Loss Function</i>	<i>Categorical Cross Entropy</i>

Tabel 3.2 Parameter Model LSTM.

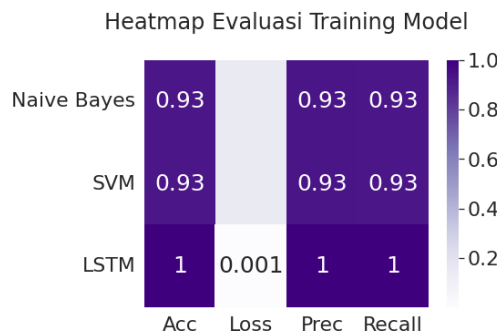
Penggunaan 5 *epochs* untuk menghindari banyaknya resource komputasi yang digunakan, *max sequence lengths* 250 didapat dari hasil analisis sebaran dataset. Fungsi *loss* yang digunakan *Categorical Cross Entropy* dikarenakan label memiliki struktur *one hot encoding* dan optimizer Adam. Sehingga pada akhirnya, Parameter - parameter tersebut dapat memberikan hasil yang *reliable*.

TADULAKO  
DATA CHALLENGE

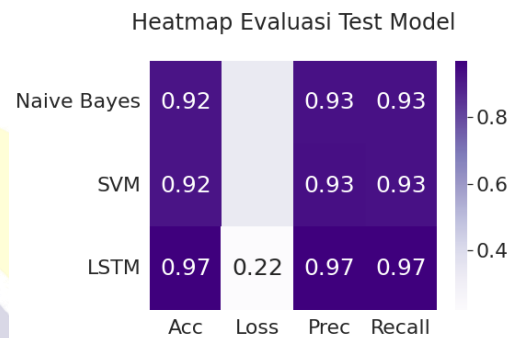
## 4. Hasil Pengujian

### 4.1 Training dan Validasi

Dalam proses training yang dapat dilihat pada [collab](#), *training* model dengan menggunakan parameter pada tabel 3.1 dan 3.2 juga menggunakan *stratified cross validation* untuk menghilangkan pengukuran yang bias yang akan diambil rata-rata dari hasil model menggunakan *stratified cross validation* mendapatkan hasil sebagai berikut:

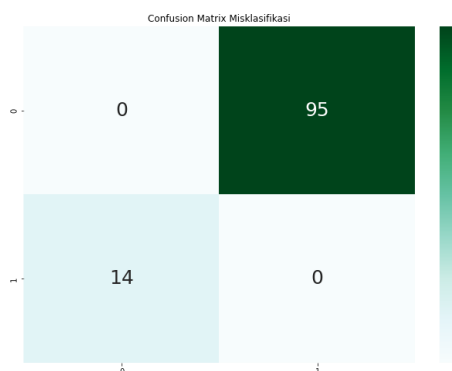


Gambar 4.1. Heatmap hasil evaluasi training model

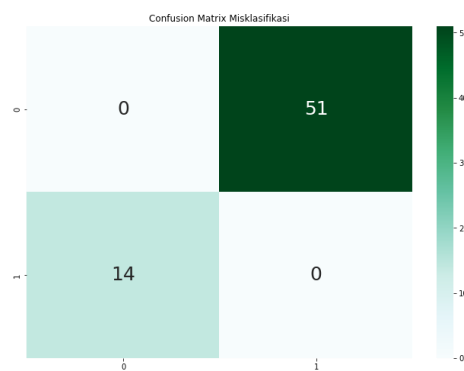


Gambar 4.2. Heatmap hasil evaluasi test model

Matriks utama dari penelitian ini adalah akurasi dikarenakan *dataset* telah *balanced*. LSTM menunjukkan performansi *training dataset* maupun *test dataset* dengan *loss training* hanya 0.001 dan *loss test* 0.22. Sedangkan model SVM dan model *naive bayes* memiliki performansi dibawah LSTM dengan akurasi *training* 0.93 dan akurasi *test* 0.92.

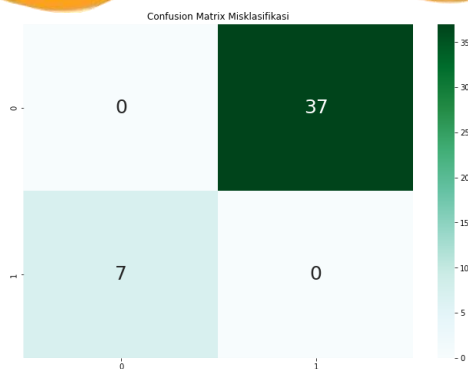


Gambar 4.3. Misclassification *confusion matrix* Naive bayes



Gambar 4.4. Misclassification *confusion matrix* SVM





Gambar 4.3. *Misclassification confusion matrix LSTM*

Gambar 4.3, gambar 4.4 dan gambar 4.5 merupakan *confusion matrix* dari misklasifikasi prediksi model, sumbu y menggambarkan label aktual dan sumbu x menggambarkan label prediksi. Dari ketiga model memiliki kecenderungan memprediksi label *bullying* dilihat dari banyaknya *false negative* yaitu kalimat yang seharusnya diprediksi sebagai normal akan tetapi diprediksi sebagai *bullying*.

#### 4.2. Hasil Klasifikasi

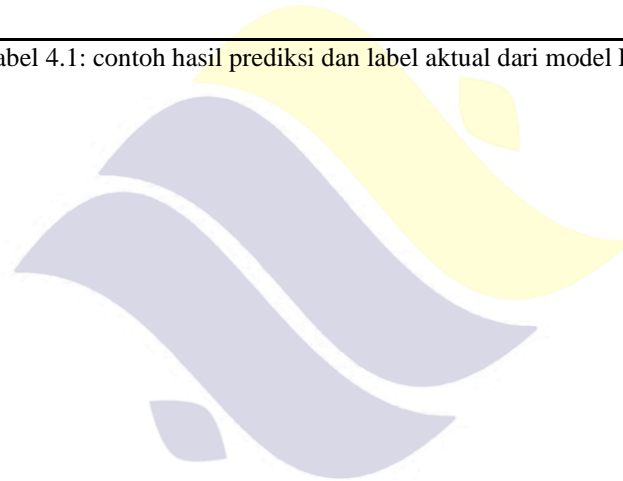
Tabel berikut sampel hasil klasifikasi dari model terbaik yaitu model LSTM dengan label 0 merupakan label normal dan 1 merupakan label *bullying*.

No	Teks	Prediksi
1	dasar artis abal abal enggak laku layar kacaudh enggak duit tidak punya uang miskinjadi main dukun minta ilmu pelet perebut suami orang kaya	1
2	restoran berbentuk kompleks berupa saungsaung berada kebunkebun air mancur aliran sungai bergemerik indah malam hari pijaran api peneangan jalan memperindah suasana makanan cukup enak enak suasana menonjol tempat	0
3	wakil gubernur tingkalakunya berpakaian anak pejabat pemerintahan stres kali orang	1
4	tempat nya asyik nongkrong ngobrol barengbareng teman teman roti bakar nya enak ditambah kopi susu dikunjungi enak kumpul barengbareng temanteman kumpul barengbareng temanteman harga terjangkau	0
5	sekeluarga ngefans kagum lihat nya dehtapi lihat anak nya jelekbuluk hermansyah kutuk menimpa bunda kd untung allah	1



	swt menggantikan amora adik nyagemez lihat nya cocok banget jadi anak artis cla bunda kd	
6	buta hati buta mata lihat tenaga kerja aseng ilegal berkeliaran negeri dungu piara	1
7	edan orang udah makan uang rakyat tidur	1
8	langsung bilang maksud klu gini namanya munafik	1
9	respect istri anak nya menunggu sabar dirumah insya allah allah swt menjaga nya	0
10	kfc umum nya ruangan nya cukup besar ruang khusus acara ulang tahun nya tempat bermain anak lumayan bawa anak	0

Tabel 4.1: contoh hasil prediksi dan label aktual dari model LSTM



TADULAKO  
DATA CHALLENGE



## 5. ANALISIS

Berdasarkan observasi dan analisis pada bab sebelumnya, ditemukan bahwa ketiga model secara garis besar dapat mengklasifikasikan data dengan tepat. LSTM merupakan model terbaik dibanding dengan model SVM dan model *naive bayes* dikarenakan memiliki performansi yang lebih baik. Model LSTM dapat mencapai hingga akurasi 1.00 di *train* dan 0.97 pada *test set*. Ketiga model tersebut juga berhasil memperkecil *false negative* dengan hasil model *naive bayes*=14, model SVM=14 dan LSTM=7. Hal ini menunjukkan bahwa ketiga model tersebut berhasil mengidentifikasi label *bullying* dengan baik, akan tetapi ketiga model tersebut juga lebih cenderung memprediksi label *bullying* yang dapat dilihat dari *false positive* dengan hasil model *naive bayes*=95, model SVM=51 dan LSTM=37 dari total data 7411. Dalam *train set* terdapat kalimat yang terlabel *bullying* sementara menurut peneliti merupakan kalimat normal, hal ini bisa terjadi karena kesalahan pelabelan dari sumber data yang bisa mempengaruhi model. Jumlah kalimat yang bermasalah ini tidak banyak sejauh observasi oleh peneliti.”

TADULAKO  
DATA CHALLENGE



## 6. KESIMPULAN

Berdasarkan observasi dan analisis, secara garis besar model *support vector machine*, *naive bayes*, dan LSTM memiliki performansi yang baik dalam melakukan klasifikasi apakah suatu kalimat berpotensi sebagai kalimat *bullying* atau tidak. Performansi ketiga model tersebut dikatakan baik dikarenakan pengukuran *metrics* menunjukkan akurasi diatas 90% untuk ketiga model pada *training set* maupun *test set* dan juga dilihat dari Keberhasilan model untuk memperkecil *false positive*. Pemberlakuan analisis teks yang lebih dalam perlu dilakukan mengingat urgensi penggunaan *tools* ini untuk meminimalisir kasus *cyberbullying* di Indonesia. Penggabungan dataset yang dilakukan dari 5 sumber yang berbeda memberikan data yang kaya akan informasi dan hal ini dapat dimanfaatkan oleh penelitian lainnya. Beberapa pengembangan lebih lanjut dapat dilakukan untuk menyempurnakan model ini, antara lain penambahan dataset dari sumber-sumber lain, perbaikan label dapat dilakukan untuk meningkatkan performa dan juga bisa mencoba model yang lain seperti *transformers*, BiLSTM, CNN-LSTM dan sebagainya untuk melihat perbandingan performansi lebih jauh antara model tersebut. Selain itu, model dapat dikembangkan lebih lanjut untuk aplikasi lain seperti *question answering* ataupun *chatbot*.

TADULAKO  
DATA CHALLENGE



## DAFTAR PUSTAKA

- [1] Beran, T., & Li, Q. (2005). Cyber-harassment: A study of a new method for an old behavior. *Journal of Educational Computing Research*. 32(3), 265-277.
- [2] Asosiasi Penyelenggara Jasa Internet Indonesia (APJII). (2017). Infografis: Penetrasi dan perilaku pengguna internet Indonesia. Teknopreneur. Diakses pada tanggal 27 September 2021 dari <https://www.apjii.or.id/content/read/39/342/Hasil-Survei-Penetrasi-dan-Perilaku-Pengguna-Internet-Indonesia-2017>.
- [3] Ningrum, F.S dan Amna, Z. (2020). *Cyberbullying* Victimization dan Kesehatan Mental pada Remaja. *Jurnal Psikologi dan Kesehatan Mental*. 2528-5181.
- [4] Sam, D. L., Bruce, D., Agyemang, C. B., Amponsah, B., & Arkorful, H. (2017). *Cyberbullying* victimization among High School and University Student in Ghana. *Deviant Behavior*. 1-17
- [5] Badan Pusat Statistik. Jumlah Penduduk Menurut Kelompok Umur dan Jenis Kelamin, 2020. Diakses pada tanggal 27 September 2021 dari [https://www.bps.go.id/indikator/indikator/view\\_data/pub/0000/api\\_pub/YW40a21pdTU1cnJxOGt6dm43ZEoZz09/da\\_03/1](https://www.bps.go.id/indikator/indikator/view_data/pub/0000/api_pub/YW40a21pdTU1cnJxOGt6dm43ZEoZz09/da_03/1)
- [6] Pratama, K.R. (2021). Instagram, media Sosial Pemicu “Cyberbullying”. Diakses pada tanggal 27 September 2021 dari <https://tekno.kompas.com/read/2021/03/29/07164137/instagram-media-sosial-pemicu-cyberbullying-tertinggi>.
- [7] Hinton, Geoffrey E. (1992). How Neural Networks Learn from Experience. *Scientific American*, vol. 267, no. 3, 1992, pp. 144–151.
- [8] Stauffer, S., Heath, M. A., Coyne, S. M. & Ferrin, S., 2012. High School Teachers Perceptions of Cyberbullying Prevention and Intervention Strategies. *Psychology in the Schools*, Volume 49.
- [9] Lestari, A.R.T, Perdana, R.S, dan Fauzi, M.A. (2017). Analisis Sentimen Tentang Opini Pilkada DKI 2017 Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naïve Bayes dan Pembobotan Emoji.
- [10] Primartha, R (2018). Belajar *Machine learning*. Informatika. Bandung.
- [11] Samsudiney. (2019). Penjelasan Sederhana tentang Apa Itu SVM?. Diakses pada tanggal 27 september 2021 dari <https://medium.com/@samsudiney/penjelasan-sederhana-tentang-apa-itu-svm>.
- [12] Chazar, C dan Widhiaputra, B.E. (2020). *Machine learning* Diagnosis Kanker Payudara Menggunakan Algoritma *Support vector machine*. *Jurnal Informatika dan Sistem Informasi*, 12(1).
- [13] Lipton, Zachary C., et al. (2015). A Critical Review of Recurrent Neural Networks for Sequence Learning. arXiv:1506.00019.





- [14] Ibrohim, M, O and Budi, I. (2019). Multi-label Hate Speech and Abusive Language Detection in Indonesian Twitter. In ALW3: 3rd Workshop on Abusive Language Online, 46-57.
- [15] Luqyana, W., Cholissodin, I., dan Perdana, R.S. (2018). Analisis Sentimen *Cyberbullying* pada Komentar Instagram dengan Metode Klasifikasi *Support vector machine*. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 2(11), 4704-4713. Diambil dari <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/3051>
- [16] Purwarianti, A. dan Crisdayanti, I. A. P. A. (2019). Improving Bi-LSTM Performance for Indonesian Sentiment Analysis Using Paragraph Vector', *Proceedings of the 2019 International Conference of Advanced Informatics: Concepts, Theory dan Applications (ICAICTA)*, pages 1-5..
- [17] Ibrohim, M, O dan Budi, I. (2018). A Dataset dan Preliminaries Study for Abusive Language Detection in Indonesian Social Media. *Procedia Computer Science*, 135(), 222–229. doi:10.1016/j.procs.2018.08.169
- [18] Ibrohim, M, O dan Budi, I. (2019). Multi-label Hate Speech dan Abusive Language Detection in Indonesian Twitter. In ALW3: 3rd Workshop on Abusive Language Online, 46-57.
- [19] S, Aliyah. et al.ef. (2018). *[IEEE 2018 International Conference on Asian Language Processing (IALP) - Bandung, Indonesia (2018.11.15-2018.11.17)] 2018 International Conference on Asian Language Processing (IALP) - Colloquial Indonesian Lexicon. , (), 226–229. doi:10.1109/IALP.2018.8629151.*
- [20] Maarif, A, A. (2015). Penerapan Algoritma TF-IDF untuk Pencarian Karya Ilmiah, Udinus Repo.
- [21] Tala, F. Z. (2003). A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia. M.Sc. Thesis. Master of Logic Project. Institute for Logic, Language dan Computation. Universiteit van Amsterdam, The Netherlands.

TADULAKO  
DATA CHALLENGE



## Lampiran

Link source code :

[https://colab.research.google.com/drive/196eI5D8uT3cerX\\_nDkdS0tKgboi7Nhjw?usp=sharing](https://colab.research.google.com/drive/196eI5D8uT3cerX_nDkdS0tKgboi7Nhjw?usp=sharing)

Link *dataset* yang dipergunakan :

[https://github.com/fendiirfan/cyberbullying\\_TCD\\_UNTAD/blob/main/Dataset.csv](https://github.com/fendiirfan/cyberbullying_TCD_UNTAD/blob/main/Dataset.csv)



**TADULAKO**  
DATA CHALLENGE