

## APLIKASI LINIER REGRESI DENGAN ALGORITMA JARINGAN SYARAF TIRUAN UNTUK SENTIMEN ANALISIS

<sup>1</sup>Amril Mutoi Siregar, <sup>2</sup>Tia Astiyah Hasan  
<sup>1</sup>amrilmutoi@ubpkarawang.ac.id, <sup>2</sup>tia.astiyacub@ubpkarawang.ac.id  
Universitas Buana Perjuangan karawang  
Jl. HS.Ronggo Waluyo, Timur, Karawang, Jawa Barat 41361

### ABSTRAK

*Penggunaan media sosial seperti twitter adalah salah satu yang populer di masyarakat Indonesia dengan kemudahan dapat diakses oleh siapapun selama ia mempunyai alat pendukung dan koneksi internet, sekarang ini orang bisa ngetweet dari mana dan kapan saja melalui gadgetnya, Karena begitu mudahnya diakses semua orang dari berbagai kalangan dan latar belakang bisa berkomentar atau membuat status untuk mengomentari suatu produk, kejadian, ataupun tokoh. Dengan banyaknya isi kicauan atau tweet yang berhubungan dengan seorang tokoh, khususnya Pemilihan calon pemimpin Presiden, Pilkada yang ada di Indonesia. Isi tweet atau kicauan setiap orang pastinya berbeda – beda tetapi muatannya hanya dibatasi oleh tiga sentiment Positif, Negatif dan Netral. Dengan adanya keputusan sentiment analisis dapat mencari isi kampanye yang sehat, mana yang negatif dan kampanye hitam. karakter untuk seleksi fitur menggunakan algoritma Neural Network (Jaringan syaraf tiruan) untuk mengklasifikasi sentimen secara otomatis, performance hasil yang didapatkan 53.33%.*

**Kata kunci:** Twitter, tweet, sentimen, sentiment analysis, Neural Network, Jaringan saraf tiruan

### 1. PENDAHULUAN

Di era informasi global saat ini, perkembangan Teknologi Informasi berkembang sangat pesat dari waktu ke waktu. Salah satunya adalah perkembangan media sosial yang memberikan informasi yang beragam. Media sosial merupakan layanan berbasis web yang memungkinkan individu membangun profil publik dan saling terhubung antara setiap individu dengan individu lainnya. Setiap individu dapat saling bertukar informasi maupun mempublikasikan hal-hal yang bersifat pribadi.

Terhitung 21 Maret 2016, Twitter genap memasuki usianya yang ke-10. Media sosial ini secara global memiliki sekitar 332 juta pengguna bulanan, dengan 500 juta kicauan dikirim setiap hari dan 200 miliar kicauan dalam se tahun. Pada pertengahan 2015, jumlah pengguna Twitter di Indonesia sendiri sudah mencapai 50 juta. Bagaimana jumlah pengguna Twitter di Indonesia saat ini. Ketika ditemui di acara ulang tahun Twitter yang ke-10 di Jakarta, Roy Simangunsong, Country Head Twitter Indonesia enggan untuk membeberkan hal tersebut.

Berdasarkan data yang diperoleh dari cnnindonesia.com, paling tidak ada 4,1 juta tweet yang berasal dari Indonesia. 77 persen pengguna Twitter di Indonesia aktif setiap harinya. Dari 77 persen tersebut, 54 persen di antaranya melakukan 2 tweet setiap harinya.

Pemanfaatan media sosial sangat beragam, tidak hanya dimanfaatkan untuk memposting hal-hal yang bersifat pribadi, tetapi ada juga yang memanfaatkan media sosial sebagai media iklan, berjualan, membangun komunitas dan lain sebagainya. Hal tersebut mendorong berbagai kalangan untuk memanfaatkan data dan informasi yang sangat beragam tersebut untuk mengetahui informasi yang sedang menjadi tren di masyarakat khususnya yang diperbincangkan di mediasosial.

Selanjutnya dari data tersebut dapat diambil informasi tentang opini atau sentimen dari pengguna media sosial terhadap entitas tertentu. Akan tetapi untuk mendapatkan informasi tambahan seperti sentimen analisis diperlukan pra-proses terlebih dahulu, karena data yang ada dalam penulisannya tidak menggunakan kata yang baku sehingga diperlukan normalisasi teks. Setelah teks dinormalisasi analisis sentimen dapat dengan mudah dilakukan. Analisis sentimen dapat dilakukan dengan metode klasifikasi, Neural Network.

### 2. LANDASAN TEORI

#### 2.1 Jaringan Syaraf Tiruan

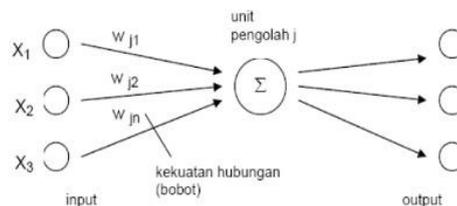
Neural Network adalah metode komputasi yang mencontoh system jaringan syaraf manusia. Metode ini menggunakan komponen perhitungan non-linear dasar yang disebut dengan neuron yang di organisasikan sebuah jaringan yang saling berhubungan, sehingga menyerupai dengan jaringan syaraf manusia.

Jaringan syaraf tiruan dibuat untuk memecahkan suatu permasalahan tertentu seperti

pengenalan pola atau klasifikasi untuk proses pembelajaran layaknya neuron manusia. JST merupakan sifat fault tolerant dalam 2 hal pertama mengenali sinyal input yang berbeda dari yang diterima sebelumnya, sebagai contoh manusia sering mengenali seseorang dari wajah yang pernah dia lihat sebelumnya dan dapat mengenali seorang yang wajahnya agak berbeda karena sudah lama tidak bertemu. Kedua mampu bekerja meskipun beberapa neuron nya tidak mampu bekerja dengan baik, jika sebuah neuron rusak neuron lain dapat dilatih untuk mengganti fungsi neuron yang rusak (sahat, 2013). JST merupakan system pemrosesan informasi yang mempunyai karakteristik menyerupai jaringan syaraf manusia (wuyandari, 2012). JST terbuat sebagai suatu generalisasi model matematis dari pemahaman manusia yang didasarkan atas pendapat sebagai berikut:

- Pemrosesan informasi yang terjadi pada komponen sederhana yang disebut dengan neuron.
- Sinyal mengalir diantara sel syaraf/neuron melalui suatu pemghubung.
- Setiap sambungan penghubung memiliki bobot yang sesuai, bobot ini digunakan untuk mendapatkan sinyal yang dikirim melaluinya.
- Setiap neuron akan menerapkan fungsi aktivasi terhadap sinyal hasil dari penjumlahan bobot yang masuk kepadanya untuk menentukan sinyal keluaran

Model struktur neuron jaringan syaraf tiruan seperti gambar berikut:



Gambar 2.1 Jaringan syaraf tiruan (JST)

Jaringan syaraf tiruan dapat belajar dari pengalaman, melakukan pembelajaran atas contoh contoh yang pernah dialaminya dan mengabstraksi karakteristik hasil masukan bahkan untuk data yang tidak terstruktur (Wuryandari, 2012).

## 22 Klasifikasi Dan Prediksi

Klasifikasi dan prediksi adalah salah satu metode yang dapat menghasilkan keputusan cerdas. Saat ini, banyak klasifikasi dan prediksi metode telah diusulkan oleh peneliti dalam pembelajaran mesin, pengenalan pola, dan statistika. Dalam studi ini, kami berfokus pada klasifikasi metode dalam data mining sebagai bagian dari proses pembelajaran mesin. Klasifikasi dan prediksi dalam data mining adalah dua bentuk analisis data yang dapat digunakan untuk mengekstrak model untuk menggambarkan data penting kelas atau untuk memprediksi tren masa depan data yang akan di prediksi. Proses klasifikasi memiliki dua tahap, tahap pertama adalah proses belajar di mana data pelatihan dianalisis dengan algoritma klasifikasi.

Model belajar atau *classifier* adalah direpresentasikan dalam bentuk aturan klasifikasi yang kedua fase adalah proses klasifikasi, di mana data pengujian yang digunakan untuk memperkirakan akurasi model klasifikasi atau classifier. Jika akurasi dianggap diterima, model dapat diterapkan pada data baru untuk mengetahui hasil prediksi. Banyak teknik yang dapat digunakan untuk klasifikasi tersebut sebagai pohon keputusan, metode Bayesian, Bayesian jaringan, berbasis peraturan algoritma, jaringan saraf, dukungan mesin vektor, asosiasi aturan pertambangan, k-tetangga terdekat-, berbasis kasus penalaran, algoritma genetika, set kasar dan logika fuzzy. Di penelitian ini kita berfokus pada tiga klasifikasi yaitu teknik pohon keputusan, jaringan saraf dan Naive Bayes. Pohon keputusan dan jaringan saraf berguna dalam mengembangkan model prediksi di banyak bidang.

## 23 Algoritma Jaringan Saraf Tiruan

*Neural Network* dapat digunakan dalam berbagai tujuan, terutama deskriptif dan *predictive* data mining. *Neural Network* awalnya dibangun dalam *machinelearning* untuk mencoba meniru neurofisiologi dari otak manusia melalui kombinasi elemen komputasi sederhana (*neuron*) dalam sistem yang saling berhubungan.

Dalam beberapa dekade terakhir, *Neural Networks* berhasil memecahkan banyak masalah di berbagai bidang antara lain ilmu saraf, matematika, statistik, fisika, ilmu komputer, teknik, biologi, dan lain-lain. *Neural Networks* (NN) diterapkan dalam pemodelan antara lain analisis *time series*, pengenalan pola, pemrosesan sinyal, teori kontrol, dan dalam pemodelan lainnya. Hal ini

dikarenakan karakteristik fundamental dari NNS yaitu kemampuan untuk belajar dari training data “dengan atau tanpa guru” (Gorunescu, 2011).

Menurut Gorunescu (2011), persamaan cara kerja dari NN dan otak manusia pada prinsipnya dikondensasikan dalam dua aspek berikut:

1. Pengetahuan diperoleh jaringan melalui proses pelatihan atau pembelajaran.
2. Intensitas koneksi antar *neuron* dikenal sebagai (*synaptic*) bobot yang digunakan untuk menyimpan pengetahuan yang diperoleh.

*Multilayer perceptron* (MLP) disebut juga *multilayer feedforward neural network* merupakan algoritma yang paling luas digunakan. Menurut Wong, Bodnovich dan Selvi (1997), sekitar 95% aplikasi bisnis yang menggunakan *neural network*, memakai algoritma ini (Vecellis, 2009). Salah satu kelebihan *neural network* adalah cukup baik dalam menangani data yang mengandung *noise* (Larose, 2005).

MLP terdiri dari *input layer*, satu atau lebih *hidden layer*, dan *output layer*. Berikut penjelasan masing-masing layer:

#### a) **Input layer**

Input layer untuk menerima nilai masukan dari tiap record pada data. Jumlah simpul input sama dengan jumlah variabel prediktor.

#### b) **Hidden layer**

Hidden layer mentransformasikan nilai input di dalam network. Jumlah hidden layer bisa berapa saja. Tiap simpul pada hidden layer terhubung dengan simpul-simpul pada hidden layer sebelumnya atau dari simpul-simpul pada input layer dan ke simpul-simpul pada hidden layer berikutnya atau ke simpul-simpul pada output layer.

#### c) **Output layer**

Garis yang terhubung dengan Output layer berasal dari hidden layer atau input layer dan mengembalikan nilai keluaran yang bersesuaian dengan variabel prediksi. Keluaran dari output layer biasanya merupakan nilai floating antara 0 sampai 1 (Kusrini, 2009).

Penemuan algoritma backpropagation untuk multilayer perceptron, merupakan metode yang sistematis untuk training sehingga bisa dilakukan dan lebih efisien. Algoritma backpropagation berasal dari learning rule Widrow dan Hoff, disusun oleh Werbos (1974), dibuat oleh Parker (1985), Rumelhart Hinton, Williams (Rumelhart dan Williams, 1986) dan peneliti lainnya (Maimon, 2005). Langkah pembelajaran dalam algoritma backpropagation adalah sebagai berikut:

Inisialisasi bobot jaringan secara acak (biasanya antara -0.1 sampai 1.0)

Hitung input untuk simpul berdasarkan nilai input dan bobot jaringan saat itu untuk setiap data pada data training, menggunakan rumus:

$$Input_j = \sum_{i=1}^n O_i w_{ij} + \Theta_j \quad (2.3)$$

Keterangan:

$O_i$  = Output simpul  $i$  dari layer sebelumnya

$w_{ij}$  = bobot relasi dari simpul  $i$  pada layer sebelumnya ke simpul  $j$

$\Theta_j$  = bias (sebagai pembatas)

Berdasarkan input dari langkah dua, selanjutnya membangkitkan output untuk simpul menggunakan fungsi aktivasi sigmoid :

$$\text{Output} = \frac{1}{1+e^{-\text{Input}}} \quad (2.4)$$

Hitung nilai Error antara nilai yang diprediksi dengan nilai yang sesungguhnya menggunakan rumus:

$$\text{Error } j = \text{Output } j \cdot (1 - \text{Output } j) \cdot (\text{Target } j - \text{Output } j) \quad (2.5)$$

Keterangan:

Output *j* = Output aktual dari simpul *j*

Target *j* = Nilai target yang sudah diketahui pada data training

Setelah nilai Error dihitung, selanjutnya dibalik ke layer sebelumnya (backpropagated). Untuk menghitung nilai Error pada hidden layer, menggunakan rumus:

$$\text{Error}_j = \text{Output}_j(1 - \text{Output}_j) \sum_{k=1} \text{Error}_k W_{jk}$$

Keterangan:

Output *j* = Output aktual dari simpul *j*

Error *k* = error simpul *k*

*W<sub>jk</sub>* = Bobot relasi dari simpul *j* ke simpul *k* pada layer berikutnya

Nilai Error yang dihasilkan dari langkah sebelumnya digunakan untuk memperbarui bobot relasi

$$\text{menggunakan rumus } W_{ij} = W_{ij} + l \cdot \text{Error } j \cdot \text{Output } i \quad (2.7)$$

Keterangan:

*W<sub>ij</sub>* = bobot relasi dari unit *i* pada layer sebelumnya ke unit *j*

*l* = learning rate (konstanta, nilainya antara 0 sampai dengan 1)

Error *j* = Error pada output layer simpul *j*

Output *i* = Output dari simpul *i*

## 24 Penelitian Terkait

Penelitian ini mempunyai keterkaitan beberapa peneliti sebelumnya dalam klasifikasi Analisa sentimen, data yang gunakan bersifat tidak terstruktur menyebabkan banyak atribut yang tidak relevan. Jika semua attribute digunakan akan mengurangi performance klasifikasi (Wang, et al., 2013). Penelitian sebelumnya melakukan klasifikasi Positif, negative dan netral (Chandani, et al. 2015)

## 25 Sentimen Analisis

Sentiment analisis adalah cara untuk menganalisa sebageian data untuk mengetahui emosi manusia. Sentimen Analisa dapat dikategorikan kedalam 3 tugas, yaitu informative text detection, information extraction dan sentiment interestingness classification (emotional, polarity indentification). Sentiment classification (negatif atau positif) digunakan untuk memprediksi sentiment polarity berdasarkan data sentimen dari pengguna (Pan, et al, 2010).

Textual sentiment analysis telah banyak digunakan, penggunaannya tidak sebatas dalam area penelitian ilmiah tetapi juga untuk kebutuhan business marketing dan teknologi (Chintala, 2012). Menurut Go (2009), sentiment analysis adalah sebuah area penelitian yang menonjol dan berkembang aktif yang dipengaruhi oleh pertumbuhan teknologi media sosial yang cepat. Melalui media sosial terdapat peluang untuk mengakses opini dari sejumlah orang pada berbagai jenis bisnis, isu dunia dan isu sosial (Go, Huang, & Bhayani, 2009).

Twitter adalah layanan microblogging yang dirilis secara resmi pada 13 Juli 2006 (Mostafa, 2013). Aktifitas utama twitter adalah mem-posting sesuatu yang pendek (tweet) melalui web atau mobile. Panjang maksimal dari tweet adalah 140 karakter, kira-kira seperti panjang karakter dari judul koran. Twitter menjadi sumber yang hampir tak terbatas yang digunakan pada text classification. Menurut Go (2009), terdapat banyak karakteristik pada tweets twitter. Pesan pada twitter memiliki banyak attribute yang unik, yang membedakan dari media sosial lainnya: 1. Twitter

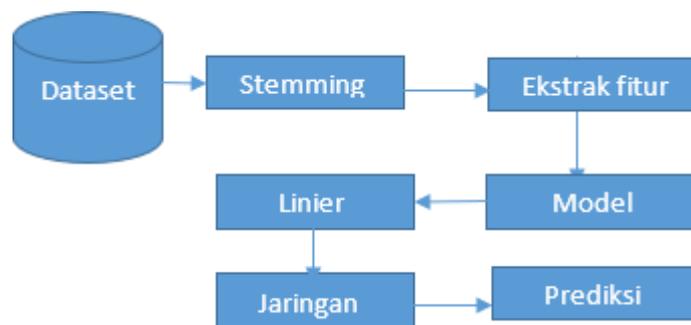
memiliki maksimal panjang karakter yaitu 140 karakter. 2. Twitter menyediakan data yang bisa diakses secara bebas dengan menggunakan Twitter API, mempermudah saat proses pengumpulan tweets dalam jumlah yang sangat banyak. 3. Model bahasa – pengguna twitter mem-posting pesan melalui banyak media berbeda. Frekuensi dari salah ejaan, bahasa gaul dan singkatan lebih tinggi daripada media sosial lainnya.

4. Pengguna twitter mengirim pesan singkat tentang berbagai topik yang disesuaikan dengan topik tertentu dan itu berlaku secara global. Selama beberapa tahun terakhir, twitter menjadi sangat populer. Jumlah pengguna twitter telah naik menjadi 190 juta dan jumlah tweet yang dipublikasikan di twitter setiap hari adalah lebih dari 65 juta.

### 3. METODE PENELITIAN

#### 3.1. METODE PENELITIAN YANG DIGUNAKAN

Metode penelitian berisi langkah-langkah yang akan digunakan dalam penelitian ini. sistematika ini menyebabkan proses penelitian lebih mudah dipahami oleh orang lain. Penelitian yang akan dilakukan adalah untuk merancang sistem yang diperoleh dari pengamatan data yang ada. Penggunaan mesin belajar untuk mempermudah mendapat keputusan atau hasil yang diharapkan. Tools yang digunakan dalam penelitian ini adalah RapidMiner sebagai alat pemodelannya. kerangka model yang diusulkan adalah hibrid dari Linear Teknik Regresi dan Artificial Neural Network, yang kategori pembelajaran diawasi (*supervised learning*).



Gambar 3.1 Sistem prediksi yang gunakan

Keterangan system prediksi yang digunakan :

#### 1. Pengumpulan dataset

Pengumpulan data dilakukan untuk mendapatkan data yang dibutuhkan dalam penelitian ini. Data yang dikumpulkan dari contoh yang sebelumnya telah ditentukan. Contoh tersebut terdiri kumpulan tweet para pengguna twitter, dan sebagai sasarannya adalah sentiment dari isi tweet tersebut.

Teks	Sentimen
We love you pak ??	Positif
kasihan Ahok, cuma krn ia maju Pilkada jadi Cagub DKI ia disasar kriminalisasi.. tidak betul cara2 spt itu	Positif
telah MENJALANKAN semua program2 rakyat so nothing left for other candidates :-)	Positif
Gak usah dibayangkan kan skrg gubernur nya siapa? Ahok kan? Buktinya rakyat jkt berbondong2... <a href="https://t.co/8EGBho5wRn">https://t.co/8EGBho5wRn</a>	Positif
katanya hasil kerja foke.,hiiii..hiiii..ngomongnya gagal pake data..semangat pak @basuki_btp .	Positif
lebih bahagia melihat @basuki_btp yang disandra&gt; @BuniYani &lt;It;jasamu kami ukir tuk bangsa dn agama	Positif
nov yg demo nuntut kosistensi ucapan @basuki_btp . ingat tanggapannya ttg video boby mah... <a href="https://t.co/J2OMf8eZvW">https://t.co/J2OMf8eZvW</a>	Positif
karena kasih ahok kepada kaum ibu tulus dan penuh kasih ....????	Positif
karena berkesan mau mengalikan kasus ahok mk di demo, tidak bulat tp lebih banyak yg menyatakan berslh	Positif
kenapa enggak semua pendukung Ahok dilaporkan ya mbah ..hayoo laporkan kalau berani..!	Positif
seru banget kak liat debat kakak dengan reply-an orang2. Hahaha sabar banget kak, keren!! Aku setuju banget Ahok adalah benar!	Positif
minta maaf krn pernyataan ahok sdh bikin gaduh...mslh hukum ga selesai dgn kata "maaf"	Negatif
krn penistaan agama atau sama @KPK_RI krn korupsi ?	Negatif
Ini masalahnya adalah @basuki_btp yg menghina Alquran dan Ulama,bukan masalah agama Islam dengan agama lain	Negatif
Dikemas dg topik apapun namanya, demo 4 Nopember itu, demo sebenarnya demo ANTI Ahok @basuki_btp dan anti @jokowi,	Negatif

Gambar 4.2 Contoh dataset learningset yang digunakan

teks	Sentimen
@AbaAbuSalim: @marwan_jafar : NAH, harusnya penegak hukum segera tangkap @basuki_btp , krn dia yg mulai. Bukan menggiring opini utk memu...	
@Alie_Bin_Alie: Abaikan mereka yg hebat memvonis seseorang @BektiAlbadri mari kita hadapi dgn cerdas @Woelan27 @suwanto111 @basuki_btp @...	
@BULLFRONT: Komentar2 di postingan Didi Arifin tentang Kalijodo yang diubah menjadi skate park dan bike park oleh Ahok. Semua mendukung...	
@HerdianaKiehl: @temanAhok #Salam2Jari Gubernur @basuki_btp #2periode Jakarta Pasti Hebat! <a href="https://t.co/JhR6ctRKJv">https://t.co/JhR6ctRKJv</a>	
rakyat utk rakyat #Salam2Jari @ruhutsitompul @basuki_btp @WagubDKI @Mongol_Stres @musashielrahman @tbayupatti @unapu...	
@didimening: @yofit_08 @AwakMedanCoy @basuki_btp minta maaf krn pernyataan ahok sdh bikin gaduh...mslh hukum ga selesai dgn kata "maaf"	
@Fr33P413st1n3: Ayo dulu2an, @basuki_btp ditahan @DivHumasPolri krn penistaan agama atau sama @KPK_RI krn korupsi ?	
@saenol_huda: Ini masalahnya adalah @basuki_btp yg menghina Alquran dan Ulama,bukan masalah agama Islam dengan agama lain @TofaLemon @jo...	
@Toni_Romdoni: Dikemas dg topik apapun namanya, demo 4 Nopember itu, demo sebenarnya demo ANTI Ahok @basuki_btp dan anti @jokowi ,	
@yanurdisah: Apa kabar mas @basuki_btp ? Bahagiakah anda kini melihat Indonesia bergejolak gara2 ucapan anda ☺ Merawat NKRI itu susah lh...	
@MANISPOLOS: Umat Islam menggelar unjuk rasa karena proses hukum terhadap @basuki_btp jalan di tempat.	
#BeritaHarian: Ahok Ingin Membuat Klinik Khusus Perempuan di Pasar <a href="https://t.co/u3NpH4nggh">https://t.co/u3NpH4nggh</a> · By #SbyCuk	
@Juang_Modjo siapa yang tau kesibukan ahok jadi gubernur. ah ente.	
@o_pripot @AgungWidrajat ya jelas lah kalo ke taman makam pahlawan ditemani pejabat pejabat ya Ahok pakai sepatu, yg Laen juga sepatuan.	
@SetiardoHarrie @wakilgubernurKW @ridwankamil @PAS_Jabar Ahok sdh beberapa kali minta maaf apakah sdh dimaafkan?	

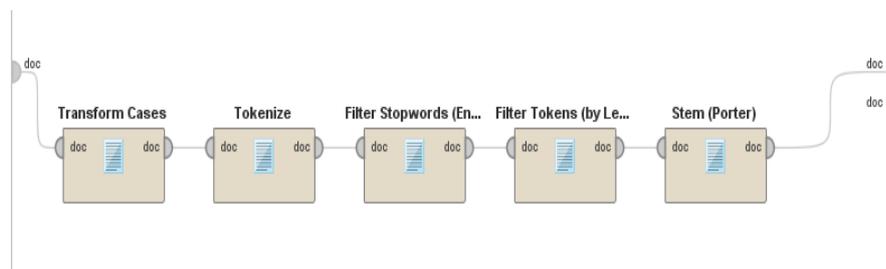
Gambar 4.3. Testing set yang digunakan

## 2. Studi Pustaka

Dalam penelitian ini menggunakan beberapa literatur-literatur, studi pustaka dilakukan untuk melengkapi pengetahuan dasar dan teori teori yang digunakan dalam penelitian ini.

## 3. Preprocessing/Stemming

Tahap preprocessing adalah merupakan tahap seleksi data yang bertujuan untuk mendapatkan data yang siap diolah untuk tahap berikutnya. Tahapan yang dikerjakan adalah untuk melakukan perubahan tipe data dan attribute atau dikenal dalam text mining adalah proses data dari yang tidak terstruktur menjadi terstruktur.



Gambar 4.4 Tampilan tahapan preprocessing text mining

## 4. Ekstrak fitur

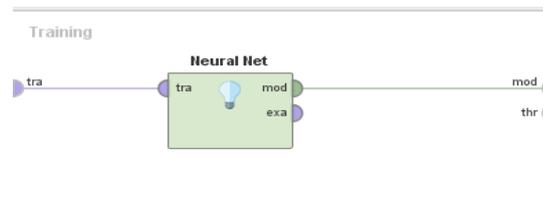
Tahapan ekstrak fitur meliputi proses tahap memotong/stem kalimat menjadi kata, dan membuat kata kata yang tidak diperlukan dan membuat symbol misalnya: @, #, & dan lain lain.

## 5. Penentuan Model

Tahap ini akan melakukan penentuan model jaringan syaraf tiruan dengan backpropagation. Hasil dari tahap ini adalah untuk mendapatkan ola yang terbaik dalam jaringan syaraf tiruan.

## 6. Pengujian Jaringan Syaraf tiruan.

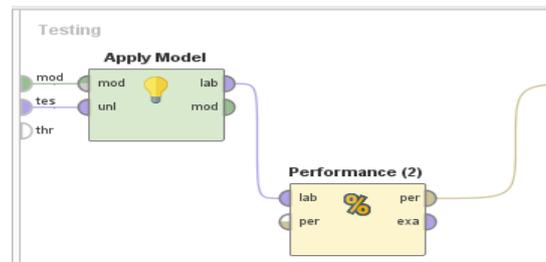
Setelah mendapatkan data yang telah diproses sebelumnya maka proses pengujian dan pelatihan data diolah dengan menggunakan algoritma Backpropagation.



Gambar 4.5. algoritma yang digunakan

7. Pegujian hasil/ prediksi output

Tahap menampilkan hasil dari pembelajaran dalam penelitian ini, akan menampilkan performance algoritma yang digunakan berupa accuracy, precision, recall setiap jenis sentiment yang digunakan.



Gambar 4.5 Testing untuk menampilkan performance sentiment analysis

4. LOKASI DAN WAKTU PENELITIAN

Penelitian ini dilakukan di Universitas Buana Perjuangan Karawang, dataset diambil dari twitter dalam masalah pilkada DKI Jakarta.waktu melakukan penelitian ini pada bulan Mei – Oktober 2018.

5. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan tools Rapidminer studio 7.3 untuk menganalisa data sentiment yang diambil dari twitter, hasil yang ditampilkan dalam penelitian berupa prediction sentiment, Confusion matrix untuk accuracy dan grafik.

Data yang digunakan dari tweet dengan learning dataset 75 tweet dengan testing 15 tweet. Learning set. Setiap data secara manual diberikan label sentimen (positif, netral, negatif). Data training yang telah memiliki label diuji dengan Rapidminer dengan menggunakan jaringan syaraf tiruan. Hasil penelitian ini seperti dibawah ini:

No	Predicti on Sentim en	Confidence Positif	Confidence Negatif	Confidence Netral	Tweet
1	Negatif	0.20657385	0.68516503	0.10826112	@AbaAbuSalim: @marwan_jafar : NAH, harusnya penegak hukum segera tangkap @basuki_btp ,
2	Negatif	0.23907647	0.6010861	0.15983743	@Alie_Bin_Alie: Abaikan mereka yg hebat memvonis seseorang @BektiAlbadri mari kita
3	Negatif	0.39662571	0.57054416	0.03283013	@BULLFRONT: Komentar2 di postingan Didi Arifin tentang Kalijodo yang diubah menjadi skate park
4	Negatif	0.27201167	0.65316371	0.07482463	@HerdianaKiehl: @temanAhok #Salam2Jari Gubernur @basuki_btp #2periode Jakarta Pasti
5	Netral	0.2445509	0.30846692	0.44698218	@itschacha_ID: dari rakyat utk rakyat #Salam2Jari @ruhutsitompul @basuki_btp @WagubDKI
6	Negatif	0.34589507	0.51111085	0.14299408	@didimening: @yofit_08 @AwakMedanCoy @basuki_btp minta maaf krn pernyataan ahok sdh

7	Negatif	0.27508499	0.66109092	0.06382408	@Fr33P4l3st1n3: Ayo dulu2an, @basuki_btp ditahan @DivHumasPolri krn penistaan agama
8	Negatif	0.34010717	0.57595987	0.08393296	@saenol_huda: Ini masalahnya adalah @basuki_btp yg menghina Alquran dan
9	Positif	0.48260648	0.37830123	0.13909229	@Toni_Romdoni: Dikemas dg topik apapun namanya, demo 4 Nopember itu, demo
10	Negatif	0.25052462	0.55803019	0.1914452	@yanurdisah: Apa kabar mas @basuki_btp ? Bahagiakah anda kini melihat Indonesia
11	Negatif	0.25539599	0.66587302	0.07873099	@MANISPOLOS: Umat Islam menggelar unjuk rasa karena proses hukum terhadap @basuki_btp jalan
12	Negatif	0.35521008	0.44763299	0.19715692	#BeritaHarian: Ahok Ingin Membuat Klinik Khusus Perempuan di Pasar https://t.co/u3NpH4nggh · By
13	Negatif	0.45267113	0.46524012	0.08208876	@Juang_Modjo siapa yang tau kesibukan ahok jadi gubernur. ah ente.
14	Positif	0.43372156	0.35149771	0.21478073	@o_pripot @AgungWidrajat ya jelas lah kalo ke taman makam pahlawan ditemani pejabat
15	Negatif	0.39650389	0.53115167	0.07234444	@SetiardjoHarrie @wakilgubernurKW @ridwankamil @PAS_Jabar Ahok sdh beberapa

Gambar 4.1 Hasil prediction sentiment

Hasil penelitian ini dapat kita bandingkan antara hasil dari performance algoritma jaringan syaraf tiruan dengan hasil judgement manusia. Hasilnya jika di bandingkan dari 15 tweet testing set yang digunakan mempunyai kesamaan sentiment adalah 4 tweet. Jadi lebih banyak yang tidak sesuai, algoritma neural network kurang baik performance untuk Analisa yang berbentuk teks (test mining) karena Algoritma Neural network lebih bagus ketika datasetnya bebrbentuk angka (numerik). untuk lebih jelas bisa dilihat tabel berikut yang warna hijau yang memiliki kesamaan sentiment.

Text	Judgement manusia	Performance Algoritma
@AbaAbuSalim: @marwan_jafar : NAH, harusnya penegak hukum segera tangkap @basuki_btp , krn dia yg mulai. Bukan menggiring opini utk memu...	Positif	Negatif
@Alie_Bin_Alie: Abaikan mereka yg hebat memvonis seseorang @BektiAlbadri mari kita hadapi dgn cerdas @Woelan27 @suwanto111 @basuki_btp @...	Positif	Negatif
@BULLFRONT: Komentar2 di postingan Didi Arifin tentang Kalijodo yang diubah menjadi skate park dan bike park oleh Ahok. Semua mendukung...	Positif	Negatif
@HerdianaKiehl: @temanAhok #Salam2Jari Gubernur @basuki_btp #2periode Jakarta Pasti Hebat! https://t.co/JhR6ctRKJv	Positif	Negatif
@itschacha_ID: dari rakyat utk rakyat #Salam2Jari @ruhutsitompul @basuki_btp @WagubDKI @Mongol_Stres @musashielrahman @tbayupatti @unapu...	Positif	Netral
@didimening: @yofit_08 @AwakMedanCoy @basuki_btp minta maaf krn pernyataan ahok sdh bikin gaduh...mslh hukum ga selesai dgn kata "maaf"	Negatif	Negatif
@Fr33P4l3st1n3: Ayo dulu2an, @basuki_btp ditahan @DivHumasPolri krn penistaan agama atau sama @KPK_RI krn korupsi ?	Negatif	Negatif
@saenol_huda: Ini masalahnya adalah @basuki_btp yg menghina Alquran dan Ulama,bukan masalah agama Islam dengan agama lain @TofaLemon @jo...	Negatif	Negatif
@Toni_Romdoni: Dikemas dg topik apapun namanya, demo 4 Nopember itu, demo sebenarnya demo ANTI Ahok @basuki_btp dan anti @jokowi ,	Negatif	Positif
@yanurdisah: Apa kabar mas @basuki_btp ? Bahagiakah anda kini melihat Indonesia bergejolak gara2 ucapan anda 😊 Merawat NKRI itu susah lh...	Negatif	Negatif
@MANISPOLOS: Umat Islam menggelar unjuk rasa karena proses hukum terhadap @basuki_btp jalan di tempat.	Netral	Negatif
#BeritaHarian: Ahok Ingin Membuat Klinik Khusus Perempuan di Pasar https://t.co/u3NpH4nggh · By #SbyCuk	Netral	Negatif
@Juang_Modjo siapa yang tau kesibukan ahok jadi gubernur. ah ente.	Netral	Negatif
@o_pripot @AgungWidrajat ya jelas lah kalo ke taman makam pahlawan ditemani pejabat pejabat ya Ahok pakai sepatu, yg Laen juga sepatuan.	Netral	Positif

@SetiardoHarrie @wakilgubernurKW @ridwankamil @PAS_Jabar Ahok sdh beberapa kali minta maaf apakah sdh dimaafkan?	Netral	Negatif
--	--------	---------

Gambar 4.3 Perbandingan Hasil prediction dengan judgement sentiment

Dalam penelitian juga hasil dapat ditampilkan berupa confusion matrix sebagai tolak ukur keberhasilan penelitian ini, yang terdiri dari accuracy, recall dan precision. Berikut hasilnya dibawah ini:

accuracy: 53.33%

	true Positif	true Negatif	true Netral	class precision
pred. Positif	4	1	3	50.00%
pred. Negatif	1	3	1	60.00%
pred. Netral	0	1	1	50.00%
class recall	80.00%	60.00%	20.00%	

Gambar 4.2 Tampilan hasil confusion matrix

## 5.1. KESIMPULAN DAN RENCANA KERJA KEDEPAN

Setelah melakukan penelitian ini ada beberapa hasil yang didapat. Secara umum hasil telah didapatkan adalah sebagai berikut:

1. Tingkat accuracy yang didapatkan adalah 53.33%
2. Penelitian sentiment dalam Bahasa Indonesia belum bisa maksimal karena wordnet kata kata Bahasa Indonesia belum tersedia.
3. Melihat hasil pembicaraan melalui media sosial terutama twitter, masih banyak ditemukan menggunakan sentiment negative dilihat dari jumlah tweet nya cukup besar, yang dapat dikategorikan masyarakat Indonesia belum menggunakan media sosial dengan sehat.

## 5.2. SARAN

Walaupun penelitian ini telah dilakukan, peneliti masih harus mengembangkan analisis dan hasil lebih lanjut, baik berupa algoritma/model lain untuk kasus linear regresi berdasarkan sentiment dengan menggunakan tools, matlab, Python.

## REFERENCES

- Anyama, Oscar Uzoma and Igiri, Chinwe Peace” An Application of Linear Regression & Artificial Neural Network Model in the NFL Result Prediction” IJERT Januari 2015
- Hera Wasiati, Sistem Pendukung Keputusan Penentuan Kelayakan Calon Tenaga Kerja Indonesia Menggunakan Metode Naive Bayes (Studi Kasus: Di P.T. Karyatama Mitra Sejati Yogyakarta), *IJNS* – Indonesian Journal on Networking and Security - Volume 3 No 2- 2014 –<http://ijns.org>
- Liu, Bing. 2012. *Sentiment Analysis And Opinion Mining*. Chicago: Morgan & Claypool Publisher. <http://www.dcc.ufrj.br/~valeriab/DTMSentiment-AnalysisAndOpinionMining-BingLiu.pdf>
- Maimon, Oded, and Lior Rokach, eds. *Data mining and knowledge discovery handbook*. Vol. 2. New York: Springer, 2005.
- Chandani V et al. 2015. Komparasi algoritma klasifikasi machine learning dan feature selection pada analisa sentimen film
- Wang, et al, 2013. Sample cutting method for imbalanced text sentiment classification based on BRC Knowledge Based System, 37, 451-461.
- Sahat, 2013. Metode jaringan syaraf tiruan dalam prediksi serangan jantung yang efektif, jurnal SNATIKOM tunas bangsa siantar.
- Wuryandari, 2012, Perbandingan Metode Jaringan Saraf Tiruan Backpropogation dengan Learning Vector Quantization pada pengenalan wajah, Jurnal KOMPUTA edisi 1 1 Volume 1. Bandung.
- Pan, S., Ni, X., Sun, J., Yang, Q., & Chen, Z. (2010). Cross-domain sentiment classification via spectral feature alignment. International World Wide Web Conference Committee, 751-760.