

Analisis Sentimen terhadap Radikalisme berdasarkan Opini Masyarakat di media sosial Twitter

Fail Amir

Fakultas ilmu komputer jurusan Informatika , universitas amikom Yogyakarta
E-mail: fail.amir@students.amikom.ac.id

Abstract

Paham radikalisme memiliki banyak pengaruh negatif terhadap persatuan negara Indonesia, bahkan dapat menghancurkan NKRI yang berpedoman pada Pancasila. Organisasi yang berpaham radikal memiliki sifat antoleransi yaitu tidak menerima golongan lain yang tidak sepaham dengannya. Organisasi ini mencekoki masyarakat Indonesia dengan dogma-dogma yang mengatas-namakan agama, agar tujuan mereka yang ingin menjadikan negara Indonesia sebagai negara khilafah dapat terwujud. Seiring dengan majunya teknologi pada masa modern ini, media sosial menjadi alternatif baru untuk menebarkan paham yang bisa mengancam kesatuan NKRI ini. Oleh sebab itu, perlunya analisis sentimen terhadap suatu postingan seseorang yang terduga sebagai konten yang mengandung ajaran radikalisme, Salah satu media sosial yang masih aktif dan banyak digunakan dalam penyebaran paham ini adalah twitter. Pada penelitian ini dilakukan pengolahan data menggunakan text mining sebagai awal memperoleh data numerik dari data dokumen yang nantinya akan diklasifikasi menggunakan model ANN dengan metode backpropagation yang sudah diperoleh. Model ANN yang diperoleh terbilang cukup baik dengan akurasi sebesar 77%, presisi sebesar 77,1%, recall sebesar 73,4%, dan f-measure 75,25%. Analisis hasil menggunakan graph wordcloud dan graph network menunjukkan kecenderungan kategori sentimen opini tergantung pada semboyan dan kata yang digunakan pada tweet.

Keywords: islam radikal, twitter, text mining, ANN, backpropagation

1. Pendahuluan

Pada pertengahan tahun 2017 lalu, negara Indonesia disibukkan dengan berita pembubaran organisasi masyarakat (organisasi masyarakat) yang dinilai menentang kedaulatan negara Indonesia yaitu Pancasila. Berita ini tersebar setelah dikeluarkannya perubahan atas UU Nomor 17 Tahun 2013 Tentang Organisasi Kemasyarakatan (Peraturan Pemerintah, 2017). Penentangan yang dilakukan beberapa organisasi masyarakat islam radikal ini berupa gerakan dengan tujuan menjadikan Indonesia sebagai negara khilafah. Organisasi ini memiliki pemahaman terhadap agama Islam yang condong ke kanan (radikal) tanpa memahami sejarah bermulanya persatuan Indonesia yang tercantumkan dalam konstitusi NKRI yaitu UUD 1945.

Paham radikal yang dianut oleh sebagian organisasi masyarakat di Indonesia dapat mengancam kedudukan Pancasila sebagai ideologi bangsa (Khamid, 2016). Paham ini mengganti isi ideologi dengan dogma-dogma yang mengatasnamakan agama sebagai dasar pembentukan ideologi dengan paham radikal. Akibatnya, masyarakat yang tergabung pada organisasi masyarakat ini jauh dari rasa nasionalisme terhadap bangsanya

sendiri. Pada kehidupan sosialnya pun paham radikalisme kerap kali mewujudkan perilaku ekstrem, intoleransi, dan destruktif dalam merealisasikan tujuannya (Agung & Santoso, 2018).

2. Tinjauan Pustaka

Paham radikalisme yang dianut oleh sebagian organisasi masyarakat di Indonesia dapat mengancam kedudukan Pancasila sebagai ideologi bangsa (Khamid, 2016).

Pada kehidupan sosialnya pun paham radikalisme kerap kali mewujudkan perilaku ekstrem, intoleransi, dan destruktif dalam merealisasikan tujuannya (Agung & Santoso, 2018).

Di antaranya terdapat penelitian yang dilakukan oleh Quanzhi Li dan tim. Dalam penelitiannya, Quanzhi Li dan tim menggunakan metode text feature selection untuk mendukung analisis sentimennya (Li, et al., 2016).

Seperti penelitian yang dilakukan oleh Novitasari yang mengklasifikasikan sinyal egg berupa data gelombang dengan memadukan metode fuzzy dan metode modifikasi dari backpropagation (Novitasari, 2015). ANN juga dapat digunakan dalam klasifikasi image, seperti penelitian yang dilakukan oleh Leavline.

3. Dasar Teori

Klasifikasi ANN dapat dikatakan sebagai pengamatan terhadap sub populasi baru terhadap suatu set pelatihan data lama yang sudah dilakukan pengamatan sebelumnya. Data yang digunakan merupakan data non linear karena dapat menghasilkan pemetaan yang sesuai antara input dan output (Naik, et al., 2015). Dalam melakukan klasifikasi digunakan algoritma berbasis komputasi agar implementasi dari proses identifikasi dilakukan dengan cepat dan praktis. Pada dasarnya klasifikasi ANN dilakukan untuk pengenalan pola yang dilatih untuk identifikasi masalah yang diberikan dengan tahap 2 fase yaitu training dan testing. Selama proses training yang terjadi, jaringan dilatih untuk mengklasifikasi data berdasarkan model yang sudah dibangun selama pelatihan agar menghasilkan error 0% pada tahap validasinya (Li, 2017).

Klasifikasi menggunakan ANN memerlukan 3 tahapan utama yaitu input layer sebagai sumber data masukan, hidden layer sebagai prosesnya dan output layer sebagai hasil akhir dari klasifikasinya

Input layer merupakan bagian masukan data yang akan diklasifikasi untuk mendapatkan pola dan model. Lapisan hidden layer merupakan lapisan tersembunyi yang menerima data dari input, sedangkan besar kecilnya nilai input tergantung pada perkalian nilai bobotnya. Nilai bobot data yang dihasilkan akan diteruskan pada tahap output layer. Prinsip kerja pada output layer ini, hampir sama dengan prinsip hidden layer yaitu menggunakan fungsi sigmoid sebelum hasil klasifikasi keluar. Keberhasilan penerapan ANN membutuhkan kehati-hatian dalam memilih parameter yang sesuai dan signifikansi variasinya tergantung pada permasalahan dan subjek yang berbeda (Bashashati, et al., 2015).

2.8.1 Arsitektur Artificial Neural Network

Artificial neural network memiliki beberapa macam algoritma yang berbeda yang digunakan dalam menyelesaikan permasalahan atau prediksi pada suatu penelitian (Fauset, 1994) diantaranya:

a) Jaringan Tunggal

Jaringan tunggal hanya memiliki satu lapisan yang terdiri dari 1 layer input, 1 layer bobot koneksi, dan 1 layer output. Neuron pada jaringan tunggal ini terhubung lurus dengan

sekumpulan keluaran. Model yang menganut jaringan tunggal adalah ADALINE, perceptron, hopfield, dan LVQ.

b) Jaringan Multi Layer

Jaringan multi memiliki 3 tahapan yaitu input layer, hidden layer, dan output layer yang tampak seperti gambar Gambar 2.3. Jaringan multi ini biasanya lebih banyak digunakan karena dapat menyelesaikan permasalahan yang bersifat kompleks, meskipun dalam masa training membutuhkan waktu yang cukup lama dikarenakan banyaknya perhitungan node pada setiap layer-nya.

2.8.2 Backpropagation

Backpropagation terkenal dengan metode yang memiliki pelatihan dengan tipe arsitek multi layer dengan akurasi dan fleksibilitas yang baik dibandingkan dengan metode jaringan saraf tiruan lainnya, seperti algoritma perceptron (Angkoso, et al., 2018). Backpropagation merupakan salah satu jenis jaringan saraf tiruan ANN yang memiliki sepasang pola, yaitu pola masukan dan pola yang diinginkan. Backpropagation akan melatih jaringan agar seimbang dalam mengenali pola yang digunakan selama masa pelatihan. Metode ini mampu menangani pola-pola kompleks untuk mendapatkan model yang optimal (Effendi, 2013).

Berbeda dengan jaringan saraf tiruan tunggal, backpropagation menambahkan

hidden layer dalam melatih jaringan untuk memperoleh output yang sesuai. Untuk melatih jaringan backpropagation harus melalui empat fase, yaitu inisialisasi bobot, feedforward, backpropagation error, dan penyesuaian bobot. Fase inisialisasi bobot dilakukan secara random sesuai dengan rentang kisaran angka yang dikehendaki. Fase feedforward adalah fase dimana dari pola masukan hingga keluaran dihitung maju menggunakan fungsi aktivasi. Fase ketiga merupakan fase dimana pengulangan perhitungan yang dihitung mundur. Perhitungan mundur inilah menjadikan adanya selisih hasil dari outputnya. Selisih dari hasil output ini dikatakan sebagai backpropagation error. Fase terakhir merupakan fase penyesuaian bobot yang bertujuan untuk mengurangi error sebelumnya (Siang, 2005).

Pada dasarnya, backpropagation merupakan metode pengembangan dari feedforward. Metode ini sangat memperhatikan selisih atau error dalam masalah analisis regresi dengan square error sebagai fungsi kesalahan yang dihasilkan antara target yang sebenarnya dengan keluaran unit output dari network seperti pada Persamaan (2.4).

$$(e_j)^2 = (t_j - o_j)^2 \quad (2.4)$$

Sama dengan metode klasifikasi lainnya, backpropagation melakukan tahap training dan testing dalam prosesnya. Pada tahap training, keseluruhan error untuk semua unit output dideskripsikan dengan Persamaan (2.5).

4. Implementasi dan Pembahasan

kode program

```
import tweepy
from textblob import TextBlob
```

```

import csv
ulang = 10
pos = 0
neg = 0

# Step 1 - Authenticate
consumer_key= 'he8TzJJectSQsrQ3gC09XQuH3'
consumer_secret= 'zMn8uScwocSxiBdWZo6rBKARoUa6wHcky71XFh86PTlYasiH8c'

access_token='2351884572-UhucAenFEvtznzPLHz4czNrRah5lStNEomN83nzB'
access_token_secret='wblXNMdAK5kUbvhaiETSBmjbzztZb01zlWlsVaqmItY9K'

auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)

api = tweepy.API(auth)

# Open/create a file to append data to
csvFile = open('result.csv', 'a')

#Use csv writer
csvWriter = csv.writer(csvFile)

for tweet in tweepy.Cursor(api.search,
                            q = "radikal",
                            since = "2018-02-14",
                            until = "2019-01-8",
                            count = 100,
                            lang = "en").items():

    # Write a row to the CSV file. I use encode UTF-8
    csvWriter.writerow([tweet.created_at, tweet.text.encode('utf-8')])
    print tweet.created_at, tweet.text
csvFile.close()

#Step 2 - Retrieve Tweets
public_tweets = api.search(q=['radikal', 'islam'], count=100)

all_polarity = 0
for tweet in public_tweets:
    print(tweet.text)
    analysis = TextBlob(tweet.text)
    an = analysis.translate(from_lang='id', to='en')
    print(an.sentiment)

```

```

    if (an.polarity > 0.5):
        pos += an.polarity
    if (an.polarity < 0.5):
        neg += an.polarity
    all_polarity += an.polarity

    # kelsi = (an.sentiment)
    # # buka file untuk ditulis
    # file_bio = open("catatan.txt", "a")
    # # tulis teks ke file
    # file_bio.write(kelsi)
    # # tutup file
    # file_bio.close()

# tentukan lokasi file, nama file, dan inisialisasi csv
#f = open('dataset.csv', 'w')
#w = csv.writer(f)
#w.writerow(isn)

# menulis file csv
#for s in rang
#e(ulangi):
#    w.writerow(s)

# menutup file csv
#f.close()

print("")
print("total positif :")
print pos

print("")
print("total negatif :")
print neg

isn = (pos + neg) / 2

print("")
print("rata-rata :")
print isn

print("")

```

```

#Step 4 - To Check Positive or Negative Sentiment from Overall Tweets
if (all_polarity/100 > 0):
    print(all_polarity/100)
    print("")
    print("kecenderungan : ")
    print('Positive')
else:
    print(all_polarity/100)
    print("")
    print("kecenderungan : ")
    print("Negative")

# import matplotlib.pyplot as plt
# import numpy as np

# x = np.linspace(0, 10, 100)

# plt.plot(x, np.sin(x))
# plt.plot(x, np.cos(x))

# plt.show()

```

4.1 Crawling Data

Data yang diperoleh sebanyak 1478 tweet, dengan pengambilan paling banyak 50 tweet per harinya. Data tweet yang diperoleh akan dianalisis sentimennya dengan empat kategori yaitu, positif, negatif, netral, dan outlier. Contoh data tweet yang diperoleh dari hasil crawling data dari twitter seperti pada Tabel 4.1

Tabel 4.1 Sampel dari Hasil Crawling Data Tweet

Pada Tabel 4.1 terlihat bahwa hasil crawling yang diperoleh mendapatkan empat informasi yang terdiri dari D, R, F, dan text. Dengan D sebagai tanggal akses data, R sebagai jumlah retweet dari tweet yang diunggah, F sebagai like dari pembaca terhadap tweet, dan text merupakan isi dari tweet yang diunggah.

4.1 Pelabelan Data

Langkah selanjutnya setelah melakukan crawling adalah melakukan pelabelan terhadap data tweet yang sudah diperoleh. Pelabelan dilakukan secara manual dengan empat kategori dengan ketentuan berikut:

- 1) Positif = tweet yang berisi dukungan terhadap ideologi pancasila dan ketidaksetujuan terhadap pembentukan ideologi islam radikal
- 2) Negatif = tweet yang berisi dukungan terhadap perubahan dan pembentukan terhadap ideologi baru dan setuju terhadap pembentukan ideologi islam radikal
- 3) Netral = tweet yang sentimennya tidak dikeduanya atau berada dikeduanya

- 4) Outlier = tweet yang berisi sekedar informasi, diluar konteks tema permasalahan, dan menggunakan bahasa lain selain Bahasa Indonesia.

4.3 Pre-processing

Langkah awal dari pengolahan data merupakan tahap pra-proses. Tahap ini penting dilakukan sebelum melakukan ekstraksi dan seleksi fitur agar mendapat nilai dari masing-masing kata dari dokumen yang dimiliki. Pada tahap ini dilakukan

proses cleansing, case folding, penghapusan stopwords, dan tokenisasi.

a) Cleansing

Proses cleansing adalah proses pembersihan kalimat. Pembersihan dilakukan dengan maksud menghapus semua karakter web dan html. Pada proses ini juga menghapus angka, tanda baca, emoticon, dan hashtag pada kalimat. Penghapusan ini dilakukan karena karakter dan tanda baca tersebut tidak dibutuhkan dan dapat mengganggu nilai term pada saat proses klasifikasi. Output dilihat pada Tabel 4.4.

Tabel 4.4 Sampel Data Sesudah Tahap Cleansing

No	Label kelas	Text	Setelah dilakukan cleansing data
1	Positif	Survei: 79% WNI tidak Setuju NKRI Diganti Menjadi Khilafah	Survei WNI tidak Setuju NKRI Diganti Menjadi Khilafah
2	Negatif	Memperjuangkan Khilafah itu adalah kewajiban #khilafahajaranislam #ReturnTheKhilafa	Memperjuangkan Khilafah itu adalah kewajiban hilafahajaranislam ReturnTheKhilafa
5	Negatif	Jejak Khilafah di Nusantara #ReturnTheKhilafah	Jejak Khilafah di Nusantara ReturnTheKhilafah

b) Case folding

Tahap selanjutnya merupakan tahap case folding. Pada tahap ini dilakukan perubahan semua huruf kapital pada kalimat dari A-Z dijadikan huruf kecil semua (lower case), dan menghasilkan output yang dapat dilihat pada Tabel 4.5

Tabel 4.5 Sampel Data dengan Sesudah Tahap Case Folding

No	Label kelas	Text	Setelah dilakukan case folding
1	Positif	Survei: 79% WNI tidak Setuju NKRI Diganti Menjadi Khilafah	survei wni tidak setuju nkri diganti menjadi khilafah
2	Negatif	Memperjuangkan Khilafah itu adalah kewajiban #khilafahajaranislam #ReturnTheKhilafa	memperjuangkan khilafah itu adalah kewajiban hilafahajaranislam returnthekhilafah
5	Negatif	Jejak Khilafah di Nusantara #ReturnTheKhilafah	jejak khilafah di nusantara returnthekhilafah

c) Penghapusan stopwords

Tahap berikutnya merupakan tahap penghapusan stopwords. Pada tahap ini semua kata pada kalimat yang mengandung kata stopwords akan otomatis dihapus. Kata “tidak” digabungkan dengan kata setelahnya dengan tanda penghubung (_). Penggabungan kata ini dilakukan untuk membedakan kata negatif yang diikuti kata “tidak” dengan kata positif dari kata itu sendiri, sehingga perolehan akhir dari pra- proses terlihat pada Tabel 4.6.

Tabel 4.6 Sampel Data Sesudah Tahap Stopword

No	Label kelas	Text	Setelah dilakukan penghapusan stopwords
1	Positif	Survei: 79% WNI tidak Setuju NKRI Diganti Menjadi Khilafah	surveiwni tidak_setuju nkri
2	Negatif	Memperjuangkan Khilafah itu adalah kewajiban #khilafahajaranislam #ReturnTheKhilafa	memperjuangkan kewajiban returnthekhilafah
5	Negatif	Jejak Khilafah di Nusantara #ReturnTheKhilafah	jejak nusantara returnthekhilafah

d) Tokenisasi

Proses tokenisasi merupakan proses dilakukannya partisi terhadap kalimat menjadi perkata. Tokenisasi dilakukan agar mempermudah proses pembobotan pada masing-masing kata yang menjadi term seperti pada Tabel 4.7.

Tabel 4.7 Sampel Data Sesudah Tahap Tokenisasi

No	Label kelas	Text	Setelah dilakukan tahap tokenisasi
1	Positif	Survei: 79% WNI tidak Setuju NKRI Diganti Menjadi Khilafah	survei - wni - tidak_setuju - nkri
2	Negatif	Memperjuangkan Khilafah itu adalah kewajiban #khilafahajaranislam #ReturnTheKhilafa	memperjuangkan - kewajiban – returnthekhilafah
5	Negatif	Jejak Khilafah di Nusantara #ReturnTheKhilafah	jejak - nusantara -returnthekhilafah

4.4 Pembobotan TF IDF

Tahap selanjutnya ialah melakukan pembobotan kata terhadap term-term yang sudah diperoleh dari tahap tokenisasi terhadap dokumennya. Tahap ini dimulai dengan mencari nilai TF term dari banyaknya dokumen atau data. TF merupakan banyaknya kata term yang meuncil pada masing-masing dokumen seperti pada Tabel 4.8.

Tabel 4.8 Pembobotan TF

	Do	kumen 1	Dokumen 2	Dokumen 3
Term	Survei	1	0	0
	Wni	1	0	0

tidak_setuju	1	0	0	
Nkri	1	0	0	
kewajiban	0	1	0	
memperjuangkan		0	1	0
jejak	0	0	1	
nusantara	0	0	1	
returnthekhilafah		0	1	1

Setelah diketahui nilai TF, selanjutnya menghitung nilai DF dan IDF. DF merupakan jumlah munculnya term pada dokumen, sedangkan IDF merupakan perhitungan dari Persamaan (2.2) yang akan menghasilkan Tabel 4.9.

Tabel 4.9 Pembobotan IDF

Term	Df	IDF
survei	1	$\text{Log}(3/1) = 0,477$
Wni	1	$\text{Log}(3/1) = 0,477$
tidak_setuju	1	$\text{Log}(3/1) = 0,477$
Nkri	1	$\text{Log}(3/1) = 0,477$
kewajiban	1	$\text{Log}(3/1) = 0,477$
memperjuangkan	1	$\text{Log}(3/1) = 0,477$
jejak	1	$\text{Log}(3/1) = 0,477$
nusantara	1	$\text{Log}(3/1) = 0,477$
returnthekhilafah	2	$\text{Log}(3/2) = 0,176$

Dengan diperolehnya bobot TF dan IDF, selanjutnya dapat diperoleh bobot TF IDF dengan mengalikan bobot TF masing-masing dokumen dengan bobot IDF, maka akan menghasilkan bobot akhir

4.5 Klasifikasi ANN

Bobot yang sudah didapatkan dari perhitungan TF IDF selanjutnya akan dilakukan pengujian klasifikasi menggunakan ANN dengan metode backpropagation. Berdasarkan Tabel 4.10, didapatkan sampel data yang terdapat pada Tabel 4.11.

Tabel 4.11 Sampel Bobot Hasil TF IDF untuk Klasifikasi ANN

Keterangan:

(1) = positif

(0) = negatif

= inputan

dokumen = dokumen (tweets) input

Berdasarkan Tabel 4.11, dapat dilakukan klasifikasi backpropagation secara manual, sebagai gambaran dalam perhitungan klasifikasi menggunakan program Rstudio. Dengan inputan X1 sampai X9, digunakan satu hidden layer dengan 2 node. Terdapat sinapsis

dengan bobot () yang terhubung dari node input ke node hidden. Kemudian neuron akan diteruskan menuju daerah output layer dengan 1 node. Terdapat bias yang terhubung dengan node-node hidden dan node output, yang dapat dilihat pada Gambar 4.1.

Gambar 4.1 Arsitektur Klasifikasi untuk Sampel Data

Gambar 4.1 menunjukkan arsitektur ANN dari sampel data yang digunakan, arsitektur ANN yang diperoleh dijadikan sebagai langkah awal untuk perhitungan

manual keempat langkah yaitu inisialisasi bobot, feedforward, backpropagation,

dan penyesuaian bobot.

a) Inisialisasi bobot

Misalkan mengambil 1 data yaitu dokumen1 pada Tabel 4.11 untuk hitung manual. Misalkan bias pada node hidden dan output adalah 1 dengan $1 = 0,477$,

$2 = 0,477$, $3 = 0,477$, $4 = 0,477$, $5 = 0$, $6 = 0$, $7 = 0$, $8 = 0$, $9 = 0$, dan $t = 1$,

seperti pada Tabel 4.12 dan 4.13 berikut:

Tabel 4.12 Inisialisasi Bobot Input Iterasi Pertama

	1 (hidden)	2 (hidden)
1	-0,09	-0,08
2	-0,07	-0,06
3	-0,05	-0,04
4	-0,03	-0,02
5	-0,01	0,3
6	0,01	0,02
7	0,03	0,04
8	0,05	0,06
9	0,07	0,08
1	0,89	-0,1

Tabel 4.13 Inisialisasi Bobot Output Iterasi Pertama

b) Feedforward

Langkah selanjutnya adalah menghitung sinyal input dengan Persamaan (2.36) yang akan menjadi nilai bobot yang menghubungkan node input ke node hidden

$$= 0 + \sum$$

$$1 = 0,89 + 0,477(-0,09) + 0,477(-0,07) + 0,477(-0,05) + 0,477(-0,03) + 0(-0,01)$$

$$+ 0(0,01) + 0(0,03) + 0(0,05) + 0(0,07) = 0,77752$$

$$2 = -0,1 + 0,477(-0,08) + 0,477(-0,06) + 0,477(-0,04) + 0,477(-0,02) + 0(0,3) + 0(0,02) + 0(0,04) + 0(0,06) + 0(0,08) = -0,1954$$

Kemudian menerapkan fungsi aktivasi, yaitu fungsi sigmoid kepada masing-masing sinyal input,

sedangkan untuk nilai output sementara dapat diketahui dengan menghitung unit

terapkan fungsi aktivasi kembali untuk hasil perhitungan unit keluaran yang sudah diperoleh.

$$= \frac{1}{1 + 0,62408}$$

$$= 0,3488532$$

Nilai keluaran yang didapat untuk hanya sebesar 0,3488532, sedangkan nilai

yang diinginkan adalah 1. Maka dengan ini, proses dilanjutkan dengan tahap backpropagation untuk mencari nilai faktor kesalahan dan perubahan bobot untuk iterasi selanjutnya agar nilai mendekati .

c) Backpropagation

Langkah selanjutnya melakukan perhitungan factor sebagai unit keluarannya

$$= (t -) () (1 -)$$

$$= (1 - 0,3488532) (0,3488532) (1 - 0,3488532)$$

$$= 0,147911$$

Hasil faktor yang sudah diperoleh digunakan untuk menghitung perubahan bobot keluaran dengan nilai learning rate () = 0,2

$$\Delta =$$

$$\Delta_{01} = 0,2 (-0,147911) (1) = 0,029582$$

$$\Delta_{11} = 0,2 (0,147911) (0,684714) = 0,020255$$

$$\Delta_{21} = 0,2 (0,147911) (0,451305) = 0,013351$$

dengan menggunakan perjalanan rumus dan proses yang sama, faktor kesalahan bobot dari unit tersembunyi juga dapat diketahui.

$$=$$

$$1 = (0,147911) (-0,3) = -0,04437$$

$$2 = (0,147911) (-0,13) = -0,00958$$

$$= (\quad) (\quad) (1 - \quad)$$

$$1 = -0,04437 (0,684714) (1-0,684714) = -0,00958$$

$$2 = -0,027 (0,451305) (1-0,451305) = -0,00476$$

Faktor yang diperoleh masing-masing inputan pada hidden tersembunyi dapat diketahui perubahan bobotnya seperti perhitungan berikut.

$$\Delta \quad =$$

$$\Delta = 0,2 (-0,00958) (1) = -0,00191587$$

Perhitungan perubahan bobot pada sel Δ_{02} , Δ_{03} , dan seterusnya sama, sehingga diperolehlah perubahan bobot yang dapat dilihat pada Tabel 4.14.

	Z1 (hidden)	Z2 (hidden)
1	-0,00091	-0,00045
2	-0,00091	-0,00045
3	-0,00091	-0,00045
4	-0,00091	-0,00045
5	0	0
6	0	0
7	0	0
8	0	0
9	0	0
1	-0,00191587	-0,00095

d) Penyesuaian bobot

Langkah terakhir adalah mendapatkan bobot baru unit keluaran:

$$(\text{baru}) = (\text{lama}) + \Delta$$

$$10(\text{baru}) = 10(\text{lama}) + \Delta_{10} = -0,36 + 0,029582 = -0,334178$$

$$11(\text{baru}) = 11(\text{lama}) + \Delta_{11} = -0,3 + 0,020255 = -0,27974$$

$$12(\text{baru}) = 12(\text{lama}) + \Delta_{12} = -0,13 + 0,013351 = -0,3436649408$$

$$(\text{baru}) = (\text{lama}) + \Delta$$

$$01(\text{baru}) = 01(\text{lama}) + \Delta_{01} = 0,89 - 0,00191587 = 0,888084$$

Perhitungan update bobot pada sel 02(), 03(), dan seterusnya sama, sehingga diperoleh Tabel 4.16 dan Tabel 4.17 sebagai update bobot baru untuk iterasi selanjutnya.

	Z1 (hidden)	Z2 (hidden)
X1	-0,09091	-0,08045
X2	-0,07091	-0,06045
X3	-0,5091	-0,04045
X4	-0,03091	-0,02045
X5	-0,01 0,3	
X6	0,01 0,02	
X7	0,03 0,04	
X8	0,05 0,06	
X9	0,07 0,08	
1	0,0888084	-0,10095

Proses ini akan terus berlanjut hingga sebanyak iterasi yang diinginkan dan nilai

mendekati . Untuk data penuh yang dimiliki, diaplikasikan menggunakan program Rstudio untuk mempermudah perhitungannya. Data yang dimiliki sebanyak 1141 dokumen dengan penilaian 1 untuk berlabel positif dan 0 untuk berlabel negatif. Setelah diproses dengan menggunakan TF IDF untuk mendapatkan bobot setiap term, dilakukan seleksi fitur untuk variabel atau term yang dinilai kurang mempengaruhi dalam proses klasifikasi. Term yang dihilangkan merupakan term yang hanya memiliki bobot dari setiap pencocokan dokumen kurang dari 2 kali kemunculan. Term yang dihilangkan diantaranya adalah kata “harem”, “islamtaubah”, “kertas”, “tingkat”, “kasasi”, dan banyak lagi. Hingga term yang dimiliki sebanyak 579 sebagai variabelnya dari 1141 dokumen. Data inilah yang kemudian dilanjutkan ke proses klasifikasi.

4.6 Hasil Pengujian Model dan Akurasi

Proses klasifikasi menggunakan backpropagation memunculkan output berupa bobot-bobot pada setiap lapisan, nilai sum square error, dan maksimal iterasi pada saat proses training dilakukan untuk setiap proses pengelompokan data tweet terhadap sentimennya. Iterasi berhenti ketika nilai MSE training mendekati threshold yang sudah ditentukan yaitu 0,01. Proses training dan testing terbagi menjadi dua grup yaitu:

- Pembagian data 80% untuk training dan 20% untuk testing
- Pembagian data 60% untuk training dan 40% untuk testing

Masing-masing dari setiap proses training, baik pada grup pertama maupun grup kedua memasukan batas iterasi maksimum sebanyak 100000 kali iterasi. Selain itu uji coba dilakukan pada jumlah hidden dan node hidden beserta perubahan learning rate untuk mencari model terbaik dengan nilai akurasi yang paling tinggi.

4.6.1 Pengujian Jumlah Node pada Hidden layer

Jumlah node hidden pada setiap layer berpengaruh terhadap perolehan akurasi saat melakukan klasifikasi data tweet terhadap sentimennya. Uji coba dilakukan dengan

menggunakan 1 hidden layer dengan jumlah node yang diuji coba sebanyak 1 sampai 10 node hidden. Model yang dipilih merupakan model dengan 1 hidden dan jumlah node yang memiliki nilai akurasi paling tinggi. Kemudian jumlah node pada model yang terpilih digunakan pada uji coba klasifikasi saat menggunakan 2 hidden layer.

a) Pembagian data 80% untuk training dan 20% untuk testing

Tabel 4.17 Uji Coba Jumlah Node pada Hidden layer pada Grup Pertama

Node	Jumlah node	Maksimal Iterasi	Akurasi
------	-------------	------------------	---------

1 hidden layer 1	333	0,73245
2	785	0,73245
3	6396	0,75438
4	1187	0,76754
5	3139	0,69736
6	1528	0,72808
7	4533	0,74561
8	2924	0,71053
9	1029	0,71054
10	2845	0,70615

2 hidden layer (4,2)	31281	0,77200
(4,5)	11660	0,73245
(4,10)	4289	0,73245
(4,50)	3697	0,74561

Ketika menggunakan 1 hidden layer, terpilih jumlah node hidden sebanyak 4 node karena memiliki nilai akurasi paling tinggi yaitu sebesar 76%. Jumlah node ini akan dijadikan sebagai jumlah node pada hidden layer pertama ketika uji coba

menggunakan 2 hidden layer. Pada Tabel 4.17 didapatkan model terpilih dengan 2 hidden layer, 4 node pada hidden pertama, dan 2 node pada hidden kedua. Akurasi yang diperoleh sebesar 77%.

b) Pembagian data 60% untuk training dan 40% untuk testing

Uji coba klasifikasi pada grup kedua (60% training, 40% testing), sama dengan proses yang dialami uji coba grup pertama. Hasil uji coba jumlah hidden dan jumlah node hidden seperti pada Tabel 4.18.

Tabel 4.18 Uji Coba Jumlah Node pada Hidden layer pada Grup Kedua

Node	Jumlah node	Maksimal Iterasi	Akurasi
------	-------------	------------------	---------

1 hidden layer 1	325	0,72369
2	235	0,71942

3	1960	0,73027
4	1130	0,72810
5	1246	0,72936
6	550	0,71701
7	1822	0,71930
8	1150	0,72149
9	882	0,72588
10	641	0,70615

2 hidden layer (3,2)	397	0,72365
(3,5)	118	0,72365
(3,10)	922	0,72587
(3,50)	558	0,71492

Berdasarkan hasil uji coba model pada Tabel 4.18, ketika menggunakan 1 hidden layer, terpilih jumlah node hidden sebanyak 3 node dengan akurasi sebesar 73%. Jumlah node ini akan dijadikan jumlah node pada hidden layer pertama ketika uji coba menggunakan 2 hidden layer. Namun dari uji coba menggunakan 2 hidden layer hanya memiliki akurasi tertinggi 72% dengan 3 node pada layer pertama dan 10 node pada layer kedua (0,725878 < 0,73027). Model yang diambil tetap 1 hidden layer dengan 3 node hidden.

4.6.2 Uji Coba Learning Rate

Uji coba jumlah hidden dan node hidden mendapatkan model yaitu:

- a) Grup pertama = 2 hidden layer (4,2)
- b) Grup kedua = 1 hidden layer (3)

Setelah mendapatkan model untuk kedua uji coba tersebut, selanjutnya dilakukan uji coba learning rate untuk melengkapi model terbaik yang terlihat pada Tabel 4.19. Model yang dipilih adalah model dengan learning rate yang memiliki nilai akurasi paling tinggi.

Tabel 4.19 Uji Coba Learning Rate

Grup Pertama

2 Hidden layer (4,2) Grup Kedua

1 Hidden layer (3)

Learning rate	Maks. Iterasi	Akurasi	Maks. Iterasi	Akurasi
0,1	8764	0,77200	347	0,71711
0,2	32612	0,75439	1432	0,71711
0,3	27802	0,76316	833	0,77149
0,4	5749	0,73246	4014	0,70900
0,5	42534	0,75877	836	0,72200

4.6.3 Penentuan Model

Berdasarkan hasil uji coba learning rate pada Tabel 4.19, didapatkan model terbaik untuk masalah klasifikasi data tweet terhadap sentimen menggunakan backpropagation adalah:

- a) Pembagian training dan testing: 80%, 20%
- b) Jumlah input: 579 node
- c) Jumlah hidden layer: 2 hidden (4,2)
- d) Learning rate: 0,1

Model tersebut memiliki confusion matrix dari proses testing seperti berikut.

Tabel 4.20 Confusion Matrix

Aktual			
Prediksi	Positif	Negatif	Total Baris
Positif	126	37	163
Negatif	15	50	141
Total Kolom	65	163	228

Berdasarkan Tabel 4.20, performansi sistem yang pertama bisa dinilai melalui perhitungan akurasi, presisi, recall, dan f-measure.

$$= \frac{126}{228} \cdot 100\% = 55,26\%$$

$$= \frac{126}{126+37} \cdot 100\% = 77,3\%$$

$$= \frac{50}{50+15} \cdot 100\% = 76,9\%$$

$$\begin{aligned}
 &= \frac{77,3 + 76,9}{2} \% \\
 &= 77,1\% \\
 &= \frac{126}{126 + 15} \cdot 100\% \\
 &= 89,4\% \\
 &= \frac{50}{50 + 37} \cdot 100\% \\
 &= 57,47\% \\
 &= \frac{(89,4 + 57,47)}{2} \% = 73,4\% \\
 &= \frac{(73,4 + 77,1)}{2} \% = 75,25\%
 \end{aligned}$$

Berdasarkan model backpropagation yang didapat, klasifikasi opini islam radikal dengan menggunakan metode artifial neural network, memperoleh nilai akurasi sistem sebesar 77%, rata-rata presisi sebesar 77,1%, rata-rata recall sebesar 73%, dan memperoleh F-Measure sebesar 75,25%. Dengan demikian dapat disimpulkan bahwa model cukup baik untuk digunakan pada klasifikasi data tweet yang mengandung sentimen islam radikal.

4.7 Analisis dan Visualisasi

Analisis sentimen terhadap data tweet yang mengandung opini islam radikal telah memperoleh hasil pengelompokan data dengan pelabelan yang maksimal dan memperoleh hasil akurasi yang cukup baik. Analisis dari hasil penelitian ini dapat dilakukan melalui visualisasi dengan menggunakan grafik wordcloud dengan frekuensi minimal 30 kali. Karena pada penelitian ini menggunakan dua kategori sentimen positif dan negatif, maka pada Gambar 4.2 menampilkan wordcloud untuk frekuensi data positif dan Gambar 4.3 menampilkan frekuensi data negatif.

Gambar 4.2 menunjukkan bahwa frekuensi tertinggi berupa kata “khilafah”, karena kata ini digunakan sebagai kata kunci untuk pengambilan data tweet pada saat crawling data di twitter. Pada wordcloud yang berkategori positif ini, kata yang muncul secara relevan

diantaranya kata “islam”, “negara”, “pancasila”, “ideologi”, “indonesia”, “nkri”, dan “hti”. Kata ini menjadi term yang berpengaruh terhadap klasifikasi data yang berkemungkinan tergolong pada sentimen positif. Maksud dari kata yang muncul adalah tweet yang mengandung beberapa kata ini cenderung memegang teguh ideologi pancasila sebagai ideologi tetap untuk persatuan NKRI di negara Indonesia. Selain itu kata “hti” muncul bersamaan dengan kata “anti” yang menyimpan suatu opini bahwa mereka tidak menginginkan beberapa organisasi masyarakat yang dinilai membawa ideologi baru untuk negara Indonesia. Gambar 4.3 menampilkan frekuensi wordcloud dengan kemunculan kata yang berpengaruh terhadap kategori tweet negatif.

Gambar 4.3 Wordcloud Data dengan Kategori Negatif

Gambar 4.3 menunjukkan bahwa beberapa kata yang berkontribusi untuk menjadikan suatu dokumen bersentimen negatif diantaranya kata “islam”, “muslim”, “khilafahajaranislam”, “ajaran”, “htilanjutkanperjuangan”, “htilayakmenang”, “tegak”, dan lain sebagainya. Beberapa dari kata ini merupakan semboyan dalam bentuk hashtag bagi sebagian organisasi masyarakat yang menginginkan perubahan terhadap ideologi Indonesia menjadi khilafah. Berdasarkan kedua gambar tersebut ada beberapa kata secara bersamaan muncul diantaranya kata “hti”. Namun kata “hti” yang berdekatan dengan kata “anti” cenderung memiliki sentimen positif, sedangkan jika berdekatan dengan kata “sistem” maka kecenderungan tweet memiliki sentimen negatif. Selain itu terdapat kata “islam” yang berdekatan dengan kata “pancasila” termasuk pada kategori sentimen positif dan jika berdekatan dengan kata “khalifah” masuk pada sentimen negatif. Selain kedua kata tersebut terdapat kata lain yang sama-sama muncul pada wordcloud kedua kategori sentimen diantaranya ada kata negara, sistem, dan indonesia. Analisis visualisasi data pada penelitian ini juga menggunakan jaringan graf yang melakukan analisis dari hubungan antar kata pada suatu dokumen menggunakan software VOSviewer yang dibedakan berdasarkan dua kategori pula yaitu positif pada Gambar 4.4 dan kategori negatif untuk Gambar 4.5.

Gambar 4.4 Hasil Visualisasi Kategori Positif

Visualisasi pada Gambar 4.4 terbagi menjadi dua cluster. Cluster pertama berwarna hijau yang memiliki tiga frase yaitu “nkri”, “pancasila”, dan “indonesia”. Ketiga frase ini berhubungan kuat dengan informasi masyarakat mengenai organisasi masyarakat yang berniat mendirikan negara kilafah, kebanyakan tweet yang ditulis oleh beberapa masyarakat ini menekankan bahwa Indonesia yang merupakan negara NKRI tetap pada ideologi sebelumnya yaitu Pancasila. Pernyataan ini senada dengan isi UUD 1945 dan UU yang berlaku secara di Negara Indonesia.

Pada cluster yang kedua memiliki 3 frase dengan warna merah yaitu khilafah, islam, dan positif. Arti pada keterhubungan dari tiga frase ini hanya menunjukkan bahwa data tweet dengan frase islam termasuk dikategorikan positif. Hal ini dikarenakan warga negara Indonesia mayoritas beragama Islam baik yang mendukung berdirinya negara khilafah maupun bukan. Frase ini ditunjukan oleh semboyan “Islam Yes, Khilafah No” yang banyak di retweet oleh beberapa akun.

Visualisasi terhadap data tweet yang berkategori negatif ditampilkan pada Gambar 4.5.

Gambar 4.5 Hasil Visualisasi Kategori Negatif

Hasil visualisasi pada Gambar 4.5 memiliki tiga cluster. Cluster pertama memiliki kasus dan penjelasan yang sama seperti cluster kedua pada hasil visualisasi kategori positif. Bedanya hanya terletak pada sentimennya saja. Cluster kedua yang berwarna biru dengan dua frase yaitu “khilafahajaranislam” dan “returnthekhilafah”. Berdasarkan frase ini suatu tweet yang beropini ingin mendirikan negara khilafah menganggap bahwa khilafah berasal dari ajaran islam melalui sejarah kekhalifaan pada zaman dahulu dan ingin diwujudkan pada masa yang akan datang di negara Indonesia. Kedua frase ini digunakan dalam bentuk hashtag, ini terlihat dengan penggabungan kata tanpa spasi pada frase tersebut.

Cluster ketiga divisualisasikan dengan tiga frase yang berwarna merah. Sama halnya dengan cluster kedua, frase-frase pada cluster ini juga berupa hashtag yang ditandai dengan penggabungan kata tanpa spasi yaitu “islamselamatkan negerinya”, “allahbersamahtinya”, dan “antikhilafahantiislam”. Frase ini muncul pada tweet yang secara terang-terangan menyebutkan ikon organisasi mereka. Tweet yang mengandung frase ini sudah jelas memiliki sentimen negatif terhadap NKRI yang artinya mendukung berdirinya negara khilafah. Penyebutan ikon dari organisasi masyarakat ini muncul seiring dengan fakta berita yang dilansir dari beberapa web media resmi seperti detik.com dan tribunnews.com mengenai pengesahan pembubaran organisasi masyarakat yang mereka naungi.

5. Kesimpulan dan Saran

5.1 Simpulan

Berdasarkan hasil pembangunan model klasifikasi dari uji coba sistem dengan menggunakan artificial neural network terhadap data tweet yang diperoleh, dapat disimpulkan bahwa:

1. Artificial neural network berperan dalam melakukan klasifikasi opini islam radikal dari sentimen data tweet pada twitter. Proses ekstraksi dokumen menggunakan metode text mining untuk memperoleh pola inputan yang terstruktur berdasarkan sejumlah data yang dimiliki untuk diklasifikasi sentimennya.
2. Membangun model klasifikasi sentimen tweet islam radikal menggunakan metode backpropagation dapat dilakukan berdasarkan uji coba jumlah node hidden, jumlah hidden layer, besar learning rate, dan pembagian data training-testing. Uji coba sistem menghasilkan model arsitektur dengan 579 node inputan, 2 hidden layer (jumlah node hidden pertama sebanyak 4 node dan 2 node untuk hidden layer kedua), besar nilai learning rate adalah 0.1, pembagian data sebanyak 80% untuk data training dan 20% untuk testing. model dikatakan cukup baik karena memperoleh performa akurasi sebesar 77%, presisi sebesar 77,1%, recall sebesar 73,4%, dan f-measure sebesar 75,25%.
3. Hasil analisis menggunakan graph wordcloud dan graph network, kecenderungan kategori sentimen opini tergantung pada semboyan dan kata yang digunakan pada tweet. Tweet yang mengandung sentimen islam radikal menggunakan frase “returnthekhilafah”, “islamselamatkan negerinya”, “antikhilafahantiislam”, “allahbersamahtinya”, dan “khilafahajaranislam”. Dengan demikian pengguna twitter yang berperan sebagai penerima informasi dapat membedakan antara tweet yang mengandung opini islam radikal atau tidak sebagai bentuk waspada agar tidak mudah mempercayai dan terhasut oleh opini tweet dari orang lain.

Daftar Pustaka

Abraham, F., 2014. *Pemanfaatan Twitter Sebagai Sarana Komunikasi Massa*, Bajarmasin: Kementrian Komunikasi dan Informatika Republik Indonesia.

Agung, I. & Santoso, A., 2018. *Pola Deteksi Dini dalam Menangkal Faham Radikalisme di Lingkungan Pendidikan*. Jakarta: Penerbit Edu Pustaka.

Angkoso, C. V., Hendrawan, Y. F., Kusumaningsih, A. & Wahyuningrum, R. T., 2018. *Performance Analisis of Color Cascading Framework on Two Different Classifier in Malaria Detection*. EECSI, pp. 1-4.

Araque, O., Zhu, G., Amado, M. G. & Iglesias, C. A., 2016. *Mining the Opinion web: Classification and Detection of Aspect Contexts for Aspect Based Sentiment Analysis*. IEEE 16th International Conference on Data Mining Workshops.

Asiyah, S. N. & Fithriasari, K., 2016. *Klasifikasi Berita Online Menggunakan Metode Support Vector Machine dan K-Nearest Neighbor*. Jurnal Sains dan Seni ITS.

Azizan, S. A. & Aziz, I. A., 2017. *Terrorism Detection Based on Sentiment Analysis Using Machine Learning*. Journal of Engineering and Applied Sciences.

Balikas, G., Moura, S. & Mini, M.-R. A., 2017. *Multitask Learning for Fine- Grained Twitter Sentiment Analysis*. Grenoble Alps Univercity, Volume 1, pp. 1-5.

Bashashati, H., Ward, R. K., Birch, G. E. & Bashashati, A., 2015. *Comparing Different Classifiers in Sensory Motor Brain Computer Interfaces*. PLoS One, Volume 10, p. 6.

Damarjati, D., 2018. *detiknews.com*. [Online] Available at: <https://news.detik.com/berita/d-4358370/terorisme-terlaknat-2018-bom-sekeluarga-mengguncang-surabaya> [Diakses 20 Juli 2019].

Effendi, A., 2013. *Penggunaan Artifial Neural Network untuk Mendeteksi Kelainan Mata Miopi pada MANusia dengan Metode Backpropagation*, Malang: Universitas Islam Negeri Maulana Malik Ibrahim.

Fauset, L., 1994. *Fundamentals of Neural Networks*. New Jersey: Prentice Hall. Fink, C. R., Chou, D. S. & Kopecky, J. J., 2011. Llorens, Ashley J. JOHN HOPKINS APL TECHNICAL DIGEST, Volume 30, pp. 22-30.

Gonzales, C. B., Garcia-Nieto, J., Navas-Delgado, I. & Aldana-Montes, J. F., 2015. *A Fine Sentiment Analysis with Semantics in Tweets*. International Journal of Interactive Multimedia and Artificial Intelligence, Volume 3, pp. 22-28.

Gurney, K., 1997. *An Introduction to Neural Network*. London: UCL Press.

Haewoo, K., Changhyun, L., Hosung, P. & Moon, S., 2010. *What is Twitter, a Social Network or a News Media*. IW3C2, pp. 591-600.

Novitasari, D. C. R., 2015. *Klasifikasi Sinyal EGG Menggunakan Metode Fuzzy C-Means Clustering (FCM) dan Adaptive Neighborhood Modified Backpropagation*. Jurnal Matematika "MANTIK&qout, Volume 1, pp. 31-36.

Ozturk, N. & Ayvaz, S., 2017. *Sentiment Analysis on Twitter: A text Mining Approach to The Syrian Refugee Crisis*. Telematics and Informatics.

Peraturan Pemerintah, 2017. *Perubahan UU No 17 Tahun 2013 Tentang Organisasi Masyarakat*. s.l.:s.n.

Putra, E. D., 2014. *Menguak Jejaring Sosial*. Surabaya: Graha Ilmu.

Rinzel, J. & Ermentrot, G. B., 1989. *Analysis of Neuronal Excitability and Oscillations. Methods In Neuronal Modelling: From Synapses to Networks*.

Sarwani, M. R. & Mahmudy, W. F., 2015. *Analisis Twitter untuk Mengetahui Karakteristik Seseorang Menggunakan Algoritma Nai'Ve Bayes Classifier*. Sminar Nasional Sistem Informasi Indonesia, 2-3 November 2015.

Setiawan, K. Y., Hidayati, H. & Gozali, A. A., 2014. *Analisis User Opinion Twitter pada Level Fine-Grained Sentiment Analysis Terhadap Tokoh Publik*. e- Proceeding of Engineering, Volume 1, pp. 639-646.

Siang, J. J., 2005. *Jaringan Saraf Tiruan dan Pemogramannya Menggunakan Matlab*. Yogyakarta: ANDI.

Sutikno, et al., 2016. *Backpropagation dan aplikasinya*. Ilmu kOmputer Studi Kasus dan Aplik, pp. 135-146.

Suzuki, K., 2013. *Artificial Neural Network-Architectures and Applications*. Croatia: InTech.

Syafi'i, I., 2018. *Peraturan Pemerintah Pengganti Undang-Undang (Perppu) Nomor 2 Tahun 2017 Tentang Organisasi Kemasyarakatan (Ormas) dalam Pandangan Pengurus Wilayah Nahdlatul Ulama (PWNU) Jawa Timur*, Surabaya: Universitas Islam Negeri Sunan Ampel.

+Thoyyib, M., 2018. *Radikalisme Islam Indonesia*. TA'LIM : Jurnal Studi Pendidikan Islam, Volume 1, pp. 90-105.

Wechmongkhlonkon, S., Poomtong, N. & Areerachakul, 2012. *Application of Artificial Neural Network to Classification Surface Water Quality*. International Journal of Environmental and Ecological Engineering, Volume 6, pp. 574-574.

Yanti, N., 2011. *Penerapan Metode Neural Network Dengan Struktur Backpropagation untuk Prediksi Stok Obat di Apotek (Studi Kasus : Apotek ABC)*. Seminar Nasional Aplikasi Teknologi Informasi, pp. 15-20.

Yuan, L. & Rachmat, A., 2016. *Sentipol: dataset Sentimen Komentar pada Kampanye PEMILU Presiden Indonesia 2014 dari Facebook Page*. Konferensi Nasional Teknologi Informasi dan Komunikasi (KNASTIK), pp. 218-228.

Yulian, E., 2018. *Text Mining dengan K-Means Clustering pada Tema LGBT dalam Arsip Tweet Masyarakat Kota Bandung*. *Jurnal MAtematika "MANTIK"*, Volume 4, pp. 53-58.