ANALISIS SENTIMEN PADA MEDIA SOSIAL *TWITTER* UNTUK KLASIFIKASI OPINI ISLAM RADIKAL MENGGUNAKAN JARINGAN SARAF TIRUAN

SKRIPSI

Disusun Oleh:

ROUDATUL ZANNAH

NIM: H72215023

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN AMPEL

SURABAYA

2019

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : Roudatul Zannah

NIM : H72215023

Program Studi : Matematika

Angkatan : 2015

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi yang berjudul "ANALISIS SENTIMEN PADA MEDIA SOSIAL *TWITTER* UNTUK KLASIFIKASI OPINI ISLAM RADIKAL MENGGUNAKAN JARINGAN SARAF TIRUAN". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 25 Juli 2019

Yang menyatakan,

METERAI TEMPEL

32249ADF867073026

6000

(Roudatul Zannah)

NIM: H72215023

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh :

NAMA : Roudatul Zannah

NIM : H72215023

JUDUL : Analisis Sentimen pada Media Sosial *Twitter* untuk Klasifikasi Opini Islam Radikal Menggunakan Jaringan Saraf Tiruan

Ini telah diperiksa dan disetujui untuk diujikan.

Surabaya, 11 Juli 2019

Dosen Pembimbing 1	Dosen Pembimbing 2
(Nurissaidah Ulinnuha, M.Kom)	(Dian Candra Rini N, M. Kom)
NIP: 199011022014032004	NIP: 198511242014032001

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh

NAMA : Roudatul Zannah

NIM : H72215023

JUDUL :Analisis Sentimen pada Media Sosial *Twitter* untuk Klasifikasi Opini IslamRadikal Menggunakan Jaringan Saraf Tiruan

Telah dipertahankan di depan tim penguji skripsi

Pada hari Selasa Tanggal 30 Juli 2019

Mengesahkan,

Tim Penguji

Penguji I

(Narissaidah Ulinnuha, M.Kom)

NIP: 199011022014032004

Penguji II

(Dian Candra Rini N, M. Kom)

NIP: 198511242014032001

Penguji III

(Dr. Moh Hafiyusholeh, M.Si)

NIP: 198002042014031001

Penguji IV

(Putroue Keumala Intan, M.Si)

NIP: 198805282018012001

Mengetahui,

Dekan Fakultas Sains dan Teknologi

UIN Sunan Ampel Surabaya

(Dr. Evi Purwati, M.Ag)

NIP: 196512211990022001

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN
Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : RAUDATUL ZANNAH
NIM : H72215023
Fakultas/Jurusan : SAINTEK / MATEMATIKA
E-mail address : jannahraudatul77@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :
☑ Skripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :

ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER UNTUK KLASIFIKASI OPINI ISLAM RADIKAL MENGGUNAKAN JARINGAN SARAF TIRUAN

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara ***fulltext*** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 27 JULI 2019

Penulis

(RAUDATUL ZANNAH)
nama terang dan tanda tangan

ABSTRAK

ANALISIS SENTIMEN PADA MEDIA SOSIAL *TWITTER* UNTUK KLASIFIKASI OPINI ISLAM RADIKAL MENGGUNAKAN JARINGAN SARAF TIRUAN

Paham radikalisme memiliki banyak pengaruh negatif terhadap persatuan negara Indonesia, bahkan dapat menghancurkan NKRI yang berpedoman pada Pancasila. Organisasi yang berpaham radikalis 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 wordcould* dan *graph network* menunjukkan kecenderungan kategori sentimen opini tergantung pada semboyan dan kata yang digunakan pada *tweet*.

Kata Kunci*: islam radikalis, twitter, text mining, ANN, backpropagation*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

ABSTRACT

SENTIMENT ANALYSIS ON TWITTER SOCIAL MEDIA FOR THE CLASSIFICATION OF RADICALS ISLAM OPINION USING ARTIFICIAL NEURAL NETWORKS

Radicalism has many negative effect on Indonesian unity. NKRI that guided by Pancasila will be destroyed if radicalism is present. Radicalist organization has intolerance, that could not accept groups who don't have same understanding. This organization suffocates the Indonesian people with dogmas on the name of religion. Because of that, the aim of those who want to change Indonesia become a khilafah state can be realized. On this advanced era, social media can be a new alternative to spread understanding that could theathen the unity of the NKRI nation. Therefore, sentiment analysis needed to identify someone's post whether it is related to radicalism content. One of social media that still active and widely used in the dissemination of this understanding is twitter. On this research, data processing using text mining method as a start to obtain numerical data from document data. The result of data pre-processing will be later classified using the ANN's architecture with backpropagation method that has been obtained. The ANN's architecture results is good enough with an accuracy of 77%, an precision of 77,1%, a recall of 73,4%, and f-measure of 75,25%. Analyze of result using graph wordcould and graph network, with trend of sentiment categories opinion based on motto tweet and word which uses on tweet.

Keyword: islam radikalis, twitter, text mining, ANN, backpropagation

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING i

PENGESAHAN TIM PENGUJI SKRIPSI ii

PERNYATAAN KEASLIAN iii

MOTO iv

HALAMAN PERSEMBAHAN v

KATA PENGANTAR vi

ABSTRAK vii

ABSTRACT viii

DAFTAR ISI ix

DAFTAR TABEL xii

DAFTAR GAMBAR xiv

BAB I PENDAHULUAN 1

1.1 Latar Belakang 1

1.1 Rumusan Masalah 6

1.2 Tujuan Penelitian 7

1.3 Manfaat Penelitian 7

1.4 Batasan Masalah 7

BAB II LANDASAN TEORI 8

2.1 Islam radikalis 8

2.2 *Twitter* 9

2.3 *Twitter* APl 9

2.4 *Twitter Fine-Grained* 10

2.5 Pelabelan Data 11

2.6 *Text Mining* 12

2.6.1 *Text Pre-processing* 13

2.6.2 Pembobotan Data 14

2.7 Analisis Sentimen 15

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2.8 *Artificial Neural Network* 16
2.8.1 Klasifikasi ANN 19
2.8.2 Arsitektur *Artificial Neural Network* 20
2.8.3 *Backpropagation* 21
2.8.4 Uji Coba Model ANN 31
2.9 Evaluasi Sistem 32
BAB III METODE PENELITIAN 34
3.1 Jenis Penelitian 34
3.2 Sumber Data 34
3.3 Pengolahan Data 35
3.4 Rancangan Sistem 36
3.4.1 *Crawling* Data *Twitter* 37
3.4.2 Pelabelan dan Penghapusan Data 38
3.4.3 *Pre-Processing* 38
3.4.4 Pembobotan TF-IDF 41
3.4.5 Klasifikasi *Artificial Neural Network* (ANN) 41
3.4.6 Pengujian Model 42
3.4.7 Analisis dan Visualisasi. 42
BAB IV HASIL DAN PEMBAHASAN 43
4.1 *Crawling Data* 43
4.2 Pelabelan Data 44
4.3 *Pre-processing* 45
4.4 Pembobotan TF IDF 48
4.5 Klasifikasi ANN 49
4.6 Hasil Pengujian Model dan Akurasi 56
4.6.1 Pengujian Jumlah *Node* pada *Hidden layer* 57
4.6.2 Uji Coba Learning Rate 59
4.6.3 Penentuan Model 59
4.7 Analisis dan Visualisasi 61
BAB V PENUTUP 66
5.1 Simpulan 66

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

5.2 Saran 67
DAFTAR PUSTAKA 68
LAMPIRAN

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR TABEL

Tabel 2.1 Contoh *Confusion Matrix* dua Kelas 32

Tabel 3.1 Sampel Data *Tweet* Berdasarkan *Mention Account* 35

Tabel 3.2 Komponen yang Dihapus 39

Tabel 3.3 Sampel Data *Stopword* 40

Tabel 3.4 Ilustrasi Hasil Pembobotan TF-IDF 41

Tabel 4.1 Sampel dari Hasil *Crawling Data Tweet* 43

Tabel 4.2 Contoh Pelabelan Data 45

Tabel 4.3 Sampel Data dengan Label Positif dan Negatif 45

Tabel 4.4 Sampel Data Sesudah Tahap *Cleansing* 46

Tabel 4.5 Sampel Data Sesudah Tahap *Case Folding* 46

Tabel 4.6 Sampel Data Sesudah Penghapusan *Stopword* 47

Tabel 4.7 Sampel Data Sesudah Tahap Tokenisasi 47

Tabel 4.8 Pembobotan TF 48

Tabel 4.9 Pembobotan IDF 48

Tabel 4.10 Pembobotan TF IDF 49

Tabel 4.11 Sampel Bobot Hasil TF IDF untuk Klasifikasi ANN 49

Tabel 4.12 Inisialisasi Bobot *Input* Iterasi Pertama 51

Tabel 4.13 Inisialisasi Bobot *Output* Iterasi Pertama 51

Tabel 4.14 Perubahan Bobot ΔV_{kj} 54

Tabel 4.15 Inisialisasi Bobot *Input* Iterasi Kedua 55

Tabel 4.16 Inisialisasi Bobot *Output* Iterasi Kedua 55

Tabel 4.17 Uji Coba Jumlah *Node* pada *Hidden Layer* pada Grup Pertama 57

Tabel 4.18 Uji Coba Jumlah *Node* pada *Hidden Layer* pada Grup Kedua 58

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.19 Uji Coba *Learning Rate* 59

Tabel 4.20 *Confusion Matrix* 60

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR GAMBAR

Gambar 2.1 Jaringan Saraf Manusia 16

Gambar 2.2 Fungsi ANN 18

Gambar 2.3 Asitektur Klasifikasi ANN 20

Gambar 2.4 Arsitektur untuk Alur *Backpropagation* 23

Gambar 3.1 *Flowchart* Pengolahan Data 36

Gambar 4.1 Arsitektur Klasifikasi untuk Sampel Data 50

Gambar 4.2 *Wordcould* Data dengan Kategori Positif 61

Gambar 4.3 *Wordcould* Data dengan Kategori Negatif 62

Gambar 4.4 Hasil Visualisasi Kateogori Positif 63

Gambar 4.5 Hasil Visualisasi Kateogori Negatif 64

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB I

PENDAHULUAN

1.1 Latar Belakang

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 tercanangkan dalam konstitusi NKRI yaitu UUD 1945.

Paham radikalis 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 radikalis. Akibatnya, masyarakat yang tergabung pada organisasi masyarakat ini jauh dari rasa nasionalisme terhadap bangsanya sendiri. Pada kehidupan sosialnya pun paham radikalis kerap kali mewujudkan perilaku ekstrim, antoleransi, dan destruktif dalam merealisasikan tujuannya (Agung & Santoso, 2018).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Salah satu organisasi yang ramai dibicarakan dan dituding sebagai organisasi islam radikalis dengan paham radikalis adalah *Hiztbut Tahrir Indonesia* (HTI). Organisasi ini secara terang-terangan menyemarakkan Indonesia menjadi negara khilafah, dianggap tidak sesuai dengan Pancasila, terutama sila pertama. Pada tanggal 19 Juli 2017 organisasi ini resmi dibubarkan (Syafi'i, 2018). Ada beberapa organisasi masyarakat yang diduga menyimpang dari nilai Pancasila, contohnya seperti kelompok organisasi Ikhwanul Muslimin (IM) (Thoyyib, 2018). Penyimpangan semacam ini sebenarnya sudah lama terjadi di awal kemerdekaan Indonesia. Salah satu organisasi yang menyimpang pada awal kemerdekaan adalah DI/II, organisasi ini menginginkan Indonesia menjadi negara islam (Khamid, 2016).

Penyimpangan yang marak dilakukan oleh organisasi masyarakat yang berpaham radikalis menjadikan negara Indonesia jauh dari kata kesatuan yang bersimbolkan Bhineka Tunggal Ika. Organisasi masyarakat yang berpaham radikalis mengedepankan golongannya sendiri dan tidak mentoleransi perbedaan di lingkungan sekitarnya, apalagi yang menyangkut masalah perbedaan agama. Misalnya kejadian pengeboman yang terjadi pada bulan mei 2018 lalu yang dilakukan oleh beberapa orang dari organisasi masyarakat islam radikal di sebuah gereja (Damarjati, 2018). Kejadian ini mencerminkan bahwa sebagian dari mereka melakukan tindakan yang melanggar hukum untuk menghilangkan perbedaan yang mereka anggap salah dan tidak sejalan dengan mereka. Permasalahan ini dapat menjadikan Indonesia terpecah-belah dan menyebabkan banyak permusuhan dimana-mana.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Seiring dengan majunya teknologi pada masa modern ini, media sosial menjadi alternatif baru untuk menebarkan paham yang bisa mengancam kesatuan bangsa NKRI. Para pelaku islam radikalis mendoktrin masyarakat dengan pengetahuan agama yang condong radikal. Banyaknya informasi kurang valid dan berbau radikal yang tersebar di media sosial, menyebabkan perlunya menganalisis sentimen suatu artikel atau postingan seseorang untuk mengetahui apakah postingan tersebut berasal dari golongan islam radikalis atau bukan. Tujuan dari tindakan ini adalah agar masyarakat tidak mudah mempercayai begitu saja terhadap opini seseorang di media sosial yang membawa konten radikal.

Salah satu media sosial yang masih aktif dan banyak digunakan oleh berbagai kalangan adalah *twitter*. Sosial media ini dinilai banyak memberikan informasi tentang sentimen postingan seseorang. Penggunaan sosial media *twitter* ini, dibuktikan dengan adanya beberapa penelitian yang menggunakan *twitter* sebagai bahan untuk menganalisis sentimen opini seseorang. Diataranya 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). Pada penelitian ini juga akan melakukan analisis sentimen terhadap opini seseorang dari *tweet* yang mereka unggah di *twitter* menggunakan *text mining*. Analisis yang dimaksud adalah mengetahui keberpihakan seseorang apakah teridentifikasi dari golongan radikalis atau tidak.

Metode *text mining* digunakan karena dalam pengerjaannya dapat mengubah setiap kata dan huruf dokumen menjadi data numerik. *Text mining* dipercaya menjadi salah satu metode untuk seleksi fitur dokumen yang baik dan banyak

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

digunakan oleh para peneliti. Salah satu peneliti yang melakukan analisis sentimen media sosial *twitter* menggunakan metode *text mining* yaitu Eko Yulian. Eko Yulian melakukan analisis arsip dengan melakukan *clustering* pada dokumen menggunakan *text mining* dengan *K-means clustering* (Yulian, 2018). Selain itu, terdapat penelitian yang dilakukan oleh Ozturk dan tim yang menganalisis sentimen *twitter* terhadap krisis yang terjadi di suriah lewat postingan yang dilakukan masyarakat (Ozturk & Ayvaz, 2017). Penelitian lain mengenai *text mining* juga dilakukan oleh Mirani dan tim yang menganalisis lokasi pengaruh ISIS melalui media sosial *twitter* menggunakan *text mining* (Mirani & Sasi, 2016), dan banyak lagi peneliti lainnya. Hal ini membuktikan bahwa metode *text mining* menjadi metode yang tepat untuk suatu kasus analisis sentimen.

Pada penelitian ini menggunakan metode *artificial neural network* (ANN) untuk melakukan klasifikasi data dokumen sekaligus mengetahui adanya postingan berupa islam radikalis. Metode ini merupakan jaringan saraf tiruan yang pada umumnya dapat digunakan dalam klasifikasi berbagai permasalahan dengan data non linear. ANN memiliki kinerja yang baik dan banyak digunakan pada masalah visi komputer untuk pengenalan pola. Banyak peneliti yang melakukan klasifikasi dalam menyelesaikan penelitiannya dengan menggunakan metode ini. 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. Leavline melakukan klasifikasi citra bunga anggrek untuk membedakan variasi tanamannya

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

(Leavline, 2015). Selain itu metode ANN digunakan oleh Wechmongkhonkon dan tim dalam mengklasifikasi permukaan air guna memanajemen kualitas air dengan jaringan saraf tiruan MLP dalam proses klasifikasinya (Wechmongkhlonkon, et al., 2012). Klasifikasi kualitas air juga pernah dilakukana sebelumnya oleh Meair dan tim yang lebih condong pada model prediksi ketahanan kualitas air berdasarkan variabel permasalahannya (Maeir & Dany, 2000). Hal ini menunjukkan bahwa klasifikasi menggunakan metode ANN dinilai masih memberikan hasil yang cukup baik.

Metode ANN dapat digunakan dalam penelitian mengenai analisis sentimen pada dokumen teks. Metode ini pernah digunakan oleh Jian dan tim dalam penelitiannya untuk mengungkapkan pendapat positif atau negatif terhadap suatu topik. Hasil klasifikasi yang diperoleh menunjukkan metode ANN memiliki akurasi lebih tinggi dari pada SVM dan HMM pada suatu analisis *review* film (Jian, et al., 2010). Penelitian lain juga menyebutkan bahwa untuk kasus analisis sentimen, metode ANN lebih unggul daripada SVM terutama untuk konteks data tidak seimbang pada 13 tes. Sedangkan SVM dapat mengungguli ANN hanya dengan 2 tes saja, meskipun perbedaan akurasi pada keduanya tidak pernah melebihi 3% (Moraes, et al., 2013). Berdasarkan penelitian yang membuktikan bahwa ANN berhasil dalam melakukan klasifikasi, maka pada penelitian ini menggunakan metode ANN agar memperoleh model dan akurasi klasifikasi yang baik dan optimum. Penelitian ini melakukan analiasis sentimen terhadap *tweet* seseorang di *twitter* untuk mengetahui apakah opini yang seseorang unggah teridentifikasi dari golongan islam radikalis. Analisis sentimen yang dilakukan bertujuan untuk

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

mencegah adanya penyebaran paham radikalis yang tengah marak di media sosial dan diharapkan masyarakat dapat dengan tepat menanggapi berita yang tersebar di media sosial. Dalam Alqur'an pun telah dijelaskan mengenai pentingnya menganalisis suatu berita yang diterima (*tabayyun*) seperti firman Allah dalam QS Al-Hujarat ayat 6 yang berbunyi:

يَـٰٓأَيُّهَا ٱلَّذِينَ ءَامَنُوٓاْ إِن جَآءَكُمۡ فَاسِقُۢ بِنَبَإٖ فَتَبَيَّنُوٓاْ أَن تُصِيبُواْ قَوۡمَۢا بِجَهَٰلَةٖ فَتُصۡبِحُواْ عَلَىٰ مَا فَعَلۡتُمۡ نَٰدِمِينَ ٦

Artinya: "Hai orang-orang beriman, jika datang orang fasik membawa suatu berita, maka periksalah dengan teliti agar kamu tidak menimpakan suatu musibah kepada suatu kaum tanpa mengetahui keadaannya yang menyebabkan kamu menyesal atas perbuatanmu itu".

Pada penelitian ini melakukan analisis sentimen dengan menggunakan metode *text mining* sebagai metode yang terpercaya untuk mengatasi data teks dan ANN sebagai metode klasifikasinya, diharapkan akan diperoleh model dan akurasi yang optimum agar menghasilkan visualisasi dan analisis yang baik dan tepat.

1.1 Rumusan Masalah

Rumusan masalah dari penelitian ini antara lain:

a) Bagaimana penerapan *artificial neural network* dalam mengklasifikasikan opini *tweet* islam radikalis?
b) Bagaimana mengevaluasi model *artificial neural network* untuk klasifikasi opini islam radikal?
c) Bagaimana hasil analisis sentimen di media sosial *twitter* berdasarkan hasil klasifikasi opini islam radikal ?

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

1.2 Tujuan Penelitian

Adapun penelitian yang dilakukan memiliki tujuan sebagaimana berikut:

a) Mengetahui penerapan *artificial neural network* dalam mengklasifikasikan opini *tweet* islam radikalis
b) Membangun model *artificial neural network* untuk klasifikasi sentimen
c) Menganalisis sentimen media sosial *twitter* berdasarkan hasil klasifikasi opini islam radikalis

1.3 Manfaat Penelitian

Penelitian yang dilakukan memberikan manfaat secara teoritis dan praktis sebagaimana berikut:

a) Manfaat Teoritis

   Sebagai sarana pengembangan penerapan keilmuan di dunia sains dan *engineering* dengan metode yang tepat untuk mencari solusi untuk permasalahan analisis sentimen data dokumen.
b) Manfaat Praktis

   Sebagai acuan untuk melakukan penelitian serupa dan dapat diterapkan di masa yang akan datang

1.4 Batasan Masalah

Penelitian ini memiliki batasan masalah diantaranya:

a) Data yang digunakan merupakan *tweet* yang berbahasa Indonesia
b) Memakai metode pembobotan *TF-IDF* pada tahap kedua *pre-pocessing*
c) Pada tahap pertama *pre-processing* tidak menggunakan s*temming*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB II

LANDASAN TEORI

2.1 Islam radikalis

Organisasi masyarakat Indonesia yang berideologi islam radikal dinilai menyimpang dari ideologi bangsa Indonesia, yakni Pancasila. Ideologi Islam radikal ini dipengaruhi oleh paham *Islamic State in Iraq and Syria* (ISIS) dengan misi menyebarkan dogma-dogma ajaran Islam yang disalahgunakan untuk jihad dan menginginkan Indonesia menjadi negara khilafah (Khamid, 2016). Ideologi radikal dapat menyebar dan mempengaruhi masyarakat seluruh pelosok Indonesia dengan membuat masyarakat yang menjadi targetnya, menjadi satu tujuan dan satu ideologi dengan mereka.

Masyarakat yang tergabung dalam organisasi masyarakat yang berideologi radikal ini mengesampingkan rasa nasionalisme yang wajib dimiliki setiap warga negara Indonesia. Mereka lebih memilih Indonesia menjadi negara Islam dan mengabaikan sejarah yang terkandung dalam nilai-nilai Pancasila hasil perjuangan pahlawan terdahulu yang menginginkan warga Indonesia menjadi negara Bhineka Tunggal Ika. Larangan dan tindakan keras terhadap organisasi yang tidak berlandaskan Pancasila ini disebutkan keras dalam Peraturan Pemerintah terbaru (Peraturan Pemerintah, 2017).

Beberapa hal negatif akibat paham radikalis adalah munculnya kekerasan, antoleransi, ekstrimisasi, dan destruktif. Lebih berbahaya lagi jika menyebabkan tindakan anarkis yang dapat meresahkan masyarakat.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2.2 *Twitter*

Twitter merupakan salah satu media massa jejaring sosial dengan model *microblog* daring yang menyediakan fasilitas mengirim dan membaca pesan yang digunakan lebih dari 41 juta pengguna pada bulan Juli 2009 lalu (Haewoo, et al., 2010). Pesan yang disampaikan merupakan pesan teks yang terdiri dari 140 karakter dan bertambah menjadi 280 karakter pada tahun 2017 lalu (Putra, 2014).

Twitter pertama kali dikenalkan pada tahun 2006 oleh seorang sarjana dari Universitas New York yang bernama Jack Dorsey pada acara diskusi perusahaan *podcast*. Jack Dorsey mengemukakan gagasannya tentang layanan pesan singkat untuk berkomunikasi dalam sebuah kelompok kecil. Hingga pada tahun 2007 *twitter* dapat menjadi perusahaan mandiri dan meningkat hingga 60.000 *tweet* pada tahun 2011(Putra, 2014).

Twitter merupakan jejaring sosial layanan web dengan empat pola dalam berkomunikasi yaitu, pola komunikasi berkicau, pola broker, pola komunikasi ideal, pola komunikasi interaktif (Abraham, 2014). Selain sebagai alat berkomunikasi, *twitter* juga sebagai media penyampaian informasi dan isu-isu terkini dengan menggunakan *hashtag* topik yang sedang dibahas. *Hashtag* yang paling banyak disebutkan akan menjadi trending topik. Biasanya *twitter* menampilkan 10 trending topik utama pada penggunanya agar dapat mengetahui informasi terkini (Kwak, et al., 2010).

2.3 *Twitter* APl

Twitter Aplication Progamming Language (APl) merupakan fasilitas penambangan data dari media sosial *twitter*, sehingga data dan informasi yang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

masuk dikirim ke suatu sistem informasi menggunakan akun si pengguna *twitter* (Iswahyudi & Hasan, 2016). Seorang pengguna *twitter* dalam penggunaan media sosial *twitter* pasti memiliki profil singkat. Pengguna *twitter* dapat melakukan post dan melihat postingan orang lain. Data dan informasi seputar kegiatan pengguna selama aktif di media sosial *twitter* ini dapat diperoleh dengan menggunakan *twitter* API.

Terdapat beberapa informasi yang dapat diambil menggunakan API yaitu, teman dan *followers* si pengguna, postingan si pengguna, hasil pencarian pada *twitter*, tempat, dan lokasi (Kwak, et al., 2010). Desain *twitter* sendiri adalah menyediakan metode tunggal untuk mengakses data dari API *twitter* dengan cara membuat aplikasi di apps.*twitter*.com dan menghasilkan kunci *oAuth* bagi masing-masing pengunanya.

2.4 *Twitter Fine-Grained*

Twitter fine-grained merupakan teknik untuk melakukan analisis sentimen terhadap unggahan seseorang di media sosial *twitter*. Secara umum, *fine-grained* itu sendiri bertujuan menentukan sentimen pada suatu kalimat (Gonzales, et al., 2015). Obyek yang akan diklasifikasi pada teknik *fine-grained* merupakan kalimat pada data dokumen. Kalimat tersebut dibedakan menjadi dua jenis, yaitu positif dan negatif. Kalimat yang dianalisis dan memiliki informasi obyektif merupakan kalimat yang memiliki nilai perasaan atau makna yang tersembunyi dalam kalimat si penulis. Kalimat seperti ini biasanya muncul pada opini, pendapat, ataupun komentar seseorang yang dapat bernilai positif atau negatif (Setiawan, et al., 2014).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Nilai positif ataupun negatif yang tersirat dalam suatu kalimat opini, pendapat, atau komentar seseorang dinamakan sebagai polaritas, sedangkan untuk opini, pendapat, atau komentar seseorang yang dapat bernilai positif atau negatif dinamakan sebagai sentimen. Polaritas dari sentimen inilah yang nantinya digunakan dalam klasifikasi atau prediksi pada suatu penelitian (Fink, et al., 2011). Berdasarkan teknik *fine-grained* yang sudah dilakukan, selain mempermudah proses analisis dan klasifikasi, kemudahan juga akan diperoleh dalam melakukan langkah-langkah selanjutnya (Balikas, et al., 2017) diantaranya yaitu:

a) Dapat menggunakan ekstraksi fitur sama untuk mewakili bentang teks yang dinilai hampir mirip.
b) Jika seseorang mengetahui nilai sentimen apa pada suatu opini, maka dapat mempersempit keputusan klasifikasi.

2.5 Pelabelan Data

Dataset berupa teks yang ditambang dari media sosial seperi *twitter*, *facebook*, *instagram*, atau media sosial lainnya merupakan data yang menyimpan sentimen tertentu pada setiap kalimat opini yang diunggah. Data teks yang ditambang untuk kasus yang dialami negara tertentu lebih baik menggunakan data teks dengan bahasa pada negara tersebut. Misalkan kasus politik yang terjadi di negara Indonesia, maka data teks yang di*crawling* dan dilabelkan berupa data teks dengan Bahasa Indonesia. Penambangan dengan bahasa yang sesuai ini dapat mempermudah dalam pelabelan suatu kalimat (Yuan & Rachmat, 2016).

Dataset yang digunakan dalam proses klasifikasi memerlukan mekanisme pelabelan yang tepat agar dapat diproses dengan baik pada tahap *training* dan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

testing data. Ada beberapa cara dalam melabelkan data. Pertama, melabelkan data secara manual. Cara ini terbilang sederhana dan memerlukan pakar khusus dalam menentukan pelabelan sesuai dengan kasus atau tema permasalahan dari data teks yang ditambang. Pelabelan dilakukan dengan mengelompokan data menjadi beberapa kategori. Pembagian ini dilakukan untuk mneghindari bias dalam proses kategori opini. Namun cara ini mempunyai kelemahan, yaitu membutuhkan waktu cukup lama pada saat dataset terbilang besar dan banyak (Matsubara, et al., 2008). Cara yang kedua menggunakan teknik *crowsourced labelling*, teknik ini dapat melakukan pelabelan secara cepat bahkan dengan data dalam jumlah besar sekalipun. Namun pelabelan dengan teknik *crowsourced labelling* ini, akan menghasilkan label yang bervariasi dan kurang baik dalam kategori pasti. *Crowsourced labelling* membutuhkan banyak pelabel (responden) untuk meningkatkan kualitas label. Cara ini pernah dilakukan oleh Welinder dan Perona dalam penelitiannya yang memberi pelabelan pada sekumpula gambar dengan memanfaatkan layanan *Amazon Mechanical Turk*.

2.6 *Text Mining*

Text mining merupakan metode yang bisa dikatakan hampir mirip dengan *data mining*. Hanya saja data yang diolah merupakan data dalam bentuk teks. Tujuan awal dari *text mining* lebih mengutamakan untuk menemukan suatu pola-pola yang dapat dijadikan informasi dari suatu data teks yang sifatnya kurang terstruktur (Asiyah & Fithriasari, 2016). Pengolahan menggunakan *text mining* ada beberapa fase diantaranya adalah:

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2.6.1 *Text Pre-processing*

Tahap *pre-processing* merupakan tahapan pertama untuk melakukan suatu filtrasi pada data teks yang terdiri dari langkah-langkah. Langkah awal pada *pre-processing* ialah melakukan *case folding* dan *removing*. *Case folding* merupakan tahapan mengubah semua huruf kapital dari abjad 'a-z' menjadi huruf kecil semua, sedangkan *removing* merupakan tahapan menghapus *noise* yang tidak diperlukan (Mujilahwati, 2016). Penghapusan yang dilakukan adalah:

1. *Remove Username* (@), bertujuan meghapus nama dari akun lain yang di *tag* pada data teks.
2. *Remove Hashtag* (#), bertujuan menghapus tanda pagar karena termasuk *noise* yang tidak diperlukan untuk ekstraksi data teks. Karena biasanya *hashtag* hanya digunakan untuk topik pembicaraan dan pengelompokkan pembicaraan.
3. Menghapus angka dan tanda baca, yaitu proses penghapusan angka dan tanda baca seperti titik (.), koma (,), tanda tanya (?), tanda petik (' '), dan garis miring (/) yang terdapat pada data teks.
4. *Clean One Character*, berfungsi menghilangkan satu huruf yang terdapat pada teks baik itu memang sudah ada dari saat awal pengambilan data, maupun huruf tunggal yang muncul setelah penghapusan tanda baca.
5. *Remove Stopwords*, penghapusan kata-kata yang tidak perlu seperti kata penguhubung dan kata-kata keterangan waktu maupun tempat.

Penghapusan *noise* dibutuhkan, agar tidak mengganggu *term* pada saat proses pembobotan (Kurniawan, et al., 2012).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tahap *pre-processing* berikutnya adalah tahap *tokenizing* yang merupakan proses mengubah kata dari suatu kalimat menjadi berpisah-pisah. Kemudian dilanjutkan dengan tahap *stemming*. Tahap ini menghilangkan kata himbuhan baik awalan, akhiran, maupun sisipan sehingga memperoleh kata dasarnya saja (Sarwani & Mahmudy, 2015).

2.6.2 Pembobotan Data

Setelah melakukan semua langkah pada tahap *pre-processing*, selanjutnya hasil data diubah dalam bentuk *term* matriks untuk dilakukan pembobotan. Pembobotan yang dilakukan bisa menggunakan pembobotan *TF-IDF* (*term frequency inverse document frequent*) (Asiyah & Fithriasari, 2016). Tahap *TF* merupakan cara menemukan jumlah kata yang ada pada dokumen. Pencocokkan dengan kata yang sama untuk melihat seberapa banyak kata tersebut muncul merupakan tahap *IDF*. Persamaan (2.1) dan (2.2) merupakan rumus yang digunakan untuk melakukan pembobotan.

$$w_{ij} = TF_{ij} \times IDF \tag{2.1}$$

$$IDF = log\left(\frac{N}{DF_j}\right) \tag{2.2}$$

keterangan:

i = 1,2, ..., p (jumlah variabel)

j = 1,2, ..., N (Jumlah data)

Untuk w_{ij}sebagai bobot kata (i) pada dokumen (j), sedangkan N adalah jumlah keseluruhan data. Untuk TF_{ij} merupakan banyaknya kata (i) yang muncul pada

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dokumen (j) dan DF_j sebagai jumlah dokumen (j) yang mengandung kata (i). Dalam melakukan pembobotan, $TF\ IDF$ dapat memberikan informasi penting mengenai nilai frekuensi kata yang muncul pada suatu dokumen.

2.7 Analisis Sentimen

Analisis sentimen merupakan studi fokus komputasi untuk mengekspresikan sikap, opini, pendapat, emosi, subjektifitas, pandangan dan penilaian seseorang yang dilihat terhadap apa yang ditulis (Liu, 2012). Pengetahuan tentang opini yang diberikan seseorang kini penting diketahui. Dalam dunia industri pun, sebagian besar penyedia jasa memperhatikan opini masyarakat terhadap produk yang dijualnya. Penyedia jasa sangat memperhatikan opini masyarakat terkait produknya, sehingga dapat menyesuaikan dan mengembangkan produk yang agar dapat diterima oleh masyarakat dan penyedia jasa dapat memperoleh keuntungan yang cukup besar. Selain itu, analisis sentimen juga dapat digunakan dalam memprediksi kemenangan calon legislatif melalui *twitter* dari beberapa *tweet* selama masa kampanye (Hakimi, 2018).

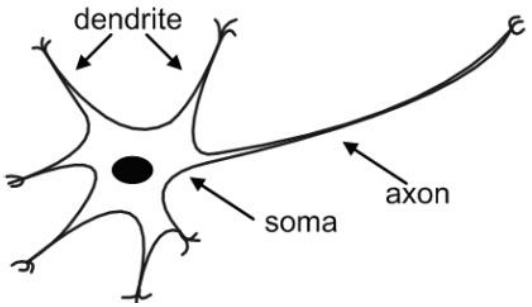
Pengetahuan mengenai opini seseorang kini dapat melalui media sosial, dengan melibatkan postingan yang seseorang sering tulis di akun media sosialnya. Analisis sentimen dapat dilakukan dengan hanya melibatkan media sosial yang sering digunakan oleh seseorang tersebut seperti *twitter*, *facebook*, *blog*, dan beberapa media sosial lainnya. Analisis sentimen dapat dilakukan dengan melakukan pendekatan model *bag of words*, yang mana untuk setiap data dokumen akan diubah menjadi fitur vektor yang nantinya akan diteruskan pada algoritma klasifikasi (Araque, et al., 2016). Analisis yang dilakukan berupa opini atau argumen

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

diidentifikasi berdasarkan opini yang sudah ditentukan sebelumnya. Secara umum, analisis sentimen menggunakan dua tipe teks informasi yaitu fakta dan opini. Opini-opini yang dianalisis akan diklasifikasi berdasarkan data dokumen yang mengandung fakta (Azizan & Aziz, 2017). Dokumen yang dianalisis bukan hanya dokumen terbatas dalam Bahasa Inggris, melainkan bahasa apapun yang dapat diterjemahkan (Ozturk & Ayvaz, 2017).

2.8 *Artificial Neural Network*

Artificial neural network merupakan jaringan saraf tiruan berbasis komputasi yang pertamakali di kemukakan pada tahun 1943 oleh Warren MacCulloh bersama dengan Walter Pitts. MacCulloh merupakan seorang ahli saraf dan Pitts merupakan seorang ahli logika (Suzuki, 2013)**.** Keduanya bekerjasama merancang model perhitungan sederhana yang menyatukan keilmuan biologi terkait jaringan saraf pada manusia dengan ilmu logika dalam menyelesaikan suatu permasalahan. Jaringan saraf tiruan yang dibuat merupakan kecerdasan buatan berbasis komputasi yang dianalogikan seperti jalannya sistem jaringan saraf manusia yang bekerja seperti pada Gambar 2.1.



Gambar 2.1 Jaringan Saraf Manusia (Krenker, et al., 2011)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Secara biologis, jaringan saraf manusia memiliki *neuron* yang berperan aktif menerima dan menyalurkan *impuls* dari *neuron* lain melalui dendrit dan mengirim sinyal melalui akson oleh badan sel. Akson bercabang dan terhubung dengan dendrit sel saraf lain dan mengirimkan *impuls* melalui sinapsis (Istiadi, 2018). Kunci dari metode ini adalah *neuron* yang bekerja sebagai pusat pengaturan dan informasi yang ada pada sel jaringan saraf manusia sehingga memberikan perintah pada tubuh untuk melakukan suatu gerakan. Jaringan saraf tiruan yang dianalogikan berdasarkan jaringan saraf manusia memiliki beberapa sifat (Rinzel & Ermentrot, 1989) seperti berikut:

a) Aktivasi *neuron* merupakan semua proses baik itu pada masa kinerja atau tidak
b) Sejumlah sinapsis tetap dalam keadaan kinerjanya
c) Keterlambatan sinaptik adalah satu-satunya keterlambatan signifikan dalam sistem jaringan saraf
d) Aktivitas sinapsis yang terlambat mencegah *neuron* dalam kinerjanya
e) Struktur jaringan tidak berubah sepanjang waktu

Adapun model matematika yang mewakili karakteristik tersebut tampak seperti Persamaan (2.3).

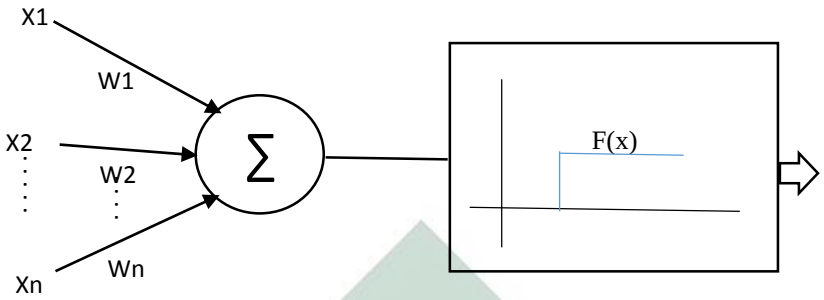
$$a = \sum_{j=1}^{p} x_j w_k \tag{2.3}$$

Dimana x_j = *input* biner ke *j*

w_k = bobot simpatik

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Logika yang membangun Persamaan (2.3) ditunjukkan seperti Gambar 2.2



Gambar 2.2 Fungsi ANN (sumber: (Gurney, 1997))

Bobot sinapis w_n merupakan bilangan real yang menjadi nilai dari suatu sinapsis yang terhubung. Sinapsis memiliki kemampuan menghambat sinyal negatif atau mengaktifkan sinyal posistif. Selain itu, sinapsis juga memiliki intensitas yang berbeda. Sedangkan sinyal a merupakan *input* fungsi aktivasi F(x) yang memberikan *output neuron*.

Secara umum jaringan saraf tiruan yang diperkenalkan oleh MacCulloh dan Pitts merupakan model matematika yang menghubungkan beberapa *neuron* agar terbentuk suatu simulasi untuk melakukan operasi aritmatika sederhana (Leavline, 2015). Oleh karena itu, pada tahun 1958 Rosn Boltt mengembangkan model jaringan untuk mengoptimalkan hasil iterasinya. Model pengembangan ini disebut dengan perceptron yang memiliki fungsi berikut:

a) Memetakan pola yang didapat dari *input* kedalam pola pada *output*
b) Penyimpanan pola akan dipanggil kembali
c) Prediksi dan klasifikasi pola

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

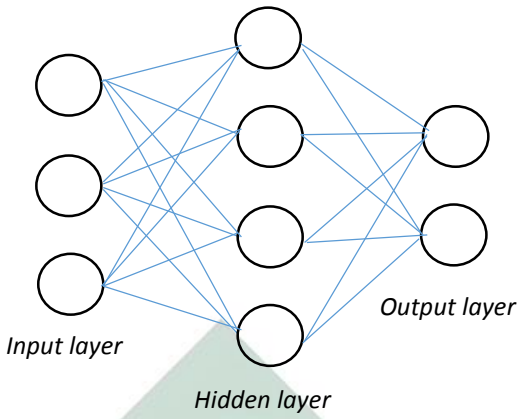
Perkembangan jaringan saraf tiruan berlanjut pada penelitian yang dilakukan oleh Hopfield pada tahun 1982. Hopfield berhasil menambahkan fungsi energi ke ANN untuk perhitungan suatu objek berdasarkan objek yang sudah dikenal sebelumnya. Tidak hanya berhenti disini, peningkatan jaringan saraf tiruan terus berlanjut, hingga pada tahun 1986 beberapa peneliti Rumelhart, Hinton, dan William menciptakan suatu algoritma yang disebut dengan *backpropagation*. Algoritma ini merupakan algoritma perbaikan dan menutupi kekurangan yang tidak dimiliki oleh algoritma perceptron (Effendi, 2013).

2.8.1 Klasifikasi ANN

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 seperti yang terlihat pada Gambar 2.3.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 2.3 Arsitektur Klasifikasi ANN

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 signifikasi variasinya tergantung pada permasalahan dan subjek yang berbeda (Bashashati, et al., 2015).

2.8.2 Arsitektur *Artificial Neural Network*

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

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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.

2.8.3 *Backpropagation*

Backpropagation terkenal dengan metode yang memiliki pelatihan dengan tipe arsitek *multi layer* dengan akurasi dan fleksibelitas 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 medapatkan model yang optimal (Effendi, 2013).

Berbeda dengan jaringan saraf tiruan tunggal, *backpropagation* menambahkan *hidden layer* dalam melatih jaringan untuk memperoleh *output* yang sesuai. Untuk

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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 *output*nya. 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).

$$L(t, y) = (\mathrm{t} - \varphi)^2 = \varepsilon \tag{2.4}$$

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).

$$E(n) = \frac{1}{2}\sum_y \varepsilon_y^2(n) \tag{2.5}$$

atau

$$E(n) = \frac{1}{2}\sum_{y=1}^{n}(t_y - \varphi_y)^2 \tag{2.6}$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Untuk mempermudah pemahaman digunakan beberapa notasi yang dipakai dalam pembahasan perhitungan *error backpropagation*.

$E(n)$= *error* data ke- n

t_y = nilai target

x_j = *input* unit ke-j

v_{jk} = bobot yang terhubung dengan *input* ke- j pada unit k (pada unit *hidden*)

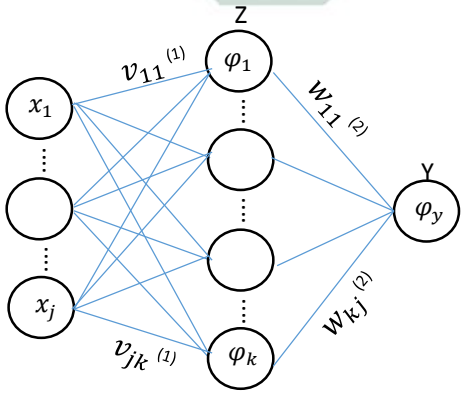
w_{kj} = bobot yang terhubung dengan *output* ke- k pada unit j (pada unit *output*)

Z_{in} = nilai *input* untuk *hidden layer*

φ_k = nilai *output* pada setiap *node hidden*

φ_y = nilai *output* pada *layer output*

Inisialisasi bobot v_{jk} dan w_{kj} dibutuhkan untuk langkah awal pengerjaan *backpropagation*. Arsitektur pada Gambar 2.4 berikut menunjukan alur perjalanan *neuron* dari sinyal *input* menuju *output*, hingga kembali lagi ke *input* dan dilakukan pengecekan kesalahan.



Gambar 2.4 Arsitektur untuk Alur *Backpropagation*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Berdasarkan Gambar 2.4, proses *feedforword* dimulai dengan mencari nilai Z_{in} yang diperoleh dari hasil kali antara nilai *input* dengan bobot *input* (v_{jk}) seperti pada Persamaan (2.7)

$$Z_{in} = xv_{jk}{}^{(1)} \tag{2.7}$$

Perhitungan dilanjutkan dengan mendapatkan hasil *output* dari *node hidden* dengan menggunakan fungsi aktifasi pada Persamaan (2.8) terhadap nilai Z_{in} yang diperoleh.

$$\varphi_k = f(Z_{in}) = \frac{1}{1+e^{-Z_{in}}} \tag{2.8}$$

Selanjutnya nilai φ_k menjadi nilai *input* untuk perhitungan *output layer* yang dikalikan dengan bobot *output* w_{kj}, seperti pada Persamaan (2.9).

$$y_k = \varphi_k w_{kj} \tag{2.9}$$

Fungsi aktifasi digunakan kembali untuk mendapatkan hasil *output* pada *layer output* sebagai keluaran akhir dari proses *feedforword*.

$$\varphi_y = f(y_k) = \frac{1}{1+e^{-y_k}} \tag{2.10}$$

Tahap selanjutnya adalah tahap *backpropagation error*. Tahap ini merupakan tahap mencari nilai *error* atau nilai perubahan bobot dengan menghitung mundur dari *output layer* ke *input layer*. Dengan memperhatikan persamaan *error* yang dimiliki pada Persamaan (2.7), diperoleh bahwa perubahan bobot pada *layer output* (Δw_{kj}) merupakan *sum square error* dari hasil selisih antara t_y (*output* target) dan φ_y (*output* prediksi) yang diturunkan pada setiap bobot *output*nya, maka substitusikan Persamaan (2.6) terhadap Persamaan (2.11)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$\Delta w_{kj} = \frac{\partial E(1)}{\partial w_{kj}} \tag{2.11}$$

$$\frac{\partial E(1)}{\partial w_{kj}} = \frac{\frac{1}{2}(t_k - \varphi_y)^2}{\partial w_{kj}}$$

$$\frac{\partial E(1)}{\partial w_{kj}} = -(t_k - \varphi_y)\frac{\partial \varphi_y}{\partial w_{kj}} \tag{2.12}$$

Karena $\varphi_y = f(y_k)$, maka

$$\frac{\partial E(1)}{\partial w_{kj}} = -\left(t_k - \varphi_y\right)\frac{\partial \varphi_y}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}}$$

$$\frac{\partial E(1)}{\partial w_{kj}} = -\left(t_k - \varphi_y\right)f'(y_k).\frac{\partial y_k}{\partial w_{kj}} \tag{2.13}$$

untuk $\frac{\partial y_k}{\partial w_{kj}}$ sama dengan aktivitas setiap sinapsis pada daerah *output* setelah layar *hidden*.

$$\frac{\partial E(1)}{\partial w_{kj}} = \begin{bmatrix} -t_1 - \varphi_1 \\ . \\ . \\ -t_k - \varphi_y \end{bmatrix} \begin{bmatrix} f'(y_1) \\ . \\ . \\ f'(y_k) \end{bmatrix} \begin{bmatrix} w_{11} \\ . \\ . \\ w_{kj} \end{bmatrix} \tag{2.14}$$

$$\begin{bmatrix} -t_1 - \varphi_1 \\ . \\ . \\ -t_k - \varphi_y \end{bmatrix} \begin{bmatrix} f'(y_1) \\ . \\ . \\ f'(y_k) \end{bmatrix} = \begin{bmatrix} \delta_1 \\ . \\ . \\ \delta_y \end{bmatrix} = \delta_y \tag{2.15}$$

dengan δ_y merupakan faktor kesalahan yang diperoleh dari *layer output* ke *layer hidden*, sedangkan $\begin{bmatrix} w_{11} \\ . \\ . \\ w_{kj} \end{bmatrix}$ merupakan aktivitas *neuron* pada setiap sinapsisnya,

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

sehingga untuk matriks $\begin{bmatrix} w_{11} \\ \cdot \\ \cdot \\ w_{kj} \end{bmatrix}$ ditranspose agar dapat dioperasikan dengan matriks δ_y, maka didapat Persamaan (2.16)

$$\frac{\partial E(1)}{\partial w_{kj}} = [w_1 \quad . \; . \; . \quad w_k] \begin{bmatrix} \delta_1 \\ \cdot \\ \cdot \\ \delta_y \end{bmatrix} = \begin{bmatrix} w_1 \delta_1 \\ \cdot \\ \cdot \\ w_k \delta_y \end{bmatrix} \tag{2.16}$$

$$\frac{\partial E(1)}{\partial w_{kj}} = {w_{kj}}^T \delta_y \tag{2.17}$$

$$\frac{\partial E(1)}{\partial w_{kj}} = {w_{kj}}^T (t_k - \varphi_y) f'(y_k) \tag{2.18}$$

Persamaan (2.17) merupakan perubahan bobot untuk *layer hidden*. Untuk nilai Y biner secara umum biasanya menggunakan fungsi aktivasi sigmoid. Untuk mendapatkan nilai kesalahan (δ_k) dari bobot *output*, maka nilai y_k pada Persamaan (2.19) diturunkan terhadap fungsi aktifasinya menggunakan turunan parsial.

$$\varphi_y = f(y_k) = \frac{1}{1+e^{-y_k}} \longrightarrow \frac{u'v - uv'}{v^2} \tag{2.19}$$

dengan $u = 1$, $u' = 0$

$v = 1 + e^{-y_k}$ $v' = e^{-y_k}$

maka

$$\varphi'_y = f'(y_k) = \frac{0(1+e^{-y_k}) - 1(-e^{-y_k})}{(1+e^{-y_k})^2}$$

$$\varphi'_y = f'(y_k) = \frac{1}{1+e^{-y_k}} \cdot (1 - \frac{1}{1+e^{-y_k}})$$

$$\varphi'_y = f'(y_k) = f(y_k)(1 - f(y_k))$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$\varphi'_y = f'(y_k) = \varphi_y\,(1 - \varphi_y) \tag{2.21}$$

Sehingga diperoleh

$$\delta_y = \left(t_k - \varphi_y\right)\varphi_y\,(1 - \varphi_y) \tag{2.22}$$

Persamaan untuk *single error* pada perhitungan *backforward* di daerah *output* ke *hidden* telah didapatkan. Langkah selanjutnya ialah mencari *single error* dari *layer hidden* ke *layer input*. Persamaan (2.17) merupakan hasil perubahan bobot dari *layer output* sebagai nilai node pada *hidden*. Untuk mencari perubahan bobot pada *layer input* (Δv_{jk}), maka hasil perubahan bobot Δw_{kj} dikalikan dengan fungsi aktifasinya yang diturunkan terhadap v_{jk} sebagai bobot *input*.

$$\Delta v_{jk} = \frac{\partial E(1))}{\partial v_{jk}} = \begin{bmatrix} \frac{\partial E}{\partial v_{11}} & \cdots & \frac{\partial E}{\partial v_{1k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial E}{\partial v_{j1}} & \cdots & \frac{\partial E}{\partial v_{jk}} \end{bmatrix} \tag{2.23}$$

$$\frac{\partial E(1)}{\partial v_{jk}} = {w_{kj}}^T \delta_y \frac{\partial \varphi_k}{\partial v_{jk}} \tag{2.24}$$

Karena $\varphi_k = f(Z_{in})$, maka

$$\frac{\partial E(1)}{\partial v_{jk}} = {w_{kj}}^T \delta_y \frac{\partial \varphi_k}{\partial Z_{in}} \cdot \frac{\partial Z_{in}}{\partial v_{jk}}$$

$$\frac{\partial E(1)}{\partial v_{jk}} = {w_{kj}}^T \delta_y f'(Z_{in}) . \frac{\partial Z_{in}}{\partial v_{jk}} \tag{2.25}$$

Nilai $\frac{\partial z}{\partial v_{jk}}$ sama dengan aktivitas pada setiap sinapsis layar *input*, maka $\frac{\partial z}{\partial v_{jk}}$ merupakan nilai bobot yang ditranspose sama dengan proses pada Persamaan (2.16) dan Persamaan (2.17)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$\frac{\partial E(1)}{\partial v_{jk}} = {w_{kj}}^T \delta_y \, f'(Z_{in}) v^T \tag{2.26}$$

dan ${w_{kj}}^T \delta_y \, f'(Z_{in})$ merupakan faktor kesalahan pada *single error hidden layer* ke *input layer* (δ_k), maka diperoleh *single error* atau perubahan bobot untuk sinyal *input* pada Persamaan (2.27).

$$\frac{\partial E(1)}{\partial v_{jk}} = V^T \delta_k \tag{2.27}$$

Kedua perubahan bobot yang didapatkan ini (Δv_{kj} dan Δw_{kj}), selanjutnya akan digunakan untuk mencari *update* bobot terbaru pada iterasi selanjutnya. Pengerjaan mencai *update* bobot pada tahap *backpropagation error* ini akan berhenti jika nilai *output* φ_y mendekati nilai target t_y. Persamaan (2.28) menunjukkan *update* bobot *input* dan Persamaan (2.29) menunjukkan *update* bobot *output*.

$$v_{jk}(baru\ (n)) = v_{jk}(lama\ (n)) + \Delta v_{jk}(n) \tag{2.28}$$

$$w_{kj}(baru\ (n)) = w_{kj}(lama\ (n)) + \Delta w_{kj}(n) \tag{2.29}$$

Metode *backpropagation* dapat dilakukan dengan langkah-langkah menggunakan empat tahapan pokok yang dapat digunakan untuk perhitungan manual (Sutikno, et al., 2016)

a) Langkah 0 : Inisialisasi bobot dengan mengambil nilai *random*

Langkah 1 : Jika salah dalam mengambil langkah henti, maka lakukan langkah 2-9

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Langkah 2 : Lakukan langkah 3-8 pada masing-masing pasangan *training*

b) *Feedforward*

Langkah 3 : Setiap unit *input* menerima sinyal x_j (j = 1,2,3, ...n) akan diteruskan ke semua unit yang berada di lapisan tersembunyi (*layer* sesudahnya)

Langkah 4 : Setiap unit pada lapisan tersembunyi menjumlahkan sinyal-sinyal *input* bersama biasnya berdasarkan Persamaan (2.30) yang sesuai dengan Persamaan (2.7) setelah itu hitung *output* dari *hidden* dengan persamaan aktivasi yang digunakan pada Persamaan (2.8)

$$Z_{in} = v_{0k} + \sum_{j=1}^{n} x_j v_{jk} \tag{2.30}$$

Langkah 5 : Menghitung semua keluaran jaringan di unit y_k bersama dengan biasnya (k= 1,2, ..., m) dengan menggunakan Persamaan (2.31). Kemudian gunakan kembali fungsi aktivasi untuk mendapatkan nilai *output*.

$$y_k = w_{k0} + \sum_{i=1}^{n} \varphi_k w_{kj} \tag{2.31}$$

c) *Backpropagation*

Langkah 6 : Hitung faktor δ sebagai unit keluaran menggunakan Persamaan (2.22).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$\delta_y = (t_y - \varphi_y)\varphi_y(1 - \varphi_y)$$

dimana δ_y sebagai unit *error* yang nantinya akan dipakai pada perubahan nilai bobotnya pada *layer* sebelumnya (langkah 7). Kemudian suku perubahan bobot dihitung dengan Persamaan (2.32)

$$\Delta w_k = \alpha\ \delta_y \varphi_y \tag{2.32}$$

Langkah 7 : Selanjutnya menghitung faktor unit tersembunyi (δ) berdasarkan *error* yang ada pada setiap unit tersembunyi menggunakan Persamaan (2.33).

$$\delta_{net\,j} = \delta_k w_{kj} \tag{2.33}$$

Kemudian hitung suku perubahan bobot menggunakan Persamaan (2.34) untuk dipakai pada fase selanjutnya.

$$\Delta v_{jk} = \alpha\ \delta_k v_{jk} \tag{2.34}$$

d) Penyesuaian Bobot:

Langkah 8 : Menghitung semua perubahan bobot mulai dari bobot garis yang menuju unit keluaran menggunakan Persamaan (2.35) dan dari bobot garis yang menuju unit tersembunyi menggunakan Persamaan (2.36).

$$w_{kj}(baru) = w_{kj} + \Delta w_{kj} \tag{2.35}$$

$$v_{jk}(baru) = v_{jk} + \Delta v_{jk} \tag{2.36}$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dengan, k = 1, 2, ..., m

j = 0, 1, ..., n

Langkah 9 : Jaringan dipakai untuk mengetahui pola sehingga keluaran jaringan dapat ditentukan dengan propagasi maju (*feedforward*). Fungsi aktivasi juga harus ditentukan apakah menggunakan sigmoid biner atau tidak agar proses pada langkah (4 dan 5) serta (6 dan 7) sesuai (Siang, 2005).

2.8.4 Uji Coba Model ANN

Uji coba metode ANN merupakan pengujian untuk mendapatkan model atau arsitektur jaringan terbaik. Arsitektur atau model jaringan yang baik adalah model jaringan yang menghasilkan *error* terkecil pada saat proses *training* dilakukan dan mendapatkan akurasi tinggi pada proses *training-testing*. Uji coba ANN dilakukan dengan menggunakan penyesuaian dari perbandingan antara nilai *learning rate* dan jumlah *node hidden* yang digunakan (Yanti, 2011). *Learning rate* sendiri merupakan parameter untuk menghitung nilai koreksi bobot dengan nilai *rate* yang dapat digunakan adalah berkisar 0-1. Semakin besar nilai *rate* yang digunakan, maka akan semakin cepat proses *training* yang terjadi. Namun, pengambilan nilai *rate* yang terlalu besar akan menyebabkan nilai *error* yang besar juga. Untuk itu untuk mendapatkan nilai *error* yang kecil, harus melakukan beberapa iterasi untuk mendapatkan *error* yang paling minimal. Hal ini dikarenakan nilai *rate* mempengaruhi ketelitian pada suatu sistem.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2.9 Evaluasi Sistem

Evaluasi sistem diperlukan untuk mendapat akurasi yang baik dari suatu proses klasifikasi atau prediksi. Akurasi merupakan bagian penting pada akhir suatu pengolahan untuk mendapatkan jangkauan yang mendekati keadaan yang sebenarmya. Nilai akurasi biasanya muncul bersamaan dengan presisi sebagai penguatan bahwa keadaan suatu model yang didapatkan menggambarkan keseragaman (Hanifah & Prastowo, 2016).

Jenis akurasi yang digunakan sesuai dengan metode dan permasalahan yang sedang diteliti. Model prediksi maupun klasifikasi yang akan dilakukan memiliki tingkat akurasi yang berbeda pula (Jayanti & Rustiana, 2017). Adapun untuk evaluasi sistem klasifikasi dapat dilakukan dengan menggunakan *confusion matrix*. *Confusion matrix* memiliki empat jenis evaluasi, diantaranya yaitu *true positive* merupakan kelas, *true negative*, *false positive*, dan *false negative*. Empat evaluasi ini dibedakan menjadi dua kelas seperti pada Tabel 2.1.

Tabel 2.1 Contoh *Confution Matrix* dengan Dua Kelas

Class		*True Class*	
		Positive	*Negative*
Predictive Class	*Positive*	*True Positive (TP)*	*False Positive (FP)*
	Negative	*False Negative (FN)*	*True Negative (TN)*

Adapun dari keempat jenis evaluasi ini, digunakan untuk melakukan evaluasi sistem klasifikasi yaitu akurasi, presisi, dan *recall*, dan *f-measure*. Akurasi merupakan perbandingan dari jumlah data yang diklasifkasi secara benar dibagi

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dengan semua jumlah data. Untuk nilai presisi diperoleh dari perbandingan antara *true positive* dengan jumlah data yang diklasifikasi memiliki nilai positif. *Recall* merupakan perbandingan dari *true positif* terhadap semua jumlah yang terkategori positif. Sedangkan *f-measure* diperoleh dari nilai rata-rata dari jumlah presisi dan *recall*. Adapun rumus sederhana untuk menentukan keempat evaluasi tersebut seperti pada Persamaan (2.43) sampai Persamaan (2.46).

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \tag{2.37}$$

$$Presisi\ Positif = \frac{TP}{TP+FN} \tag{2.38}$$

$$Presisi\ Negatif = \frac{TN}{TN+FP} \tag{2.39}$$

$$\text{Rata} - \text{rata Presisi} = \frac{Presisi\ positif + presisi\ negatif}{2} \tag{2.40}$$

$$Recall\ Positif = \frac{TP}{TP+FP} \tag{2.41}$$

$$Recall\ Negatif = \frac{TN}{TN+FN} \tag{2.42}$$

$$Rata - rata\ Recall = \frac{Recall\ positif + Recall\ negatif}{2} \tag{2.43}$$

$$F - Measure = \frac{Presisi + Recall}{2} \tag{2.44}$$

dimana:

$TP = true\ positive$ (klasifikasi yang tepat pada kelas positif)

$TN = true\ negative$ (klasifikasi yang tepat pada kelas negatif)

$FP = false\ positive$ (klasifikasi yang tidak tepat pada kelas positif)

$FN = false\ negative$ (klasifikasi yang tidak tepat pada kelas negatif)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB III

METODE PENELITIAN

3.1 Jenis Penelitian

Penelitian ini melakukan penelitian kuantitatif yang sudah terencana secara sistematis agar mendapatkan solusi terbaik dari permasalahan yang diteliti. Desain penelitian yang digunakan merupakan analisis sentimen dari permasalahan yang telah ditetapkan. Analisis dilakukan setelah klasifikasi dan ketepatan pegujian model terbaik didapatkan. Penelitian ini merupakan penelitian yang bersifat analisis dari pengolahan data sekunder yang diperoleh tanpa eksperimen secara langsung dilapangan.

3.2 Sumber Data

Data yang digunakan pada penelitian ini merupakan data sekunder tentang opini atau postingan seseorang di *twitter* mengenai radikalisme yang memprovokasi islam radikalis. Data diambil secara daring dari situs resmi www.api.*twitter*.com berdasarkan *mention account* dengan kata "khilafah" yang terdiri dari kalimat positif dan negatif. Data diperoleh dengan melakukan proses *web crawling* dengan menggunakan aplikasi *Netbeans*. *Crawling* dilakukan dengan memilih kata kunci dari data yang ingin ditambang. Karena pada penelitian ini menggunakan kata "khilafah", maka *crawling* dilakukan dengan menyebutkan satu kata tersebut. Data yang didapat merupakan data teks berupa kalimat dari postingan beberapa akun *twitter* seperti pada Tabel 3.1. Banyaknya data yang ditambang yaitu sebanyak ribuan *tweet.* Sebagian data *tweet* digunakan untuk *training* dan sisanya digunakan sebagai data *testing*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Kategori nilai sentimen berupa kalimat positif dan negatif, diambil dengan perbandingan 2:3. Data disusun secara acak tanpa memperhatikan keterurutan dari tanggal, akun, *retweet* ataupun yang lainnya.

Tabel 3.1. Sampel data *tweet* Berdasarkan *Mention Account*

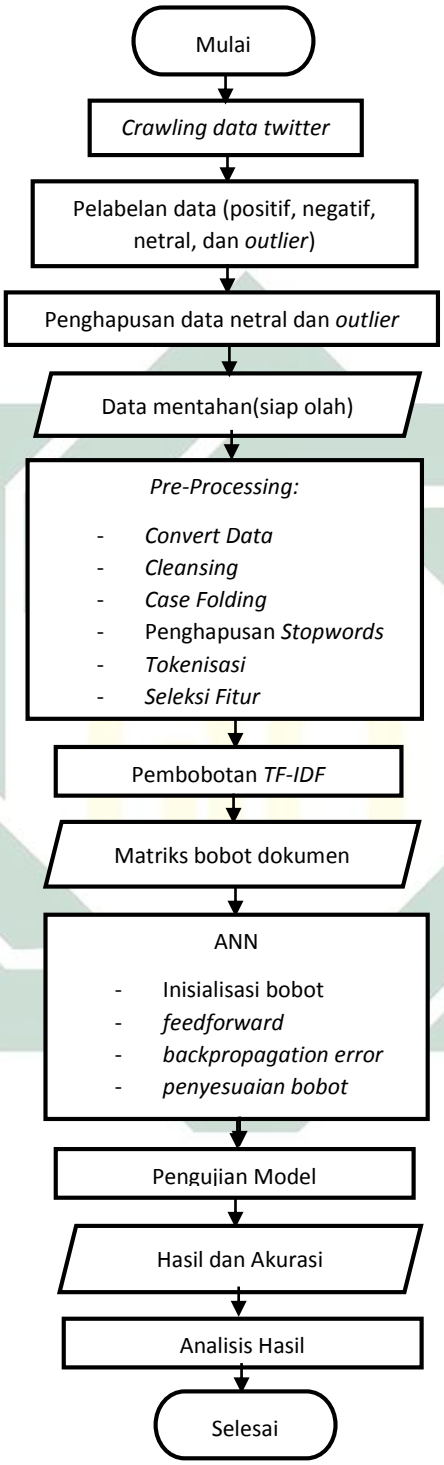
No	Akun	Kategori	*Tweet*
1.	@OktarinaMahard1	Negatif	Menuju persatuan umat yg hakiki dengan khilafah..Insyaa Allah. Allahu Akbar
2.	@muslimpribumi	Positif	Pendukung khilafah banyak yg gak bisa ngaji. Itu fakta!!!!
3.	@Dennysiregar7	Negatif	Saya sebenarnya jijik kalau tim @jokowi bawa2 narasi agama dalam kampanye. Tapi saya lebih jijik kalau pendukung @prabowo…
4.	@Airin_NZ	Positif	Ingin mnerapkan Indonesia Bersyariah. Tp menolak Baca Al-qur'an
5.	@RokhmatLabib	Negatif	Seorang kacung takan mungkin menkritik majikannya. Sudah khilafah saja...

3.3 Pengolahan Data

Data yang sudah diperoleh selanjutnya disimpan dalam format .xls dan diubah menjadi .csv. Pegubahan format ini bertujuan agar lebih mudah dalam pengolahan di bahasa R ketika mengubah data *word* menjadi *number*. Data yang sudah terkumpul, diolah menggunakan metode *text mining* yang di dalamnya terdiri dari tahap *pre-processing*, pembobotan data, hingga mendapatkan data *number* dalam bentuk matriks. Selanjutnya akan diteruskan pada tahap klasifikasi menggunakan metode ANN.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

3.4 Rancangan Sistem



Gambar 3.1 *Flowchart* Pengolahan Data

Pengolahan data berdasarkan Gambar 3.1 menjelaskan bahwa tahap-tahap yang dilalui oleh data diawali dengan tahap *pre-processing* yang di dalamnya melakukan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kegiatan *convert* data dari data mentahan. Data difilter dengan menghilangkan tanda baca, angka, dan *emoticon*, agar dapat diekstraksi fiturnya. Proses *stemming* pada tahap *pre-processing* tidak dilakukan karena peneliti menginginkan kata yang diperoleh dari tahap ini tidak berupa kata dasar, melainkan cukup dengan pemisahan kata yang sebenarnya dari proses tokenisasi suatu dokumen. Hal ini bertujuan agar pada saat tahap pembobotan kata, antara *term* dan dokumen sesuai. Proses pembobotan kata dilakukan dengan perhitungan TF-IDF. TF-IDF bertujuan memberikan pembobotan pada suatu dokumen untuk menentuan frekuensi kata pada dokumen yang disesuaikan dengan *term* yang sudah terpilih hingga menghasilkan data vektor berupa matriks.

Data dibagi menjadi dua bagian, kelompok pertama membagi 80% data *training* dan data 20% *testing*, sedangkan untuk kelompok kedua membagi 60% data *training* dan 40% data *testing*. Model dari data *training* dibangun menggunakan metode ANN dengan melalui empat tahapan dan menghasilkan bobot JST, sedangkan data *testing* digunakan sebagai sentimen aktual sebagai acuan pengujian pembangunan model.

3.4.1 *Crawling* Data *Twitter*

Data yang dibutuhkan dalam penelitian ini merupakan data sekunder berupa *tweet* di *twitter* yang diperoleh dengan cara *crawling* data *twitter API*. Untuk melakukan *crawling* data, dilakukan registrasi terlebih dahulu agar mendapat kode akses di *twitter API* melalui situs https://dev.*twitter*.com/apps/new. Setelah kode didapatkan, *crawling* data dapat dilakukan dengan menggunakan *software Netbeans* sesuai dengan kata kunci yang sudah dijelaskan sebelumnya untuk

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

mendapatkan data *tweet* yang diinginkan. saat melakukan *crawling*, pastikan internet aktif pada komputer yang melakukan *crawling*. Karena *crawling* yang dilakukan menggunakan *software* Netbeans harus dalam keadaan *online*.

3.4.2 Pelabelan dan Penghapusan Data

Data *tweet* yang sudah diperoleh selanjutnya disimpan dalam bentuk .csv. Kemudian dilakukan pelabelan secara manual yang sudah dikonsultasi terlebih dahulu kepada ahli bahasa dengan menunjukkan beberapa sampel. Data *tweet* dibedakan ke dalam empat kategori yaitu *tweet* yang mengandung sentimen positif, negatif, netral, dan yang tidak termasuk pada ketiganya (*outlier*). *Tweet* yang sudah mendapatkan label diperiksa kembali sentimennya kepada ahli dalam bidang keagamaan. Selanjutnya, dilakukan penghapusan data yang termasuk pada kategori netral dan *outlier*. Penghapusan data yang dilakukan dikarenakan tidak sesuai pada klasifikasi sentimen yang akan diteliti dan dapat mengganggu hasil klasifikasi.

3.4.3*Pre-Processing*

Pre-processing merupakan proses awal yang dilakukan pada *text mining* sebelum melakukan pembobotan terhadap data teks. Adapun prosesnya terdiri dari beberapa tahap berikut:

a) Convert data

Data yang tersimpan dalam bentuk format .xls diubah menjadi format .csv agar lebih mudah di proses dalam tahap selanjutnya pada *text mining*.

b) Cleansing

Cleansing merupakan proses penghapusan semua karakter web seperti html, www, http, dan karakter lainnya yang dianggap tidak memiliki makna

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dan tidak berguna jika diikutkan pada pengklasifikasian dokumen. Proses *cleansing* ini dilakukan karena terkadang pada suatu *tweet* di *twitter* menyertakan alamat webnya. Penghapusan dibutuhkan agar proses klasifikasi tidak terganggu. Pada tahap ini juga dilakukan tahap filtrasi data dengan menghilangkan komponen-komponen seperti tanda baca, angka, dan *emoticon* yang tidak dibutuhkan (Mujilahwati, 2016). Adapun komponen-komponen yang dihilangkan seperti pada Tabel 3.2.

Tabel 3.2. Komponen yang Dihapus

Angka yang hilang	Tanda baca yang hilang	Emoticon yang hilang
1	-	-_-
18	&	:’)
2	_	;)
20	|	:=)
2015	‘	☺
28	“	:’(
3	!	
37	?	
39	/	

c) Case-Folding

Setelah penghapusan karakter web, data yang sudah dikumpulkan dalam bentuk format .csv, selanjutnya dilakukan proses *case-folding* yaitu merubah huruf (A-Z) pada kalimat menjadi huruf kecil semua. Tahap ini bertujuan agar setiap kata dari semua dokumen sama dan tidak terjadi multi *term* pada saat tahap tokenisasi dilakukan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

d) Penghapusan *Stopwords*

Penghapusan *stopwords* merupakan penghapusan kosakata yang tidak memiliki makna penting dalam pengkategorian dokumen dan sering muncul dalam suatu kalimat seperti pada Tabel 3.3 berikut.

Tabel 3.3 Sampel Data *Stopword*

Kata penghubung	Kata tanya	Waktu
Dan	Apa	senin
Daripada	Bagaimana	Selasa
Dengan	Mengapa	Rabu
Disebut	Dimana	Kamis
Di	Kapan	Jumat
Ke	Siapa	Sabtu
Kepada		Minggu
Jika		Januari
Bagi		Februari
Sebagai		Maret
Seperti		April
walaupun		Mei
Yang		Juni
Yaitu		Juli
Yakni		Agustus

e) Tokenisasi

Tahap tokenisasi dilakukan untuk membagi dan memisahkan dokumen teks menjadi perkata atau token-token tertentu. Contohnya seperti salah satu dari sampel *tweet* pada Tabel 3.1 dokumen pertama, yaitu “menuju persatuan umat yg hakiki dengan khilafah insyaa allah allahu akbar” menjadi 11 token yang terdiri dari “menuju-persatuan-umat-yg-hakiki-dengan-khilafah-insyaa-allah-allahu-akbar”.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

f) Seleksi Fitur

Seleksi Fitur merupakan pemilihan kosakata dari hasil tokenisasi yang akan digunakan pada tahap klasifikasi. Kosakata dipilih secara manual dan diambil kosakata penting untuk dijadikan data *training* dan *testing* pada tahap klasifikasi.

3.4.4 Pembobotan TF-IDF

Kalimat yang sudah difilter akan dipisah menjadi kata-perkata atau beberapa token. Kemudian pembobotan TF-IDF dilakukan sebanyak kosakata yang dipilih dan diproses, hingga masing-masing kosakata mempunyai bobot sendiri-sendiri. Penentuan bobot yang dikenakan per kosakata dihitung berdasarkan Persamaan (2.1) dan Persamaan (2.2). Hasil dari pembobotan kosa kata ini akan membentuk suatu matriks seperti Tabel 3.4.

Tabel 3.4. Ilustrasi Hasil Pembobotan TF-IDF

Term	D1	D2	....	Dm
Tweet-1	.....	.....	.....	.....
Tweet-2	.....	.....	.....	.....
.....	.....	.....	.....	.....
Tweet-n	.....	.....	.....	.....

3.4.5 Klasifikasi *Artificial Neural Network* (ANN)

Hasil pembobotan yang sudah dalam bentuk matriks selanjutnya dilakukan klasifikasi dengan algoritma sederhana dari ANN berikut:

a) Melakukan inisialisasi pembobotan secara *random* atau acak

b) Menetapkan inisialisasi bobot dan jumlah *hidden layer* untuk semua percobaan pencarian model terbaik.

c) Melakukan perhitungan keluaran *neuron* dengan metode *feedforward*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

d) Menghitung kesalahan proses *feedforward* dengan fungsi aktivasi. Pada penelitian ini menggunakan fungsi aktivasi sigmoid biner.

e) Menghitung perubahan bobot dan bias dengan metode *backpropogation*.

f) Perhitungan kembali dengan nilai bobot dan bias baru

g) Ulangi langkah diatas hingga proses komputasi menghasilkan *error* dibawah batas minimum yang ditentukan.

3.4.6 Pengujian Model

Uji coba dilakukan untuk mendapatkan model terbaik untuk klasifikasi. Pengujian model dilakukan dengan melakukan uji coba jumlah *node* pada *hidden*, dua *hidden layer*, dan uji coba *learning rate* (0,1 – 0,5). Selain itu, pengujian dilakukan dengan membagi data pada saat melakukan proses *training* dan *testing*. grup pertama membagi data menjadi 80% untuk *training* dan 20% untuk *testing*, sedangkan grup kedua membagi data menjadi 60% *training* dan 40% *testing*. Pada masing-masing grup pembagian data tersebut dilakukan ketiga uji coba untuk mendapatkan model yang terbaik. Model yang dipilih adalah model yang memiliki nilai akurasi yang paling tinggi.

3.4.7 Analisis dan Visualisasi.

Setelah hasil dari proses klasifikasi menggunakan ANN didapatkan, selanjutnya melakukan analisis dan visualisasi dengan melakukan pembuatan *wordcloud* untuk mengetahui kosakata yang berpengaruh dengan frekuensi sesuai keinginan yang mau ditampilkan. Selain itu, analisis dan visualisasi juga dilakukan dengan melihat keterhubungan antar kata dari pembuatan graf jaringan dengan menggunakan aplikasi *VOSviewer*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB IV

HASIL DAN PEMBAHASAN

4.1 *Crawling Data*

Langkah pertama untuk penelitian ini adalah melakukan *crawling* data *tweet* dari *twitter*. Pengambilan data pada umumnya, melakukan *crawling* secara kontinu, yang dikarenakan keterbatasan ketersediaan proses *crawling* yang hanya bisa di*update* 7 hari terakhir dari waktu pencarian. *Crawling* data dilakukan secara langsung, dengan pengambilan data perhari dari tanggal 01 April 2018 sampai 30 Juni 2018. Data ini diambil bertepatan sebelum dan sesudah disahkannya salah satu organisasi masyarakat yang dinilai teridentifikasi menyimpang dari ideologi Indonesia yang terpengaruh islam radikal. Pengambilan data tanpa batasan waktu ini dapat dilakukan dengan *sourcode* dari GitHub menggunakan aplikasi *Netbeans* yang disimpan dengan format file .csv (*Comma Delimited*).

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*

D	R	F	Text
1	3	3	Saya tahu itu. Tapi coba lihat. Sikap para Petinggi di Muhammadiyah. Dan sikap Pemuda Muhammadiyah. Mereka cenderung Pro Khilafah. Konyol..!!
2	20	13	Buat yang bercita-cita dirikan khilafah, ingat Indonesia itu satu nusa, satu bangsa, satu bahasa...bukan satu agama! Bantu re*tweet* ya prof @na_dirs
3	59	75	Sistem Khilafah Tak Ada Faktanya dalam Sejarah Islam http://www.nu.or.id/post/read/77798/-sistem-khilafah-tak-ada-faktanya-dalam-sejarah-islam- …pic.*twitter*.com/6RnsEkWwcd
4	29	31	Kiai Pondok Pesantren Pancasila: Ide Khilafah Harus Dilarang http://regional.kompas.com/read/2017/06/01/14382321/kiai.pondok.pesantren.pancasila.ide.khilafah.harus.dilarang?utm_campaign=Kompascom&utm_medium=Social&utm_source=*Twitter* …
5	1	2	Untuk saat ini setuju karena kondisi Indonesia sudah dikuasai PMR (Pendidikan Moral Radikal) dan PMK (Pendidikan Moral Khilafah) https://*twitter*.com/Metro_TV/status/870164241860276224 …

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

D	R	F	Text
6	9	6	Munculnya kecintaan yang militan terhadap Pancasila hari-hari ini bisa dimaklumi, bahkan harus, karena Pancasila kita menghadapi ancaman nyata: Khilafah.
7	10	6	Masyarakat Bali Serahkan Dokumen Deklarasi Anti Khilafah ke Staf Khusus Presiden http://ow.ly/3JgR30cbazs #KASKUSNews
8	7	13	Itu hanya HALUSINASI.. ketika saya tanya sebutkan negara mana yang menerapkan khilafah dan bagaimana kehidupan rakyatnya? tidak ada yang bisa menjawab https://*twitter*.com/teh_tariq/status/870105587945586690 …
9	7	4	Pancasila itu sudah ada, kebetulan tidak konflik dengan humanitarian, jadi kalau memilih antara pancasila dan khilafah sepertinya saya harus memilih pancasila
10	20	22	bagi saya yang diam-diam / teriak mendukung khilafah, bubarkan NKRI itu sama buruknya dengan yang mendukung reklamasi, penggusuran, sama merusak pancasila https://*twitter*.com/Andry18916980/status/870071112759074816 …
11	2	1	Karena #PANCASILA sudah Khilafah yang sangat Islami, jadi kalau mau menegakkan khilafah tegakkan PANCASILA kita!
12	0	3	khilafah itu khayalan kaum extrimis yang ingin menghapus ideologi pancasila.
13	0	0	Berani berINDONESIA, harus berani berPANCASILA. Jangan mau ditipu pake khilafah atau NKRI Bersyariah. #HarlahPancasila #NKRIhargamati

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.

4.2 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.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Berikut contoh pelabelan terhadap data *tweet*:

Tabel 4.2 Contoh Pelabelan Data

No	Label kelas	Text
1	Positif	Survei: 79% WNI tidak Setuju NKRI Diganti Menjadi Khilafah
2	Negatif	Memperjuangkan Khilafah itu adalah kewajiban #khilafahajaranislam #ReturnTheKhilafa
3	Netral	Tulisan 4: Pancasila dan Khilafah
4	Outlier	suasana silaturahmi Hari Raya Idul Fitri ? 1439 H bersama masyarakat Indonesia di Tokyo - Jepang dan menikmati secuil rasa INDONESIA saat berlebaran di negeri orang Selamat Hari Raya Idul Fitri Minal Aidin Wal Faidzin Mohon Maaf Lahir dan Batin Meguro, 15 - 06 – 2018
5	Negatif	Jejak Khilafah di Nusantara #ReturnTheKhilafah

Setelah melakukan pelabelan kelas terhadap data *tweet*, diperoleh 671 data positif, 470 data negatif, 259 data netral, dan 78 data *outlier*. Langkah selanjutnya dilakukan tahap seleksi dengan melakukan penghapusan data *tweet* yang memiliki label netral dan *outlier*. Data *tweet* yang dimiliki hanya menyisakan data yang berlabel positif dan negatif saja. Hasil data yang diperoleh sebanyak 1141 *tweet* yang memiliki perbandingan 2:3 antara data yang berlabel positif dan negatif seperti pada Tabel 4.3.

Tabel 4.3 Sampel Data dengan Label Posiif dan Negatif

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

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

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

nilai dari masing-masing kata dari dokumen yang dimiliki. Pada tahap ini dilakukan *proses cleansing, case folding*, penghapusan *stopword*, 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 4.5Sampel 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

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

c) Penghapusan *stopword*

Tahap berikutnya merupakan tahap penghapusan *stopword*. Pada tahap ini semua kata pada kalimat yang mengandung kata *stopword* 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 *stopword*
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

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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 meuncul pada masing-masing dokumen seperti pada Tabel 4.8.

Tabel 4.8 Pembobotan TF

		Dokumen 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	Log(3/1) =0,477
Wni	1	Log(3/1) =0,477
tidak_setuju	1	Log(3/1) =0,477
Nkri	1	Log(3/1) =0,477
kewajiban	1	Log(3/1) =0,477
memperjuangkan	1	Log(3/1) =0,477
jejak	1	Log(3/1) =0,477
nusantara	1	Log(3/1) =0,477
returnthekhilafah	2	Log(3/2) =0,176

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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 seperti pada Tabel 4.10.

Tabel 4.10 Pembobotan TF IDF

Term	TF			DF	IDF	TF x IDF		
	Doc.1	Doc.2	Doc.3			Doc.1	Doc.2	Doc.3
Survei	1	0	0	1	0,477	0,477	0	0
Wni	1	0	0	1	0,477	0,477	0	0
tidak_setuju	1	0	0	1	0,477	0,477	0	0
Nkri	1	0	0	1	0,477	0,477	0	0
Kewajiban	0	1	0	1	0,477	0	0,477	0
memperjuangkan	0	1	0	1	0,477	0	0,477	0
Jejak	0	0	1	1	0,477	0	0	0,477
Nusantara	0	0	1	1	0,477	0	0	0,477
returnthekhilafah	0	1	1	2	0,176	0	0,176	0,176

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

input / dokumen	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	t_y
dokumen1	0,477	0,477	0,477	0,477	0	0	0	0	0	1
dokumen2	0	0	0	0	0,477	0,477	0	0	0,176	0
dokumen3	0	0	0	0	0	0	0,477	0,477	0,176	0

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Keterangan:

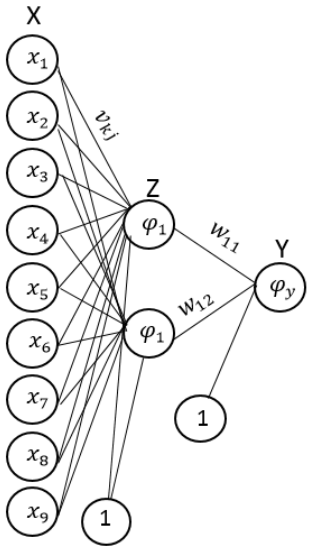
$t(1)$ = positif

$t(0)$ = negatif

X = *input*an

dokumen = dokumen (*tweets*) *input*

Berdasarkan Tabel 4.11, dapat dilakukan klasifikasi *backpropagation* secara manual, sebagai gambaran dalam perhitungan klasifikasi menggunakan program *Rstudio*. Dengan *input*an X_1 sampai X_9, digunakan satu *hidden layer* dengan 2 *node*. Terdapat sinapsis dengan bobot (v_{jk}) yang terhubung dari *node input* ke *node hidden*. Kemudian *neuron* akan dteruskan 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

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

manual keempat langkah yaitu inisialisasi bobot, *feedforward, backpropagation,* dan penyesuaian bobot.

a) Inisialiasi bobot

Misalkan mengambil 1 data yaitu dokumen1 pada Tabel 4.11 untuk hitung manual. Misalkan bias pada *node hidden* dan *output* adalah 1 dengan x_1 = 0,477, x_2 = 0,477, x_3 = 0,477, x_4 =0,477, x_5 = 0, x_6 = 0, x_7 = 0, x_8 = 0, x_9 = 0, dan t = 1, seperti pada Tabel 4.12 dan 4.13 berikut:

Tabel 4.12 Inisialisasi Bobot *Input*Iterasi Pertama

		Z_1 (*hidden*)	Z_2 (*hidden*)
Input	x_1	-0,09	-0,08
	x_2	-0,07	-0,06
	x_3	-0,05	-0,04
	x_4	-0,03	-0,02
	x_5	-0,01	0,3
	x_6	0,01	0,02
	x_7	0,03	0,04
	x_8	0,05	0,06
	x_9	0,07	0,08
	1	0,89	-0,1

Tabel 4.13 Inisialisasi Bobot *Output* Iterasi Pertama

		Y (*output*)
hidden	Z_1	-0,3
	Z_1	-0,13
	1	-0,36

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

b) Feedforward

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

Z_{in} = v_{j0} + $\sum_{j=1}^{j} x_j\, v_{jk}$

Z_{in_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

Z_{in_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 aktifasi, yaitu fungsi sigmoid kepada masing-masing sinyal *input*,

$\varphi_k = \frac{1}{1+e^{Z_in}}$

$\varphi_1 = \frac{1}{1+e^{-0,77752}}$ = 0,684714 $\varphi_2 = \frac{1}{1+e^{0,1954}}$ = 0,451305

sedangkan untuk nilai *output* sementara dapat diketahui dengan menghitung unit

y_k = w_{k0} + $\sum_{j=1}^{j} \varphi_k\, w_{kj}$

= -0,36 + (0,68714 x -0,3) + (0,451305 x -0,13)

= -0,62408

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

$\varphi_y = \frac{1}{1+e^{0,62408}}$ = 0,3488532

Nilai keluaran yang didapat untuk y hanya sebesar 0,3488532, sedangkan nilai t yang diinginkan adalah 1. Maka dengan ini, proses dilanjutkan dengan tahap

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

backpropagation untuk mencari nilai faktor kesalahan dan perubahan bobot untuk iterasi selanjutnya agar nilai y mendekati t.

c) Backpropagation

Langkah selanjutnya melakukan perhitungan factor δ sebagai unit keluarannya

δ_y = (t - φ_y) (φ_y) (1 - φ_y)

= (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* (α) = 0,2

$\Delta w_{kj} = \alpha \delta_k \varphi_k$

Δw_{01} = 0,2 (-0,147911) (1) = 0,029582

Δw_{11} = 0,2 (0,147911) (0,684714) = 0,020255

Δw_{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.

$\delta_{net_k} = \delta_y w_k$

δ_{net_1} = (0,147911) (-0,3) = -0,04437

δ_{net_2} = (0,147911) (-0,13) = -0,00958

δ_k = (δ_{net_k}) (φ_k) (1 - φ_k)

δ_1 = -0,04437 (0,684714) (1-0,684714) = -0,00958

δ_2 = -0,027 (0,451305) (1-0,451305) = -0,00476

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

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$\Delta v_{jk} = \alpha \delta_k v_{jk}$

Δv = 0,2 (-0,00958) (1) = -0,00191587

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

Tabel 4.14 Perubahan Bobot (ΔV_{kj})

	Z1 (*hidden*)	Z2 (*hidden*)
x_1	-0,00091	-0,00045
x_2	-0,00091	-0,00045
x_3	-0,00091	-0,00045
x_4	-0,00091	-0,00045
x_5	0	0
x_6	0	0
x_7	0	0
x_8	0	0
x_9	0	0
1	-0,00191587	-0,00095

d) Penyesuaian bobot

Langkah terakhir adalah mendapatkan bobot baru unit keluaran:

w_{kj}(baru) = w_{kj}(lama) + Δw_{kj}

w_{10}(baru) = w_{10}(lama) + Δw_{10} = -0,36 + 0,029582= -0,334178

w_{11}(baru) = w_{11}(lama) + Δw_{11} = -0,3 + 0,020255= -0,27974

w_{12}(baru) = w_{12}(lama) + Δw_{12} = -0,13 + 0,013351= -0,3436649408

v_{jk}(baru) = v_{jk}(lama) + Δv_{jk}

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

v_{01} (baru) = v_{01}(lama) + Δv_{01}= 0,89 - 0,00191587 = 0,888084

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

Tabel 4.15 Inisialisasi Bobot *Input* Iterasi Kedua

		Z1 (*hidden*)	Z2 (*hidden*)
Input	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

Tabel 4.16 Inisialisasi Bobot *Output* Iterasi Kedua

		Y (*output*)
hidden	Z1	-0,3304178
	Z2	-0,27974
	1	-0,346649408

Proses ini akan terus berlanjut hingga sebanyak iterasi yang diinginkan dan nilai y mendekati t. Untuk data penuh yang dimiliki, diaplikasikan menggunakan program *Rstudio* untuk mempermudah perhitungannya. Data yang dimiliki

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

sebanyak 1141 dokumen dengan penilaian 1 untuk t berlabel positif dan 0 untuk t 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:

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

b) 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.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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*.

$$Akurasi = \frac{Jumlah\ data\ benar}{Jumlah\ semua\ data} . 100\%$$

$$Akurasi = \frac{176}{228} . 100\% = 77\%$$

$$Presisi\ Positif = \frac{TP}{TP+FN} . 100\% = \frac{126}{126+37} . 100\% = 77{,}3\%$$

$$Presisi\ Negatif = \frac{TN}{TN+FP} . 100\% = \frac{50}{50+15} . 100\% = 76{,}9\%$$

$$Rata - rata\ Presisi = \frac{(77{,}3+76{,}9)\%}{2} = 77{,}1\%$$

$$Recall\ Positif = \frac{TP}{TP+FP} . 100\% = \frac{126}{126+15} . 100\% = 89{,}4\%$$

$$Recall\ Negatif = \frac{TN}{TN+FN} . 100\% = \frac{50}{50+37} . 100\% = 57{,}47\%$$

$$Rata - rata\ Recall = \frac{(89{,}4+57{,}47)\%}{2} = 73{,}4\%$$

$$F - Measure = \frac{(73{,}4+77{,}1)\%}{2} = 75{,}25\%$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Berdasarkan model *backpropagation* yang didapat, klasifikasi opini islam radikalis 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 radikalis.

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 *wordcould* dengan frekuensi minimal 30 kali. Karena pada penelitian ini menggunakan dua kategori sentimen positif dan negatif, maka pada Gambar 4.2 menampilkan *wordcould* untuk frekuensi data positif dan Gambar 4.3 menampilkan frekuensi data negatif.

Gambar 4.2 *Wordcould* Data dengan Kategori Positif

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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 *wordcould* 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 *wordcould* dengan kemunculan kata yang berpengaruh terhadap kategori *tweet* negatif.

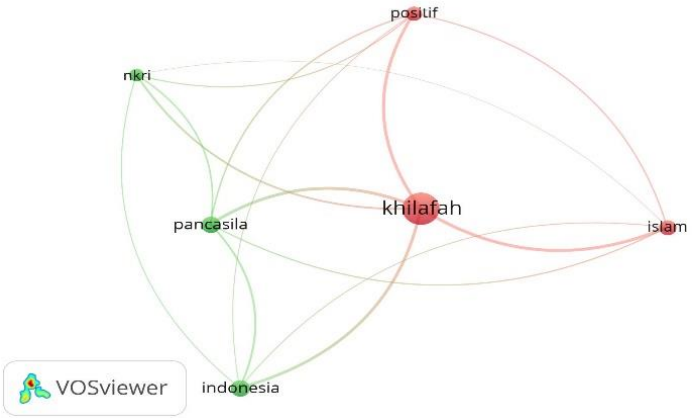
Gambar 4.3*Wordcould* Data dengan Kategori Negatif

Gambar 4.3 menunjukkan bahwa beberapa kata yang berkontribusi untuk menjadikan suatu dokumen bersentimen negatif diantaranya kata “islam”,

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

“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 *wordcould* kedua kategori sentimen diantaranya ada kata negara, sistem, dan indonesia. Analisis visualiasi 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.



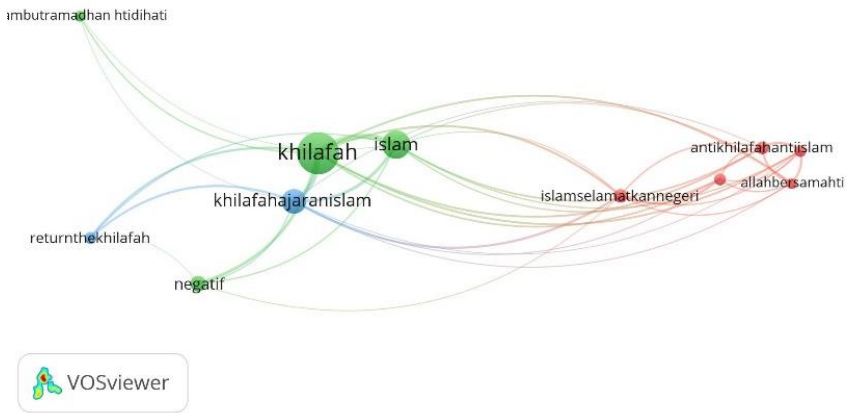
Gambar 4.4 Hasil Visualisasi Kategori Positif

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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 berkategorikan negatif ditampilkan pada Gambar 4.5.



Gambar 4.5 Hasil Visualisasi Kategori Negatif

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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 "islamselamatkannegeri", "allahbersamahti", 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 yng 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.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB V

PENUTUP

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 *input*an yang terstruktur berdasarkan sejumlah data yang dimiliki untuk diklasifikasi sentimennya.
2. Membangun model klasifikasi sentimen *tweet* islam radikalis menggunakan metode *backpropagation* dapat dilakukan berdasarkan uji coba jumlah *node hidden*, jumlah *hidden layer*, besar *learning rate*, dan pembagian data *traning-testing*. Uji coba sistem menghasilkan model arsitektur dengan 579 *node input*an, 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 wordcould* dan *graph network*, kecendurangan kategori sentimen opini tergantung pada semboyan dan kata

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

yang digunakan pada *tweet*. *Tweet* yang mengandung sentimen islam radikalis menggunakan frase “returnkhilafah”, “islamselamatkannegeri”, “antikhilafahantiislam”, “allahbersamahti”, 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.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, berikut saran untuk perbaikan penelitian selanjutnya:

1. Penelitian ini menghasilkan model yang cukup baik dan dapat digunakan sebagai acuan untuk data *tweet* yang lebih besar. Namun, masih perlu banyak pengembangan terutamanya untuk meningkatkan akurasi.
2. kata kunci “khilafah” masih terbilang cukup umum dan general, sehingga terdapat beberapa kata yang teridentifikasi positif juga ada pada kategori negatif, begitu pula sebaliknya. Hal inilah yang mempengaruhi besarnya akurasi yang didapat pada saat proses klasifikasi dilakukan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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

Hakimi, F. D. D., 2018. *Sistem Analisis Sentimen Publik Tentang Opini Pemilihan Kepala Daerah Jawa Timur 2018 pada Dokumen Twitter Menggunakan Naive Bayes Classifier,* Surabaya: Universitas Islam Sunan Ampel.

Hanifah, I. & Prastowo, B. N., 2016. Uji GPS Tracking dalam Skala Transportasi Antar Kota. *IJEIS,* Volume 6, pp. 175-186.

Istiadi, Y., 2018. *Biologi.* Revisi penyunt. Jakarta: ERLANGGA.

Iswahyudi, C. & Hasan, L. O. M., 2016. Interkoneksi Jejaring Sosial Twitter dan Sistem Informasi Berbasis Web dengan Menerapkan WEB Service, Crontab, dan API (Study Kasus Jadwal Bimbingan Dosen di IST AKPRIND Yogyakarta). *Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST),* pp. 361-368.

Jayanti, Q. & Rustiana, 2017. *Analisis Tingkat AKurasi Model-model Prediksi Kebangkrutan untuk Memprediksi Voluntary Auditor Switching,* Yogyakarta: Universitas Atma Jaya Yogyakarta.

Jian, Z., Xu, C. & Han-shi, W., 2010. Sentiment Classification Using The Theory of ANNs. *ScienceDirect,* pp. 58-62.

Khamid, N., 2016. Bahaya Radikalisme Terhadap NIKRI. *Journal of Islamic Studies and Humanities,* Volume 1.

Krenker, A., Bester, J. & Kos, A., 2011. *Introduction to The Artificial Neural Networks.* Slovenia: In Tech.

Kurniawan, B., Effendi, S. & Sitompul, O. S., 2012. KLasifikasi Konten Berita dengan Metode Text Mining. *Jurnal Dunia Teknologi Informasi,* pp. 14-19.

Kwak, H., Lee, C., Park, H. & Moon, S., 2010. What is Twitter, a Social Network or a News Media ?. pp. 591-600.

Leavline, D. E. J., 2015. Artificial Neural Network Design Flow for Classification Problem Using MATLAB. *IJARBEST,* Volume 1, pp. 22-25.

Li, Q. et al., 2016. Tweet Sentiment by Incorporating Sentiment-Specifics Word Embedding and Weighted Text Feature. *IEEE/WIC/ACM International Conference on Web Intelligence.*

Liu, B., 2012. *Sentiment Analysis and Opinion Mining.* 1st penyunt. s.l.:Morgan & Claypool Publishers.

Li, Z., 2017. A Data Classification Algorithm of Internet of Things Based on Neural Network. *iJOE,* Volume 13, pp. 28-37.

Maeir, H. & Dany, G. C., 2000. Neural Network for Prediction and Forecasting of Water Resources Variables: A Review of Modelling Issues and Applications. *Environmental Modeling and Software,* pp. 101-124.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Matsubara, E. T., Monard, M. C. & Prati, R. C., 2008. Exploring Unclassified Text Using Multiview Semisupervised Learning. *Emerging Technologies of Text Mining,* pp. 139-161.

Mirani, T. B. & Sasi, S., 2016. Sentiment analysis of Isis Related Tweets Using Absolute Location. *Interbational Conference Computional Science and Computational Intelligence.*

Moraes, R., Valiati, J. F. & Neto, W. P. G., 2013. Document-level Sentiment Classification: An Empirical Comparison Between SVM and ANN. *Expert System with Applications,* Volume 40, pp. 621-633.

Mujilahwati, S., 2016. Pre-processing Text Mining pada Data Twiitter. *Seminar Nasional dan Komunikasi.*

Naik, B., Nayak, J., Behera, H. & Abraham, A., 2015. A Self Adaptive Harmony Search Based Functional Link Higher ANN for Non Linear Data Classification. *NEUROCOMPUTING.*

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.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

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&quot*, Volume 4, pp. 53-58.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id