

Identifikasi Audio Ancaman Menggunakan Metode Convolutional Neural Network

Abstrak

Sosial media merupakan sebuah aplikasi yang berbasis internet dan dapat menunjang fungsi interaksi pada masyarakat. Berdasarkan data laporan yang diambil dari Kominfo, terdapat banyak konten negatif yang berisi ujaran kebencian. Berdasarkan masalah tersebut maka akan dibuat sebuah sistem yang bertujuan untuk mendeteksi ancaman terutama pada audio. Sehingga dapat mengurangi dan menyaring konten-konten yang berisi suara ancaman. Pada proses pembuatan sistem pendeteksi maka dibutuhkan beberapa sampel data ancaman yang akan diolah. Pengumpulan data akan dilakukan dengan menggunakan Web Scraping pada sosial media twitter. Setelah terkumpul data akan dilakukan preprocessing. Pengolahan data akan menggunakan metode Convolutional Neural Network. Akurasi yang dihasilkan dengan menggunakan metode CNN tersebut adalah 82%. Model yang didapatkan dari metode tersebut akan digunakan sebagai bahan untuk melakukan prediksi audio ancaman. Audio ancaman akan dilakukan konversi menjadi teks menggunakan speech recognition yang kemudian akan dilakukan prediksi dengan menggunakan model tersebut. Hasil dari prediksi yang dilakukan menghasilkan output berupa ancaman atau bukan ancaman.

Kata kunci: Convolutional Neural Network, Web Scraping, Audio, Speech Recognition, Text

Threat Audio Identification Using Convolutional Neural Network Method

Abstract

Social media is an internet based application that can support social interaction in society. Based on report that taken from Kominfo, there are negative content that contain hate speech. Based on that problem, A system that intended to detect any threat especially in audio so that can reduce and filter threatening contents. A sample hate speech data is needed to create detection system that will be processed. Sample data are gathered by using Web Scraping method in twitter social media. After the data is gathered, it will be preprocessed. The data that has been preprocessed will enter the processing by Convolutional Neural Network. The Result of accuracy using this CNN method is 82%. Models that obtained by this method will be used as hate speech audio prediction. Hate speech audio will be converted into text using Speech Recognition that will be predicted with this model. The result of prediction is text output that consists of threat or not threat.

Keywords: Convolutional Neural Network, Web Scraping, Audio, Speech Recognition, Text

I. PENDAHULUAN

Sosial media merupakan sebuah aplikasi yang berbasis internet dan dapat menunjang fungsi interaksi pada masyarakat. Sosial media dapat digunakan sebagai penunjang fungsi interaksi pada masyarakat. Sosial media dapat dimanfaatkan sebagai media komunikasi baik secara personal maupun secara kelompok atau grup. Sosial media juga menjadi tempat favorit untuk membuat konten baik konten positif maupun negatif.

Konten negatif yang terdapat pada media sosial salah satunya adalah ujaran kebencian. Makna ujaran kebencian secara hukum memiliki arti perkataan, perilaku, tulisan, ataupun pertunjukan yang dilarang karena dapat menimbulkan terjadinya kekerasan dan prasangka dari tindakan tersebut. Berdasarkan data laporan yang diambil dari Kominfo, terdapat banyak konten negatif yang berisi ujaran kebencian. Saat ini sudah dilaporkan beberapa konten kebencian yaitu (1) SARA 188 data, (2) Fitnah 17

data, (3) Kekerasan atau kekerasan pada anak 11 data, (4) Konten meresahkan masyarakat 23 data, Konten pelanggaran nilai social dan budaya 26 data. Jika diambil data keseluruhan dari 18 kategori pelanggaran, maka akan terdapat 1.570.541 konten yang dilaporkan.

Berdasarkan masalah tersebut maka akan dibuat sebuah sistem yang bertujuan untuk mendeteksi ancaman terutama pada audio. Sehingga dapat mengurangi dan menyaring konten-konten yang berisi suara ancaman. Pada proses pembuatan sistem pendeteksi maka dibutuhkan beberapa sampel data ancaman yang akan diolah.

Pengumpulan sample tersebut akan dilakukan melalui *Web Scraping*. *Web scraping* akan dilakukan pada sosial media *twitter* untuk mendapatkan data teks yang akan diolah. Setelah data dari *twitter* didapatkan, selanjutnya akan dilakukan pembersihan data, labeling dan mengubah data dari teks menjadi kumpulan *array* dengan menggunakan *sequence padding*.

Pada kasus ini, akan digunakan sebuah metode untuk mengolah data. Data akan diolah menggunakan metode *Convolutional Neural Network (CNN* atau *ConvNet*). Tujuannya adalah untuk mengekstraksi data sehingga dapat dibuat menjadi sebuah model. *Output* dari *CNN* adalah sebuah model yang akan dijadikan acuan oleh komputer untuk melakukan prediksi.

Setelah model terbentuk, maka komputer dapat memprediksi melalui model tersebut. Hal yang paling utama adalah melakukan input audio baik berupa rekaman langsung maupun file. Audio tidak bisa langsung dilakukan prediksi karena perlu diubah menjadi teks. Pada tahap konversi audio akan digunakan *Speech Recognition* menggunakan *Google Text-To-Speech*. Setelah dikonversi, teks dapat diubah kedalam bentuk vector dan dapat dilakukan prediksi. Hasil dari prediksi terbagi menjadi dua kategori yaitu Ancaman (1) dan Bukan Ancaman (0)

II. STUDI LITERATUR

A. Studi Literatur

Penelitian mengenai analisis sentiment telah dilakukan sebelumnya dengan metode *CNN* namun dengan data yang berbeda-beda. Salah satunya seperti penelitian yang dilakukan oleh Y. Kim [1] menggunakan dataset yang diambil dari beberapa sumber seperti *Movie Review* dan *Stanford Sentiment Treebank*. Dataset akan diolah menggunakan *CNN* dengan melakukan konversi menjadi vektor. Hasil akhir dari modelling didapatkan akurasi hingga 89%. Selain itu penelitian lain juga dilakukan oleh S. Nadhifa Ayu [2] dimana penelitian menggunakan Metode *CNN* dan data yang digunakan diambil melalui *twitter*. Tingkat akurasi yang didapatkan adalah 84%. Penelitian selanjutnya dilakukan oleh Y. Luan [3] dimana data yang digunakan sebagai berjumlah 5000 data teks subjektif dan 5000 data teks objektif. Data tersebut diolah menggunakan *CNN* dan mendapatkan akurasi dari *f1-score* sebesar 98%, *precision* 98%, dan *recall* 98%. Penelitian selanjutnya dilakukan oleh S. Pedro M dimana dilakukan *sentiment analysis* pada *twitter* dengan menggunakan beberapa metode yang satunya adalah *CNN*. Data yang digunakan berjumlah 1.578.627 data yang sudah diberi label positif atau negative. Hasil dari penelitian tersebut didapatkan akurasi sebesar 66.7%. [4]

B. Web Scraping

Web scraping merupakan teknik yang digunakan untuk mendapatkan informasi dari sebuah atau beberapa *website* secara otomatis dan tanpa harus menyalinnya secara manual.[5]

Web scraping dapat disebut dengan *intelligent, automated*, atau *autonomous agents*. Tujuan *web scraping* hanya berfokus pada cara memperoleh data melalui pengambilan dan ekstraksi data dari *website* dengan ukuran data yang bervariasi mulai dari kecil hingga besar.[6]

Hal yang paling penting pada *web scraping* adalah untuk menambang informasi data pada *website* yang berbeda beda strukturnya dan mengubahnya menjadi data yang terstruktur dan disimpan dalam *spreadsheet*, *database* atau sebuah file *Comma Separated Value (CSV)*. [7]

Web scraping cukup mudah digunakan dalam menghadapi struktur *HTML* yang berbeda beda pada setiap webnya. Selain itu, penggunaan *Web scraping* yang dapat digunakan untuk menghadapi struktur *HTML* yang berbeda antar webnya akan mempermudah dalam pengumpulan data mulai dari data yang kecil hingga data yang besar.

Salah satu library yang dapat digunakan untuk melakukan *web scraping* pada bahasa pemrograman *python* adalah *selenium*. *Selenium* merupakan library untuk *web automation test*. *Selenium* dapat dengan mudah digunakan sebagai *web scraper* sehingga dapat membantu proses *scraping*.

C. Speech Recognition

Speech recognition atau dikenal dengan *automatic speech recognition (ASR)*, *computer speech recognition* atau *speech-to-text* merupakan sebuah kemampuan yang dapat membuat sebuah program mampu melakukan proses kemampuan membaca suara manusia dan melakukan konversi ke dalam teks.[8]

Komponen yang dibutuhkan dari *speech recognition* adalah suara. Suara perlu dilakukan konversi dari suara fisik ke sinyal elektronik menggunakan *microphone* dan dilakukan konversi kembali ke sinyal digital menggunakan *analog-to-digital converter*. [9]

Speech recognition memiliki perbedaan dengan *voice recognition*. Perbedaannya terletak pada tujuan penggunaan, *speech recognition* digunakan untuk menerjemahkan dari suara manusia atau suara percakapan menjadi teks. Sedangkan *voice recognition* digunakan hanya untuk mengenali suara manusia saja.[8]

Pada bahasa *python*, *speech recognition* cukup mudah dilakukan karena adanya bantuan dari library *Google Text-To-Speech (gTTS)*. *gTTS* merupakan sebuah *API* yang dapat melakukan *speech recognition* dari kata kata yang berupa audio.[10] Library ini dapat mengubah dari bentuk audio ke teks maupun melakukan penerjemahan teks antar bahasa.

D. Convolutional Neural Network

Convolutional Neural Network (ConvNet/CNN) merupakan algoritma *deep learning* yang dapat menggunakan gambar sebagai input data yang akan diolah, menetapkan bobot dan bias yang akan digunakan untuk pemodelan pada berbagai aspek dan objek dalam gambar agar dapat membedakan satu dengan yang lain.[11]

Convolutional Neural Network atau *CNN* merupakan sebuah model dalam *machine learning* yang paling cocok digunakan dalam *image recognition* dan juga dapat memberikan hasil yang bagus dalam melakukan *Natural Language Processing* karena terdapat beberapa layer yang dapat membuat kata kata dapat di deteksi bersamaan.[12]

CNN terdiri dari beberapa lapis layer yang saling berhubungan. Pada jaringan *CNN*, setiap hasil keluaran dari sebuah *layer* akan dijadikan masukan pada layer selanjutnya sehingga operasi *Multi-Layer Convolution* digunakan untuk melakukan transformasi hasil dari setiap layer secara nonlinear sampai output layer.[13]



Gambar 1. Ilustrasi Layer pada CNN

Seperti pada Gambar 1, *CNN* terbagi menjadi enam layer yaitu:

1. *Input Layer*
Berfungsi memasukan inputan teks yang sudah diubah menjadi angka.
2. *Convolutional Layer*
Digunakan untuk ekstraksi fitur pada data teks dengan menggunakan *kernel* konvolusi.
3. *Incentive Layer*
Untuk menambahkan *mapping linear* karena convolusi merupakan pengolahan secara *linear*.
4. *Pooling Layer*
Berguna untuk melakukan *down sampling* dan mereduksi jumlah dari data yang diolah.
5. *Fully Connected Layer*
Berisi vektor *w* dimensi hasil dari *pooling layer* yang terhubung dengan *softmax layer* untuk mengurangi *feature loss*.
6. *Output Layer*
Hasil dari proses yang telah dilakukan pada layer layer sebelumnya.

E. *Confusion Matrix*

Confusion matrix digunakan sebagai alat evaluasi dari sebuah model untuk menghitung *precision*, *recall*, dan *f1-score*. [3] *Confusion matrix* memiliki bentuk seperti pada Tabel 1.

TABEL I
CONFUSION MATRIX

	Prediksi Negatif	Prediksi Positif
Actual Negatif	True Negative	False Positive
Actual Positif	False Negative	True Positive

True Negative(TN) merepresentasikan bahwa data hasil diprediksi negatif dan data *actual* yang dihasilkan negatif.

True Positive(TP) merepresentasikan bahwa data diprediksi positif dan data *actual* yang dihasilkan positif.

False Positive(FP) merepresentasikan bahwa data diprediksi positif namun data *actual negative*.

False Negative(FN) merepresentasikan bahwa data data diprediksi negatif namun data *actual* positif.

Perhitungan *precision* menunjukkan ketepatan antara prediksi dengan data *actual*. Perhitungan pada presisi dapat dilakukan dengan menggunakan persamaan(1).

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

Perhitungan *recall* digunakan untuk menggambarkan persentase informasi relevan yang berhasil diambil dan dideteksi oleh model. Perhitungan *recall* dapat menggunakan persamaan (2).

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

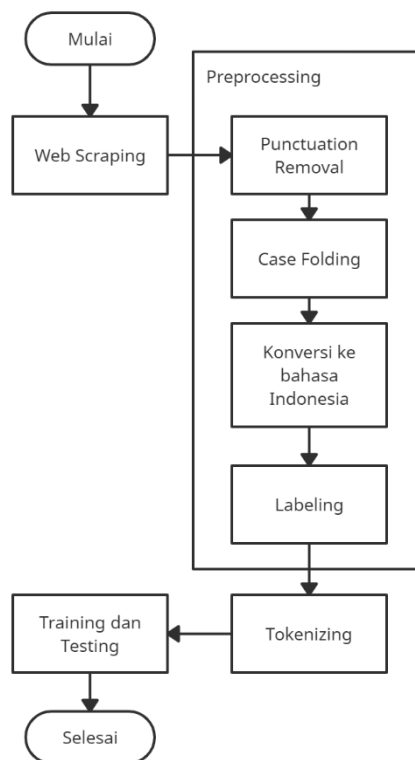
F1-score merupakan rata-ran harmonik yang diperoleh dari perhitungan *precision* dan *recall*. Nilai *f1-score* terpengaruh dari *precision* dan *recall*. Perhitungan *f1-score* dapat digunakan persamaan (3)

$$F1 = 2 \times \frac{precision \times recall}{precision+recall} \tag{3}$$

III. METODOLOGI PENELITIAN

Pada tahap ini akan dilakukan analisis untuk mengamati kebutuhan dalam melakukan klasifikasi audio ancaman. Dalam melakukan klasifikasi dibutuhkan sebuah data teks yang dikumpulkan dari *twitter*. Data tersebut berisi teks ancaman. Pada tahap ini akan dilakukan analisis metode yang tepat untuk digunakan dalam mengumpulkan data hingga melakukan klasifikasi.

Untuk tahapan penelitian yang akan dilakukan dalam melakukan analisis dapat dilihat pada Gambar 2 berikut.



Gambar 2: Tahapan Penelitian

A. *Web Scraping*

Pada tahap ini, akan dilakukan akan dilakukan pengumpulan data sebagai bahan dasar dalam melakukan analisis. Data akan disimpan dalam sebuah dataset dalam bentuk *CSV*. Data akan diperoleh melalui *twitter* dan dilakukan dengan menggunakan metode *web scraping*. *Web scraping* yang digunakan dibantu dengan *library selenium* pada *python*. Pengumpulan data dilakukan berdasarkan topik dari *hashtag* yang sedang *trending*. Atribut yang diambil ketika melakukan *scraping* adalah nama, *username* dan *tweet* teks. Data yang sudah terkumpul akan disimpan kedalam bentuk *CSV* dengan *library pandas* pada *python*. Setelah tersimpan, maka data tersebut akan dibagi menjadi dua bagian yaitu data *testing* dan data *training*. Hasil dari *scraping* dapat dilihat pada Gambar 3.

	Nama Akun	User Name	Konten
0	U9lykdsjoe2	@u9lykdsjoe2	Wake up before it's too late!\n#GangsterIstana ...
1	U9lykdsjoe2	@u9lykdsjoe2	Ane retweet...\n#GangsterIstana!\n#JokowiMelang...
2	Z1kr1y4	@Z1kr1y4	Pemimpin Jenius Dan terbaik '\n\n#GangsterIst...
3	ZK3Alph	@ZK3Alph	Cling...\n#GangsterIstana '\n#GangsterIstana
4	ZK3Alph	@ZK3Alph	Blink '\n#GangsterIstana '\n#GangsterIstana
5	K.E.Y.S.H.A #DIRumahAja	@AdindaKeyshaNew	I could start a war in 30 seconds. But some co...
6	K.E.Y.S.H.A #DIRumahAja	@AdindaKeyshaNew	All of us deserve a greater peace of mind, kno...
7	K.E.Y.S.H.A #DIRumahAja	@AdindaKeyshaNew	My family is the engine of everything & on a p...

Gambar 3: Hasil Web Scraping

B. Preprocessing

Data yang terkumpul dan sudah disimpan dalam format CSV masih berbentuk data yang mentah. Tahap selanjutnya adalah membersihkan data dengan *preprocessing* data. *Preprocessing* dilakukan untuk menghilangkan kata yang tidak dibutuhkan di dalam data seperti *hashtag*, *username*, *URL*, *mention* dan *html tag*. Pada tahap ini atribut yang diambil hanya pada atribut konten yang berisi teks. Tahap yang akan dilakukan untuk *preprocessing* adalah:

1. Punctuation removal

Pada *dataset* yang mentah masih terdapat tanda baca yang tidak diperlukan seperti titik (.), koma (,), hashtag (#), at (@), dan html tag (\n). Karakter tambahan tersebut tidak akan diproses untuk mengurangi gangguan atau *noise* pada data.

2. Case Folding

Setelah menghapus tanda baca, maka tahap selanjutnya adalah *case folding*. *Case folding* merupakan tahap untuk mengubah semua huruf dalam teks menjadi huruf kecil.

3. Konversi ke bahasa Indonesia

Pada tahap ini *dataset* yang telah dilakukan *case folding* akan diterjemahkan ke bahasa Indonesia. Banyaknya data dari warga negara Indonesia yang menggunakan bahasa Inggris serta batasan masalah pada penelitian ini hanya pada bahasa Indonesia sehingga perlu dilakukan konversi bahasa.

4. Labeling

Pada tahap ini, data sudah bersih dan akan diberi label. Labeling dilakukan untuk melakukan identifikasi pada data teks yang digunakan. Data yang sudah bersih akan diberi label berdasarkan kalimatnya. Data akan diberikan label ancaman atau bukan ancaman. Jika pada sebuah kalimat memiliki kecenderungan untuk melakukan ujaran kebencian maka akan diberi label ‘ancaman’ yang menyatakan bahwa data tersebut benar merupakan ujaran kebencian. Namun, jika data tidak memiliki kecenderungan ujaran kebencian maka akan diberikan label ‘bukan ancaman’.

5. Tokenizing

Tokenizing merupakan tahap pemotongan kalimat menjadi kata kata yang pemotongannya dilakukan dengan cara memotong spasi atau tanda baca. Pada tahap ini, *tokenizing* akan membuat kata kata menjadi sebuah *token*. Tokenisasi dilakukan untuk mengurangi besarnya data yang tersimpan dan mengurangi jarak pencarian ketika mencari kata kata [14].

C. Training dan Testing

Data yang telah melewati tahap *preprocessing*, selanjutnya akan masuk ke tahap *processing data*. *Processing data* meliputi tahap *training* dan *testing*. Tahap pertama merupakan *encoding* dari label. Pada saat *encoding* akan menggunakan *label_encoder* untuk mengubah nilai label ke bentuk angka. Label tersebut akan diubah menjadi angka 1 untuk data yang berupa ancaman dan 0 untuk data yang bukan berupa ancaman.

Kemudian data akan diubah menjadi dua bagian yaitu data *training* dan data *testing*. Data akan terbagi menjadi 75% *training* dan 25% *testing*. Setelah terbagi, masing masing data akan dibagi menjadi 2 variabel yaitu variabel X dan Y. Pada variabel X, data akan berisi konten yang akan di klasifikasi. Pada variabel Y, data akan berisi label dari konten tersebut.

Setelah itu data teks akan dikonversikan kedalam bentuk bilangan seperti pada Gambar 4. Saat dilakukan konversi, data akan berbentuk *array 2 dimensi*. Tahap konversi akan dilakukan dengan menggunakan *sequence padding*. Tujuannya adalah untuk membuat data kalimat menjadi jumlah *array* yang sama panjang antara kalimat yang satu dengan yang lainnya. *Array* disini terbentuk melalui *index* angka ketika tahap *tokenizing* yang sebelumnya dilakukan. Panjang *array* yang digunakan dalam penelitian ini dibatasi dengan jumlah 10000 kata. Jika terdapat jumlah *array* yang panjangnya kurang dari panjang maksimal *array*, maka akan digantikan dengan sebuah angka seperti pada Gambar 4.

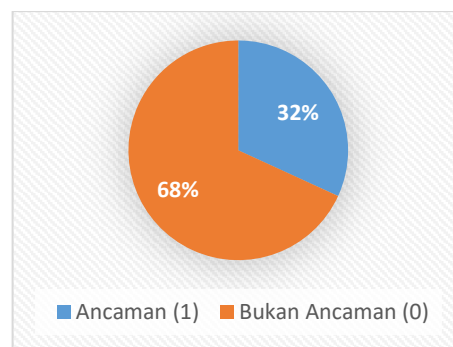
```
array([[ 0,  0,  0, ..., 121,  60, 345],
       [ 0,  0,  0, ...,  5,  24,  31],
       [ 0,  0,  0, ...,  3, 122,  11],
       ...,
       [ 0,  0,  0, ..., 24,  31, 299],
       [ 0,  0,  0, ..., 173, 173, 1084],
       [ 0,  0,  0, ...,  0, 1085, 1086]], dtype=int32)
```

Gambar 4: Sequence Padding

Setelah data teks dibuat menjadi *array*, maka selanjutnya akan dilakukan pemodelan dengan menggunakan *Convolutional Neural Network* sebagai metode yang digunakan untuk mengolah data.

IV. ANALISIS DAN PEMBAHASAN

Data yang diperoleh ketika melakukan *web scraping* berjumlah 468 data dengan persentase data yang berisi kalimat positif dan negatif seperti ditunjukkan pada Gambar 5.



Gambar 5: Pie Chart Perbandingan Data

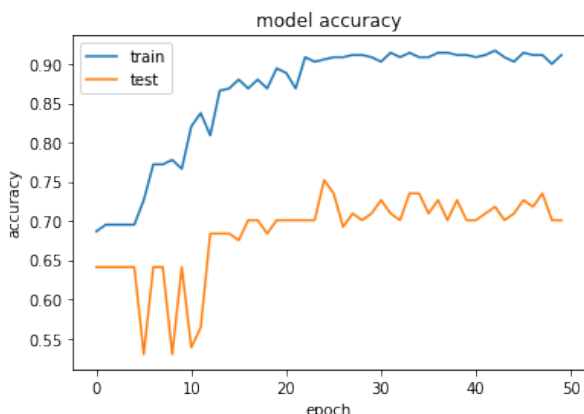
Berdasarkan diagram pada Gambar 5, data yang berupa ancaman (1) berjumlah 149 data. Sedangkan untuk data yang berupa bukan ancaman(0) terdapat 319 data. Selanjutnya adalah melakukan *modelling* dengan menggunakan metode *Convolutional Neural Network*. Pada percobaan pertama, akan digunakan struktur model CNN seperti pada Gambar 6.

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 40)]	0
embedding_1 (Embedding)	(None, 40, 20)	31380
conv1d_3 (Conv1D)	(None, 38, 32)	1952
max_pooling1d_2 (MaxPooling 1D)	(None, 12, 32)	0
conv1d_4 (Conv1D)	(None, 10, 64)	6208
max_pooling1d_3 (MaxPooling 1D)	(None, 3, 64)	0
conv1d_5 (Conv1D)	(None, 1, 128)	24704
global_max_pooling1d_1 (GlobalMaxPooling1D)	(None, 128)	0
flatten (Flatten)	(None, 128)	0
dense (Dense)	(None, 5)	645

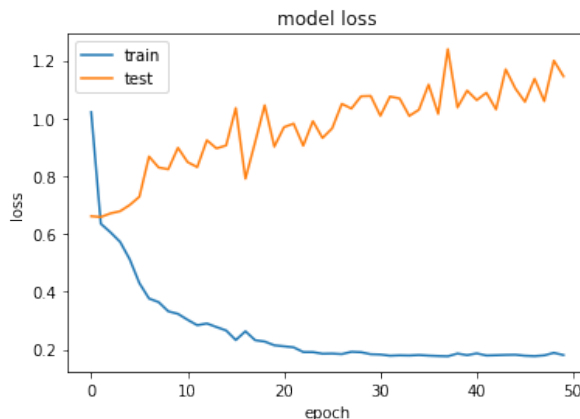
 Total params: 64,889
 Trainable params: 64,889
 Non-trainable params: 0

Gambar 6: Arsitektur CNN Uji Coba Pertama

Dengan menggunakan arsitektur tersebut, maka akan digunakan pemodelan dengan *epoch* = 50. *Epoch* merupakan sebuah perputaran yang digunakan dalam melakukan *training data* oleh *Neural Network* terutama pada *CNN*. [15] Berdasarkan penggunaan model tersebut didapatkan hasil seperti histogram model *accuracy* pada Gambar 7 dan *histogram model loss* pada Gambar 8.

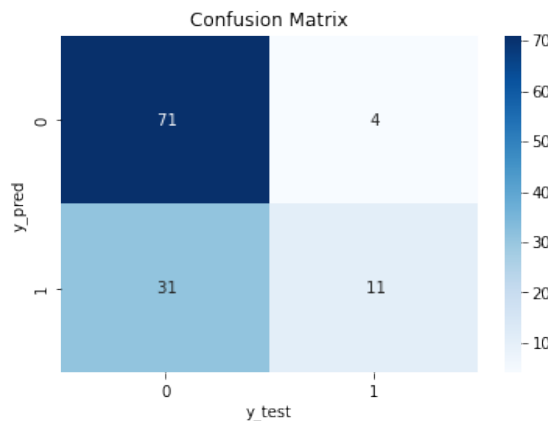


Gambar 7: Histogram Model Accuracy Uji Coba Pertama



Gambar 8: Histogram Model Accuracy Uji Coba Pertama

Berdasarkan *histogram* pada Gambar 7 dan Gambar 8 terdapat dua buah garis. Garis biru menunjukkan data *train* dan garis oranye menunjukkan data *test*. Berdasarkan model *accuracy* pada Gambar 7, proses *training* data kurang stabil pada setiap *epochnya* untuk mendapatkan akurasi karena perbandingan data yang digunakan antara data ancaman dan data bukan ancaman yang berbeda jauh. Berdasarkan *histogram model loss* pada Gambar 8 menunjukkan banyaknya *error* pada setiap *epochnya*. Pada diagram tersebut terlihat bahwa data test naik secara signifikan dan data train menunjukkan penurunan sehingga menunjukkan tanda bahwa proses *training* harus dihentikan lebih awal. Pada *confusion matrix* diperoleh hasil seperti pada Gambar 9 dan dijelaskan pada Tabel 2.



Gambar 9: Confusion Matrix Uji Coba Pertama

TABEL II
 CLASSIFICATION REPORT PERCOBAAN PERTAMA

	Precision	Recall	F1-Score	Support
0 (negatif)	0.70	0.95	0.80	75
1 (positif)	0.73	0.26	0.39	42
Accuracy			0.70	117
Macro avg	0.71	0.60	0.59	117
Weithted avg	0.71	0.70	0.65	117

Berdasarkan hasil yang diperoleh dari *Classification Report* pada Tabel 1, maka dilakukan uji coba yang kedua dengan mengurangi data dan menyeimbangkan data antara data ancaman dan data bukan ancaman. Data yang

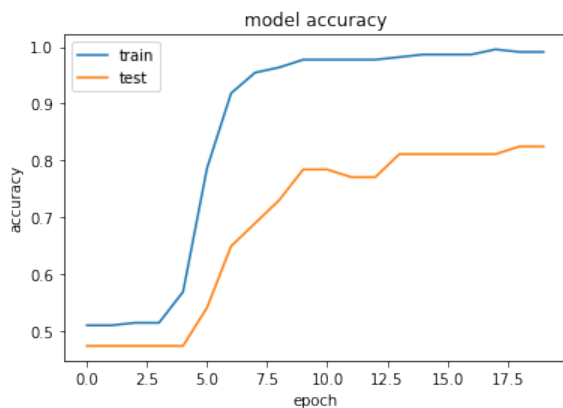
digunakan berjumlah 294 data yang kemudian akan dibagi menjadi 50% data ancaman dan 50% data bukan ancaman. Jumlah data yang digunakan adalah 147 data ancaman dan 147 data bukan ancaman.

Setelah itu dilakukan pengurangan *dense layer*, *convolutional filter* dan pengurangan *epoch* dalam membantu penanganan *overfitting* data. Epoch yang akan digunakan berjumlah 20 epoch. Arsitektur CNN model yang digunakan pada percobaan ini dapat dilihat pada Gambar 9. Dari percobaan tersebut, didapatkan hasil seperti pada Gambar 10 dan Gambar 11.

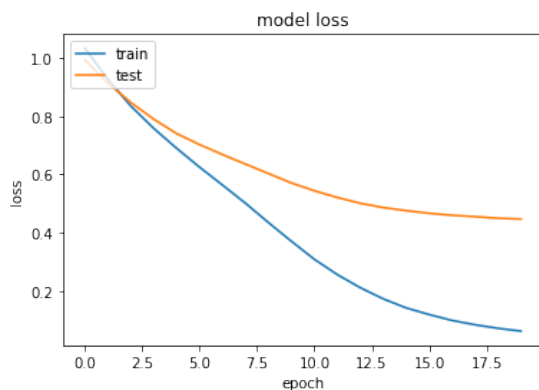
Layer (type)	Output Shape	Param #
input_3 (InputLayer)	[(None, 37)]	0
embedding_2 (Embedding)	(None, 37, 20)	20940
conv1d_2 (Conv1D)	(None, 35, 16)	976
global_max_pooling1d_2 (GlobalMaxPooling1D)	(None, 16)	0
flatten_1 (Flatten)	(None, 16)	0
dense_1 (Dense)	(None, 3)	51

 Total params: 21,967
 Trainable params: 21,967
 Non-trainable params: 0

Gambar 10: Arsitektur CNN Uji Coba Kedua

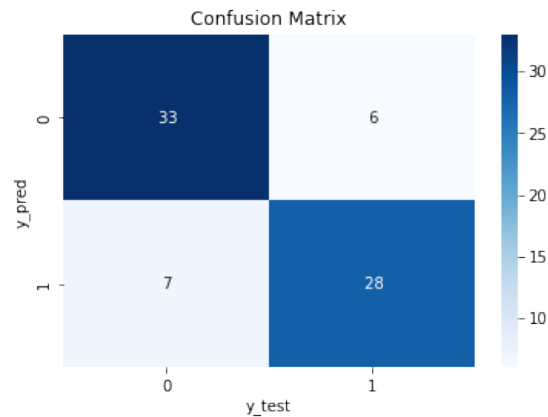


Gambar 11: Histogram Model Accuracy Uji Coba Kedua



Gambar 12: Histogram Model Accuracy Uji Coba Kedua

Seperti yang ditunjukkan oleh *histogram* pada Gambar 11 dan Gambar 12, *model accuracy* mengalami peningkatan yang stabil dengan menggunakan 20 *epoch*. Pada *model loss*, penggunaan data yang seimbang dan pengurangan *epoch* membuat peningkatan data *test* tidak naik secara perlahan dan tidak signifikan dibandingkan dengan data *train*. Sedangkan pada *confusion matrix* diperoleh hasil seperti pada Gambar 13 dan Tabel 3.



Gambar 13: Confusion Matrix Uji Coba Kedua

TABEL III
CLASSIFICATION REPORT PERCOBAAN KEDUA

	Precision	Recall	F1-Score	Support
0 (negatif)	0.82	0.85	0.84	39
1 (positif)	0.82	0.80	0.81	35
Accuracy			0.82	74
Macro avg	0.82	0.75	0.82	74
Weighted avg	0.82	0.74	0.82	74

Berdasarkan hasil dari *confusion matrix* beserta *classification report*, akurasi yang diperoleh melalui perhitungan *f1-score* meningkat hingga 82% dengan *precision* dan *recall* yang tinggi.

Hasil akhir dari model yang dibuat dalam penelitian ini akan digunakan sebagai prediksi dalam mendeteksi ancaman. Audio akan dilakukan konversi dari audio ke teks menggunakan *Google Text-To-Speech*. Setelah dilakukan konversi, maka akan diubah menjadi bentuk *token* sehingga dapat dilakukan prediksi dengan menggunakan model. Output dari prediksi ini berupa ancaman atau bukan ancaman.

V. KESIMPULAN

Berdasarkan hasil dari analisis dan uji coba yang dilakukan, terdapat beberapa kesimpulan. Pengumpulan data dapat digunakan menggunakan *Web Scraping* karena lebih fleksibel untuk digunakan. Setelah dilakukan *web scraping*, data perlu dilakukan *preprocessing* agar dapat diolah. Setelah diolah data dibagi menjadi *train* dan *test* sehingga dapat dibuat menjadi model. Pemodelan menggunakan metode *Convolutional Neural Network* yang cukup baik jika digunakan dalam mengolah teks. Berdasarkan hasil uji coba yang dilakukan, didapatkan akurasi 70% karena masalah keseimbangan data. Namun,

ketika dilakukan pengurangan data menjadi seimbang, dense layer, dan epoch didapatkan akurasi sebesar 82%.

DAFTAR PUSTAKA

- [1] Y. Kim, "Convolutional neural networks for sentence classification," *EMNLP 2014 - 2014 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, pp. 1746–1751, 2014, doi: 10.3115/v1/d14-1181.
- [2] N. A. Shaffira, "Klasifikasi Sentimen Ulasan Film Indonesia dengan Konversi Speech-to-Text (STT) Menggunakan Metode Convolutional Neural Network (CNN)," *J. SAINS DAN SENI ITS*, vol. 9, no. 1, pp. 95–101, 2020.
- [3] Y. Luan and S. Lin, "Research on Text Classification Based on CNN and LSTM," *Proc. 2019 IEEE Int. Conf. Artif. Intell. Comput. Appl. ICAICA 2019*, pp. 352–355, 2019, doi: 10.1109/ICAICA.2019.8873454.
- [4] P. M. Sosa and C. Yang, "Twitter Sentiment Analysis using combined LSTM-CNN Models Related papers Recent Trends in Deep Learning Based Natural Language Processing," 2017.
- [5] D. D. Ayani, H. S. Pratiwi, and H. Muhandi, "Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace," *J. Sist. dan Teknol. Inf.*, vol. 7, no. 4, pp. 257–262, Oct. 2019.
- [6] I. Putri Sonya, "Analisis Web Scraping Untuk Data Bencana Alam dengan Menggunakan Teknik Breadth-First Search Terhadap 3 Media Online," 2016.
- [7] A. V. Saurkar and S. A. Gode, "An Overview On Web Scraping Techniques And Tools," 2018. [Online]. Available: <http://www.ijfrcsce.org>.
- [8] I. C. Education, "Speech Recognition," *IBM Cloud Education*, 2020. <https://www.ibm.com/cloud/learn/speech-recognition> (accessed Jan. 30, 2022).
- [9] D. Amos, "The Ultimate Guide To Speech Recognition With Python," *Real Python*. <https://realpython.com/python-speech-recognition/#how-speech-recognition-works-an-overview> (accessed Jan. 30, 2022).
- [10] N. P. Durette, "gTTS (Google Text-to-Speech), a Python library and CLI tool to interface with Google Translate text-to-speech API," *pypi.org*, 2021. <https://pypi.org/project/gTTS/> (accessed Feb. 07, 2022).
- [11] S. Saha, "A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way," *Towards Data Science*, 2018. [https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53#:~:text=A Convolutional Neural Network \(ConvNet,differentiate one from the other.](https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53#:~:text=A Convolutional Neural Network (ConvNet,differentiate one from the other.) (accessed Jan. 27, 2021).
- [12] S. Liao, J. Wang, R. Yu, K. Sato, and Z. Cheng, "CNN for situations understanding based on sentiment analysis of twitter data," in *Procedia Computer Science*, 2017, vol. 111, pp. 376–381, doi: 10.1016/j.procs.2017.06.037.
- [13] Z. Wang and Z. Qu, "Research on Web text classification algorithm based on improved CNN and SVM," *Int. Conf. Commun. Technol. Proceedings, ICCT*, vol. 2017-Octob, pp. 1958–1961, 2018, doi: 10.1109/ICCT.2017.8359971.
- [14] D. Yogish, T. N. Manjunath, and R. S. Hegadi, *Review on Natural Language Processing Trends and Techniques Using NLTK*, vol. 1037. Springer Singapore, 2019.
- [15] H. Rohayani, H. L. H. Spits Warnars, M. Tuga, and E. Abdurahman, "Employing Vector Autoregression Feedforward Neural Network With Particle Swarm Optimization In Wind Speed Forecasting," *Sylwan*, vol. 164, no. 3, pp. 1–9, 2020.