

COMPARISON ALGORITHM CLASSIFICATION NAIVE BAYES, DECISION TREE, AND NEURAL NETWORK FOR ANALYSIS SENTIMENT

Ahmad Turmudi Zy

Program Studi Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa
turmudi@pelitabangsa.ac.id

Abstrak

Twitter adalah media sosial, boleh diakses bebas oleh siapapun selama ia mempunyai alat pendukung dan koneksi internet, masa-masa sekarang orang bisa ngetweet dari mana dan kapan saja melalui *gadget*-nya, baik itu Handphone, Iphone, Tablet, Ipad ataupun Tab. Karena begitu mudahnya diakses semua orang dari berbagai kalangan dan latar belakang bisa berkomentar atau membuat status untuk mengomentari suatu produk, kejadian, ataupun tokoh. Pada penulisan tesis ini akan dikhususkan membahas tentang isi kicauan atau *tweet* yang berhubungan dengan seorang tokoh, khususnya calon pemimpin Propinsi DKI Jakarta Tahun 2017. Isi *tweet* atau kicauan setiap orang pastinya berbeda-beda tetapi muatannya hanya dibatasi oleh tiga kemungkinan, positif, netral ataupun negatif. Ketiga Muatan tersebut adalah sentimen-sentimen dari tweeter. Berdasarkan data dari twitter muatan-muatan *tweet* akan diklasifikasikan sesuai sifatnya. Dalam penelitian ini akan diterapkan proses *text mining*. Karakter untuk seleksi fitur menggunakan algoritma *Naive Bayes Classifier*, *Decision Tree*, *Neural Network* untuk mengklasifikasi sentimen secara otomatis. Penelitian ini menggunakan 1200 data *tweet* tentang sentimen kepada Calon Gubernur Jakarta tahun 2017. Data tersebut diklasifikasikan secara manual dan dibagi ke dalam masing-masing 400 data untuk sentimen positif, negatif dan netral. Kemudian 1200 data digunakan untuk testing. Diharapkan dari hasil perbandingan ini didapatkan algoritma terbaik dan akurat dalam membuat klasifikasi sentimen analisis.

Kata kunci: Twitter, tweet, sentimen, *sentiment analysis*, *Naive Bayes Classifier*, *Decision Tree*, *Neural Network*.

Abstract

Twitter is a social media, that can be accessed by anyone as long as they have a supporting gadget and an internet connection, nowadays everyone can tweet anywhere and anytime using their gadget, whether it is a cellphone, iPhone, tablet, iPad, or tab. Because of the ease of access, everyone from any background can comment and create a status to comment on a product, event or character. This thesis will specialize on tweets with regards to characters, especially on the Governor election of Jakarta on 2017. Every tweets must be different but the content there is only three content in a tweet, positive, negative and neutral. These three contents reflect on the sentiments of the tweeter. Based on the contents of the tweet, it will be divided into its characters. These characters will use feature selection using algorithmic theory of *Naive Bayes Classifier*, *Decision Tree*, *Neural Network* to automatically classify the sentiment. This research will use 1200 tweet about the sentiment of the Jakarta Election. These data will be classified manually and divided into 400 data each for positive, negative and neutral. And then 1200 data will be used for testing. Hopefully the comparison will get the best algorithm that is accurate in classifying the sentiments.

Keywords: Twitter, tweet, sentiment, *sentiment analysis*, *Naive Bayes Classifier*, *Decision Tree*, *Neural Network*.

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 jejaring sosial seperti Twitter, Facebook, dan Youtube merupakan beberapa media perangkat komunikasi terpopuler di masyarakat saat ini [1]. Menjelang pemilihan kepala daerah, para tokoh publik dan

politisi sering memanfaatkan media sosial untuk meningkatkan popularitas mereka. Salah satu media sosial yang sering digunakan untuk meningkatkan popularitas seorang tokoh public adalah Twitter. Twitter telah dimanfaatkan dalam pemilihan umum di beberapa negara singapura, jerman dan amerika.

Pada penelitian ini memanfaatkan media social twitter tersebut untuk menganalisa tweet berbahasa Indonesia yang membicarakan tentang tokoh public yang menjadi kandidat calon Gubernur DKI Jakarta pada pemilihan kepala daerah tahun 2017. Analisis yang dilakukan dengan membandingkan tingkat akurasi proses klasifikasi dengan menggunakan tiga algoritma yaitu *Naive Bayes* [2], *Decision Tree* [3], dan *Neural Network* terhadap sentiment analisis.

Klasifikasi tweet yang dilakukan pada penelitian ini diperoleh berdasarkan kombinasi antara kelas sentiment. Kelas sentiment dibagi menjadi tiga popularitas yaitu sentiment positif, sentiment negative dan netral.

2. Penelitian Terkait

Penelitian tentang klasifikasi sentiment telah banyak dilakukan sebelumnya, diantaranya penelitian yang dilakukan oleh Ismail Sunni1, Dwi Hendratmo Widyantoro pada tahun 2012 terhadap sentimen tokoh public. Pada penelitiannya, Ismail dkk lebih menekankan pada praproses sebelum dimasukkan pada proses klasifikasi dengan algoritma Nive Bayes. Data tweet yang berhasil dikumpulkan dilakukan proposes dengan menghilangkan kata-kata yang tidak baku, convert emoticon dan kata yang tidak baku. Kemudian data tersebut diberi label secara manual untuk selanjutnya dilakukan proses klasifikasi. Hasil yang di dapatkan untuk akurasinya mencapai (69.4%-72.8%) [4].

Selanjutnya penelitian yang dilakukan oleh Paulina Aliandu (2013), pada papernya, Paulina membangun sebuah sistem yang bisa menentukan sentimen pada sebuah *tweet* yang mengacu pada pendapat seseorang di twitternya. Data yang diambil adalah *tweet* dari @SBYudhoyono, dan dari orang-orang yang mengomentarnya. Data yang diambil dari *twitter* dimulai dari tanggal 13 April 2013 sampai 5 Mei 2013, dan didapatkan *tweet* sebanyak 230, yang kemudian di labelin secara manual. Sistem yang dibangaun oleh paulina, yang mengguankan Naive bayes mendapatkan akurasi 79,42%.

Ahmad Fathan Hidayatullah, pada penelitiannya berhasil melakukan analisis sentimen dan klasifikasi tweet berbahasa Indonesia tentang tokoh publik. Hasil akurasi pengujian klasifikasi kategori dengan fitur term frequency diperoleh sebesar 79,91% sedangkan fitur TF-IDF didapatkan akurasi sebesar 79,68%. Klasifikasi menggunakan *tools* RapidMiner dengan Naive Bayes dan fitur term frequency diperoleh sebesar 73,81% sedangkan dengan fitur TF-IDF diperoleh sebesar 71.11%. Klasifikasi kategori dengan *Support Vector Machine* menghasilkan akurasi 83,14% untuk fitur term frequency dan 82,69% untuk fitur TF-IDF. Metode *Support Vector Machine* menghasilkan akurasi performansi yang lebih baik daripada metode Naive Bayes. Penggunaan metode *Support Vector Machine* dan *Naive Bayes* sama-sama memiliki hasil akurasi yang cukup baik untuk klasifikasi *tweet* [5].

3. Metode Penelitian

3.1 Pengumpulan Data

Penelitian dilakukan selama kurang lebih 3 bulan. Dataset yang akan digunakan dalam penelitian adalah public timeline tweet bahasa Indonesia yang merupakan hasil pencarian berdasarkan percakapan seseorang terhadap account resmi tokoh politik yang akan maju sebagai calon Gubernur Dki Jakarta 2017. Beberapa account resmi yang diambil diantaranya, @Agusyodhoyono (Agus Yudhoyono), @Basuki_btp (Basuki Tjahya Purnama), @Aniesbaswedan (Anies Baswedan).

Alat yang digunakan untuk pengambilan data menggunakan perangkat lunak aplikasi Tag V6.1.

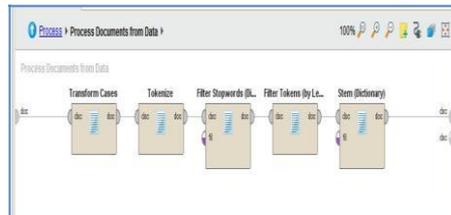


Gambar 1. Gambar tampilan program TAGS v6.1

3.2 Praproses

Tujuan dilakukannya preprocessing dokumen adalah untuk menghilangkan noise, menyeragamkan bentuk kata dan mengurangi volume kata [6]. Tahapan yang dilakukan dari dokumen preprocessing dapat dilihat pada gambar 3.2.

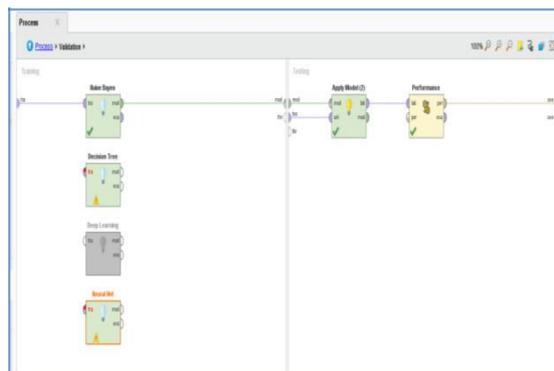
- a. **transform case**: Adalah proses merubah setiap huruf menjadi huruf kecil
- b. **tokenize** : Adalah proses memecah kalimat menjadi perkata
- c. **filter stopwords**: Adalah proses menyaring setiap kata yang tidak punya arti untuk di hilangkan.
- d. **filter tokens by length**: Adalah proses pembuangan kata yang tidak jelas maknanya berdasar panjang katanya.
- e. **stem (dictionary)** : Menyetandakan kata-kata berimbuhan kedalam kata aslinya mengacu pada kamus [7].



Gambar 2. tampilan praproses RapidMiner 7.4

3.3 Penerapan Algoritma

Algoritma yang digunakan dalam penelitian ini antara lain: Naive bayes, decision tree, neural network. Ketiganya digunakan sebagai perbandingan saja, yang tujuannya untuk mengetahui tingkat akurasi dari setiap algoritma tersebut.



Gambar 3. Proses penerapan algoritma pada RapidMiner 7.4

3.4 Evaluasi dan Pengujian

Evaluasi dan pengujian yang dilakukan pada penelitian ini dengan membagi data menjadi 2 bagian. 10% data testing dan 90% data training. Untuk proses evaluasi menggunakan confusion matrix. Confusion matrix merupakan salah satu tools penting dalam metode visualisasi yang digunakan pada mesin pembelajaran yang biasanya memuat dua kategori atau lebih. Tabel 3.1 menggambarkan contoh hasil confusion matrix prediksi dua kelas

Tabel 1. Contoh Hasil Confusion Matrix Prediksi Tiga Kelas

		Actual Class		
		Class-1	Class-2	Class-3
Predicted Class	Class-1	True positive	False negative	False Neutral
	Class-2	False positive	True negative	False Neutral
	Class-3	False Neutral	False Neutral	True Neutral

Nilai true positive (TP), true negative (TN) dan true neutral (TN2) adalah hasil klasifikasi yang benar. Nilai false positive (FP) adalah nilai dimana hasilnya diprediksi sebagai class-1 namun sebenarnya merupakan class-2 sedangkan false negative (FN) adalah nilai dimana prediksi mengklasifikasikan sebagai class-2 namun faktanya termasuk dalam klasifikasi class-1. Nilai false neutral (FN2) adalah hasil diprediksi sebagai class-3 namun faktanya termasuk dalam klasifikasi class-1 dan class-2. Nilai akurasi confusion matrix berdasarkan Tabel 2 diperoleh dengan persamaan (3.1).

$$(3.1)$$

4. Hasil dan Pembahasan

4.1 Dataset

Penelitian ini mengambil data dari Twitter pada tanggal 1 November 2016, dari 3 calon Gubernur DKI Jakarta, sebanyak 3000 Data, pengambilan tweet dari tweeter.com menggunakan Tag V6.1.

Tabel 2. Data Set positif, negatif, netral dari calon gubernur Jakarta

Tweet	Data Set	Positif	Negatif	Netral	Data Akhir
@AgusYudhoyono	3000	317	48	35	400
@Basuki_btp	3000	181	187	32	400
@AniesBaswedan	3000	239	144	17	400
Total	9000	737	379	84	1200

Dari tabel 4.1 terlihat bahwa data yang semula 3000 untuk setiap calon, setelah di anotasi menjadi 400 data set percalon.

4.2 Proses Eksperimen

Proses eksperimen dilakukan terhadap masing-masing tokoh public calon Gubernur DKI pada pemilihan tahun 2017 yaitu, Agus Yudhoyono (Agus), Basuki Tjahaya Purnama (Ahok) dan Anies Baswedan (Anis). Proses klasifikasi dilakukan uji coba terhadap masing-masing algoritma untuk mendapatkan hasil kalasifikasi dengan pengukuran *recall*. **Recall** atau juga disebut sensitivitas merupakan pengambilan data yang relevan dari kumpulan data (wikipedia).

1. Eksperimen terhadap calon nomor urut 1, yaitu Agus Yudhoyono (Agus)
Total data tweet yang berhasil dikumpulkan adalah 400 tweet, dengan pembagian 40 data testing dan 360 data training, yang akan diuji menggunakan *tool* rapidminer, sehingga didapatkan data hasil dari praproses, recall, prediksi dan klasifikasi yang dilakukan di rapidminer.

Tabel 3. Data tweet Agus

Data <i>training</i>	360
Data test	40
Tweets Positif	317
Tweets Negatif	48
Tweets Netral	35
Tweets Total	400

Tabel 4. Hasil eksperimen recall Agus

	Manual	NB	DT	NN
Positif	32	28	32	31
Negatif	5	3	0	2
Netral	3	0	0	1

Kemudian dari semua perhitungan di hitung rata-ratanya (*average*), hasil *averagenya* sebagaimana di table 4.4. Terlihat bahwa untuk recall di eksperimen ini algoritma Neural Network mendapatkan hasil paling tinggi.

Tabel 5. hasil Average recall eksperimen Agus

Average Recall	NB	DT	NN
	0,88	1,00	0,97
	0,60	0,00	0,40
	0,00	0,00	0,33
	49%	33%	57%

Selanjutnya dilakukan perhitungan tingkat akurasi proses klasifikasi terhadap tiga algoritma. Nilai akurasi dari percobaan tersebut diuraikan dalam table 4.5.

Tabel 6. Akurasi algoritma

Algoritma	Akurasi
Nive Bayes (NB)	77,5%
Decision Tree (DT)	80%
Neural Network (NN)	85%

Dari nilai-nilai akurasi pada tabel 4.5, maka dalam hal ini akurasi terbaik untuk eksperimen agus, adalah *Neural Network*.

2. Eksperimen terhadap calon nomor urut 2, yaitu Basuki Tjahaya Purnama (Ahok)

Total data tweet yang berhasil dikumpulkan adalah 400 tweet, dengan pembagian 40 data testing dan 360 data training, yang akan diuji menggunakan *tool* rapidminer, sehingga didapatkan data hasil dari paproses, recall, prediksi dan klasifikasi yang dilakukan di rapidminer.

Tabel 7 Data tweet Ahok

Data training	360
Data test	40
Tweets Positif	181
Tweets Negatif	187
Tweets Netral	32
Tweets Total	400

Tabel 8. Hasil eksperimen recall Ahok

	Manual	NB	DT	NN
Positif	18	10	18	14
Negatif	19	10	0	6
Netral	3	0	0	0

Kemudian dari semua perhitungan di hitung rata-ratanya (*average*), hasil *averagenya* sebagaimana di table 4.8. Terlihat bahwa untuk recall di eksperimen ini algoritma Neural Network mendapatkan hasil paling tinggi.

Tabel 9. hasil Average recall eksperimen Ahok

Average Recall	NB	DT	NN
	0,56	1,00	0,78
	0,53	0,00	0,32
	0,00	0,00	0,00
	36%	33%	36%

Selanjutnya dilakukan perhitungan tingkat akurasi proses klasifikasi terhadap tiga algoritma. Nilai akurasi dari percobaan tersebut diuraikan dalam table 4.9.

Tabel 10. Akurasi algoritma

Algoritma	Akurasi
Nive Bayes (NB)	52,6%
Decision Tree (DT)	45%
Neural Network (NN)	50%

Dari nilai-nilai akurasi pada tabel 4.9, maka dalam hal ini akurasi terbaik untuk eksperimen agus, adalah *Naïve Bayes*.

3. Eksperimen terhadap calon nomor urut 3, yaitu Anies Baswedan (Anis)

Total data tweet yang berhasil dikumpulkan adalah 400 tweet, dengan pembagian 40 data testing dan 360 data training, yang akan diuji menggunakan *tool* rapidminer, sehingga didapatkan data hasil dari praproses, recall, prediksi dan klasifikasi yang dilakukan di rapidminer.

Tabel 11. Data tweet Anis

Data training	360
Data test	40
Tweets Positif	239
Tweets Negatif	144
Tweets Netral	17
Tweets Total	400

Tabel 14. Hasil eksperimen recall Anis

	Manual	NB	DT	NN
Positif	24	20	24	24
Negatif	14	9	0	9
Netral	2	0	0	0

Kemudian dari semua perhitungan di hitung rata-ratanya (*average*), hasil *averagenya* sebagaimana di table 4.12. Terlihat bahwa untuk recall di eksperimen ini algoritma Neural Network mendapatkan hasil paling tinggi.

Tabel 15. hasil Average recall eksperimen Anis

Average Recall	NB	DT	NN
	0,83	1,00	1,00
	0,64	0,00	0,64
	0,00	0,00	0,00
	49%	33%	55%

Selanjutnya dilakukan perhitungan tingkat akurasi proses klasifikasi terhadap tiga algoritma. Nilai akurasi dari percobaan tersebut diuraikan dalam table 4.13.

Tabel 16. Akurasi algoritma

Algoritma	Akurasi
Nive Bayes (NB)	72,5%
Decision Tree (DT)	60%
Neural Network (NN)	80%

Dari nilai-nilai akurasi pada tabel 4.13, maka dalam hal ini akurasi terbaik untuk eksperimen agus, adalah *Neural Network*.

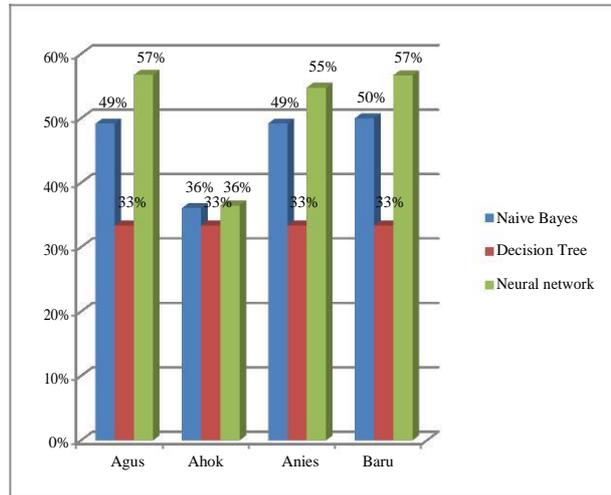
4.3 Anaisis Perbandingan tiga Algoritma

Hasil perbandingan recall dan accuracy dari ketiga algoritma yang telah diuji coba terhadap tiga tokoh public terlihat pada table 17.

NO	Nama Tes	Hasil					
		Recall			Accuracy		
		NB	DT	NN	NB	DT	NN
1	Agus	49,00%	33,00%	57,00%	78,00%	80,00%	85,00%

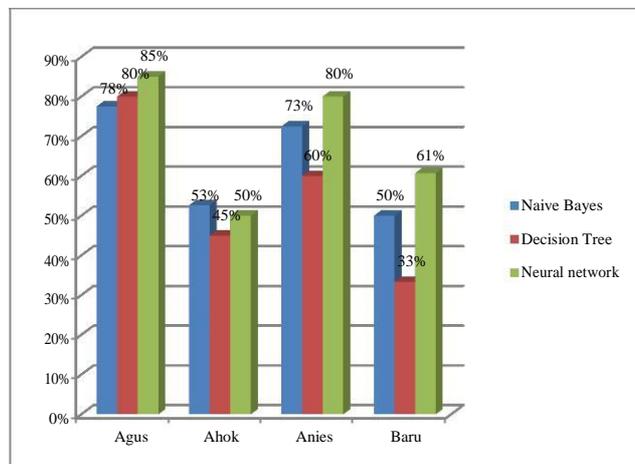
2	Ahok	33,00%	36,00%	33,00%	53,00%	45,00%	50,00%
3	Anies	49,00%	33,00%	57,00%	73,00%	60,00%	80,00%
Average		43,67%	34,00%	49,00%	68,00%	61,67%	71,67%

Pada tabel 17 terlihat hasil *recall* dan *accuracy* serta perbandingannya, Naive Bayes, Decision Tree dan Neural Network.



Gambar 4. Chart Recall dari tabel 4.14

Gambar 4.1 merupakan *chart* hasil dari perhitungan *recall* bisa kita lihat bahwa disetiap percobaan, Agus, Ahok, Anies penerapan algoritma pada setiap percobaan dihasilkan bahwa, algoritma Neural network terlihat lebih tinggi dari Naive Bayes dan Decision Tree.



Gambar 5. Chart Accuracy dari table 4.14

Gambar 4.2 merupakan *chart* dari hasil *accuracy* bisa kita lihat bahwa disetiap percobaan, Agus, Ahok, Anies penerapan algoritma pada setiap percobaan dihasilkan bahwa, algoritma Neural network terlihat lebih tinggi dari Naive Bayes dan Decision Tree.

5. Kesimpulan

Penelitian ini telah menghasilkan perbandingan algoritma Naive Bayes, Decision Tree, Neural Network untuk *recall* dan *accuracy* Neural Network adalah yang terbaik dengan tingkat akurasi mencapai 85% dengan rata-rata terhadap pengujian ketiga tokoh adalah 71,61%. Untuk mendapatkat hasil klasifikasi terbaik diperlukan

preposes yang lebih baik pula. Semakin besar jumlah data training juga akan meningkatkan hasil akurasi dan recall terhadap proses klasifikasi.

Daftar Pustaka

- [1] P. Aliandu, "Sentiment Analysis on Indonesian Tweet," *Proc. Int. Conf. Information, Commun. Technol. Syst.*, pp. 203–208, 2013.
- [2] M. H. Rasyadi, "Analisis Sentimen Pada Twitter Menggunakan Metode Naïve Bayes (Studi Kasus Pemilihan Gubernur Dki Jakarta 2017)," 2017.
- [3] S. Suryono *et al.*, "ANALISIS SENTIMENT PADA TWITTER DENGAN MENGGUNAKAN," pp. 9–15, 2018.
- [4] I. Sunni and D. H. Widyantoro, "Analisis Sentimen dan Ekstraksi Topik Penentu Sentimen pada Opini Terhadap Tokoh Publik," *J. Sarj. Inst. Teknol. Bandung Bid. Tek. Elektro dan Inform.*, vol. 1, no. 2, pp. 200–206, 2012.
- [5] A. F. Hidayatullah and A. Sn, "Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter," *Semin. Nas. Inform. 2014*, vol. 2014, no. August 2013, pp. 0–8, 2014.
- [6] B. Liu, "Sentiment Analysis and Opinion Mining," no. May, 2012.
- [7] N. Saputra, T. B. Adji, and A. E. Permanasari, "Analisis Sentimen Data Presiden Jokowi dengan Preprocessing Normalisasi dan Stemming menggunakan Metode Naive Bayes dan SVM," *J. Din. Inform.*, vol. 5, no. 1, 2015.