

APLIKASI SENTIMENT MONITORING UNTUK TWITTER DENGAN ALGORITMA NAIVE-BAYES CLASSIFIER

Ade Chandra Saputra^{a,1,*}, Agus Sehatman Saragih^{b,2}

^a Universitas Palangkaraya, Palangkaraya

^b Universitas Palangkaraya, Palangkaraya

¹ adechandra@it.upr.ac.id *; ² assaragih@it.upr.ac.id;

* corresponding author

ARTICLE INFO

Keywords

sentiment monitoring

twitter

naive-bayes classifier

ABSTRACT

Every day there are millions of opinion spread across social networks. This is often utilized by various parties to determine the opinion and sentiment of the public towards the product, brand or figures that they hold. Given the abundance of data and opinions, it is not possible to do sentiment analysis manually. In this research, author performs design and implementation of sentiment monitoring application, that could monitor people's sentiment about a particular keyword, so it is known how the people response to those keywords, whether positive, negative or neutral.

From various existing social networks, Twitter is chosen as the source of data that will be monitored. Classification algorithm used here is Naive-Bayes Classifier with Boolean Multinomial model, and feature extraction using unigram word. The training data used is 400,000 data for each type of sentiment, so the total is 1.200.000 data. In the process of classification and training, application will perform stemming to take the root words contained within the tweet. Stemming algorithm used here is Confix Stripping.

The methodology of application development that used here is staged delivery. Implementation of application is done using PHP programming language. The result of this research is a sentiment monitoring application that can monitor public sentiment about a particular keyword in a particular time frame. From testing using k-fold cross validation, obtained accuracy rate for sentiment classification amounted to 85%.

1. Pendahuluan

Setiap harinya ada jutaan opini dan pendapat yang tersebar di jejaring sosial. Hal ini sering dimanfaatkan oleh perusahaan, tim marketing maupun tim sukses untuk mengetahui bagaimana opini dan penilaian masyarakat terhadap produk, brand atau tokoh yang mereka usung. Akan tetapi, karena banyaknya data dan opini yang tersedia, tidak memungkinkan dilakukan analisis secara manual untuk mengetahui sentimen masyarakat secara akurat. Aplikasi ini digunakan untuk memonitor sentimen masyarakat mengenai suatu keyword tertentu, sehingga diketahui bagaimana tanggapan masyarakat terhadap keyword tersebut, apakah positif, negatif atau netral.

Dari berbagai jejaring sosial yang ada, Twitter dipilih sebagai sumber data dan opini yang akan dianalisis. Karena sifatnya sebagai microblogging platform, setiap tweet memiliki panjang maksimal hanya 140 karakter sehingga cukup singkat untuk dianalisis. Selain itu, Twitter merupakan salah satu jejaring sosial paling populer di dunia (Alexa, 2016) yang setiap bulannya diakses oleh 320 juta user aktif dari seluruh dunia (Twitter, 2016).

Algoritma yang digunakan untuk melakukan klasifikasi sentimen pada penelitian ini adalah algoritma Naive-Bayes Classifier. Naive-Bayes Classifier adalah algoritma klasifikasi dengan teknik prediksi berbasis probabilistik sederhana yang dikembangkan dari teorema Bayes dengan menggunakan asumsi independensi yang kuat (Prasetyo, 2012). Kelebihan dari algoritma ini adalah tangguh terhadap atribut yang tidak relevan ataupun data-data yang terisolasi. Algoritma ini juga

mampu menangani nilai atribut yang salah dengan mengabaikan data latih selama proses pembangunan dan prediksi. Selain itu, meskipun merupakan salah satu algoritma klasifikasi tertua, akan tetapi karena sifatnya yang simpel menjadikannya sebagai salah satu algoritma yang paling umum digunakan (Wu et al, 2014).

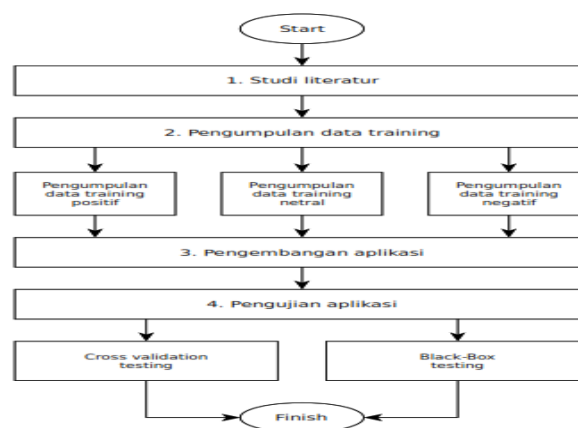
Model Naive-Bayes yang digunakan adalah model Boolean Multinomial. Pada model Multinomial, frekuensi munculnya token data mempengaruhi hasil klasifikasi (Raschka, 2014). Model ini dipilih karena berdasarkan hasil penelitian McCallum dan Nigam et al (dikutip dalam Raschka, 2014), model Multinomial secara empiris memiliki performa dan akurasi yang lebih tinggi dibandingkan model Naive-Bayes yang lain dalam melakukan klasifikasi teks, terutama jika jumlah kosakata yang digunakan cukup besar. Model Boolean berarti frekuensi kemunculan token data di setiap dokumen bersifat biner, yaitu ada (bernilai 1) dan tidak ada (bernilai 0). Model Boolean dipilih karena berdasarkan penelitian Metsis, Androutsopoulos dan Paliouras (2006) dapat memberikan hasil yang lebih akurat dibandingkan model Multinomial yang normal, walaupun dengan selisih yang relatif kecil. Selanjutnya, untuk ekstraksi token dalam kalimat digunakan teknik unigram di mana satu kata dihitung sebagai satu buah token. Hal ini juga di dukung oleh Sergios Theodoridis (2015) yang dituliskan dalam buku *A Bayesian and Optimization Perspective*.

Analisis Sentimen telah banyak dilakukan seperti mengetahui respons opini konsumen pada suatu produk (V. Ikoro dkk, 2018), preferensi politik (R. Sandoval, 2018), pendapat tentang film yang telah dirilis beserta prediksi pendapatan film dan lainnya (U. R. Hodeghatta, 2013). Penelitian terdahulu tentang Analisis Sentimen banyak dilakukan di berbagai bidang seperti bidang politik yang menganalisis sifat dan karakteristik wacana politik yang terjadi di Twitter selama pemilihan Presiden Amerika pada November 2016 (U. Yaqub, 2017). Penelitian selanjutnya tentang Analisis Sentimen untuk bidang musik yang memprediksi keberhasilan album musik, berdasarkan komentar yang diposting di jejaring sosial selama 30 hari sebelum rilis album melalui Twitter untuk mengumpulkan komentar pengguna (C. V. S. Araujo, 2017).

Analisis sentimen dalam penelitian ini untuk mengimplementasikan algoritma *Naive-Bayes Classifier* sehingga menghasilkan aplikasi yang dapat dimanfaatkan untuk memonitor sentimen masyarakat dari opini-opini dalam *tweet-tweet* berbahasa Indonesia di Twitter. Analisis sentiment ini untuk mengetahui tweet yang diposting mengandung sentimen positif, negative dan netral menggunakan klasifikasi bayes berdasarkan hastag pencarian yang dimasukan.

2. Metodologi Penelitian

Metode penelitian yang dilakukan dalam perancangan dan implementasi sentiment monitoring ini dapat digambarkan seperti gambar 1. Adapun penjelasan dari setiap tahap penelitian adalah sebagai berikut :



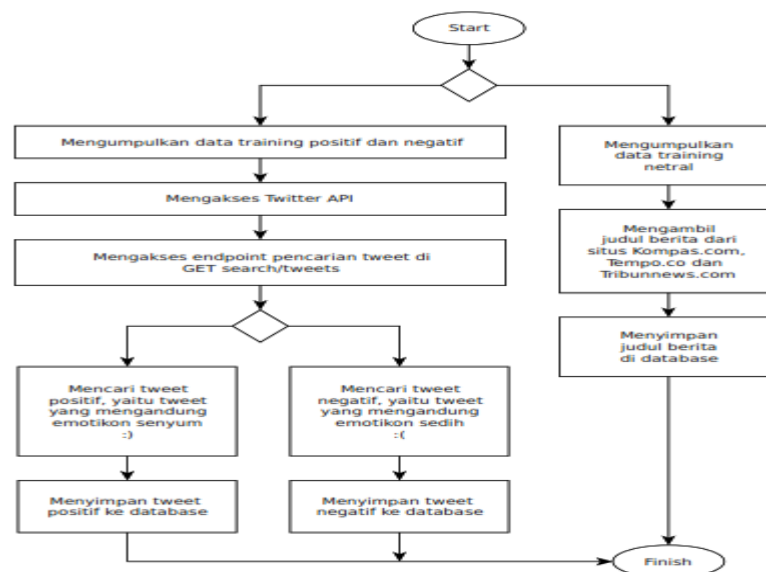
Gambar 1. Flowchart metode penelitian

2.1. Studi Literatur

Bahan referensi yang dikumpulkan adalah buku, paper, skripsi, situs internet atau literatur lainnya yang membahas teknik dan implementasi sentiment analysis, baik yang menggunakan algoritma Naive-Bayes atau yang menggunakan algoritma lainnya.

2.2. Pengumpulan Data Training

Data training yang dikumpulkan terdiri atas tiga kategori, yaitu untuk sentimen positif, negatif dan netral. Data *training* yang digunakan adalah 1.200.000 data, masing-masing 400.000 untuk setiap jenis sentimen. Untuk sentimen positif dan negatif, data *training* yang digunakan adalah *tweet* yang dikumpulkan menggunakan Twitter API. *Tweet* yang dikumpulkan adalah *tweet* yang tidak mengandung kata 'jual', 'grosir', 'eceran', 'follow' atau 'follback', bukan merupakan *retweet*, bukan merupakan jawaban dari percakapan dalam Twitter, serta tidak mengandung *link* URL. Untuk sentimen netral, data *training* yang digunakan adalah judul berita dari situs Kompas.com, Tempo.co dan Tribunnews.com. Judul berita digunakan mengikuti definisi dari penelitian Go, Bhayani dan Huang (2009), bahwa suatu *tweet* dianggap netral jika isi dari *tweet* tersebut dapat muncul sebagai judul artikel Wikipedia atau *headline* berita. Proses pengambilan data training dapat di lihat pada gambar 2.

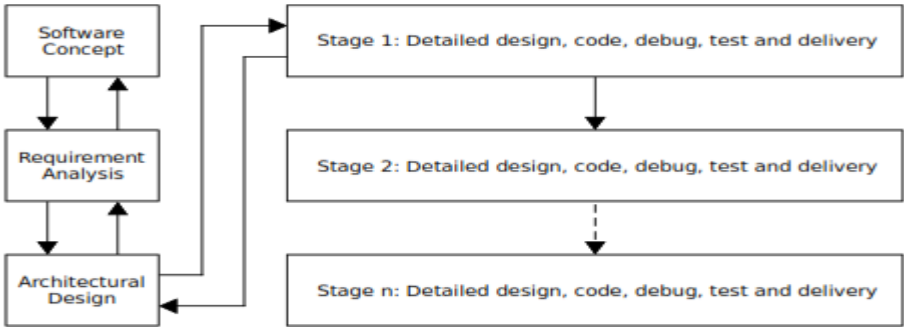


Gambar 2 Proses pengumpulan data *training*

Dari berbagai situs berita yang ada di Indonesia, situs Kompas.com, Tempo.co dan Tribunnews.com dipilih karena indeks berita di ketiga situs tersebut ditampilkan dalam bentuk HTML biasa, sehingga judul-judul berita di indeks tersebut dapat dikumpulkan dengan mudah menggunakan *web scraper*. Berbeda dengan kebanyakan situs berita lainnya yang menampilkan indeks berita menggunakan *Javascript* sehingga tidak dapat diambil dengan *web scraper* semudah ketiga situs tadi.

2.3. Pengembangan Aplikasi

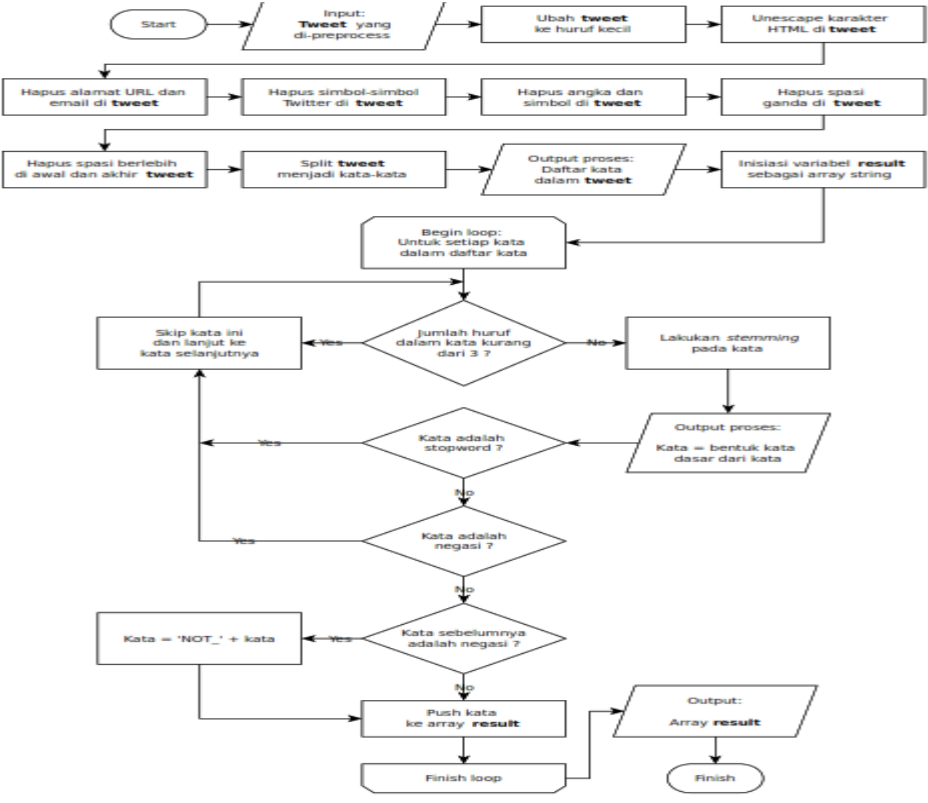
Pengembangan aplikasi dilakukan menggunakan metode staged delivery. Pada staged delivery, program dikerjakan dan dirilis secara bertahap dalam beberapa siklus sampai terbentuk sebuah aplikasi final. Model ini dapat digambarkan dengan diagram seperti gambar 3.



Gambar 3 Siklus staged delivery

2.4. Proses preprocessing data training dari twiiter

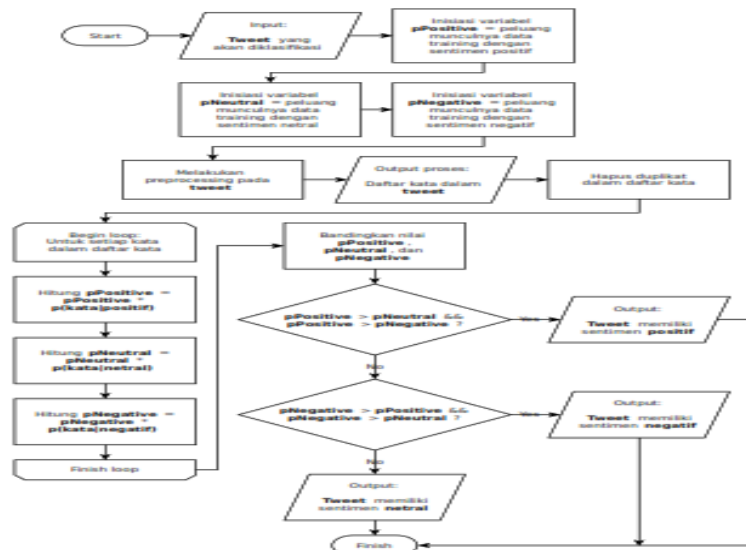
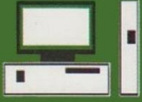
Preprocessing dilakukan sebelum melakukan training dan klasifikasi, untuk membersihkan *tweet* dan melakukan *stemming* sehingga menghasilkan output berupa *array* kata dasar yang terdapat dalam *tweet* tersebut. Flowchart preprocessing dapat dilihat di gambar 4



Gambar 4 poses preprocessing data training

2.5. Proses klasifikasi tweet

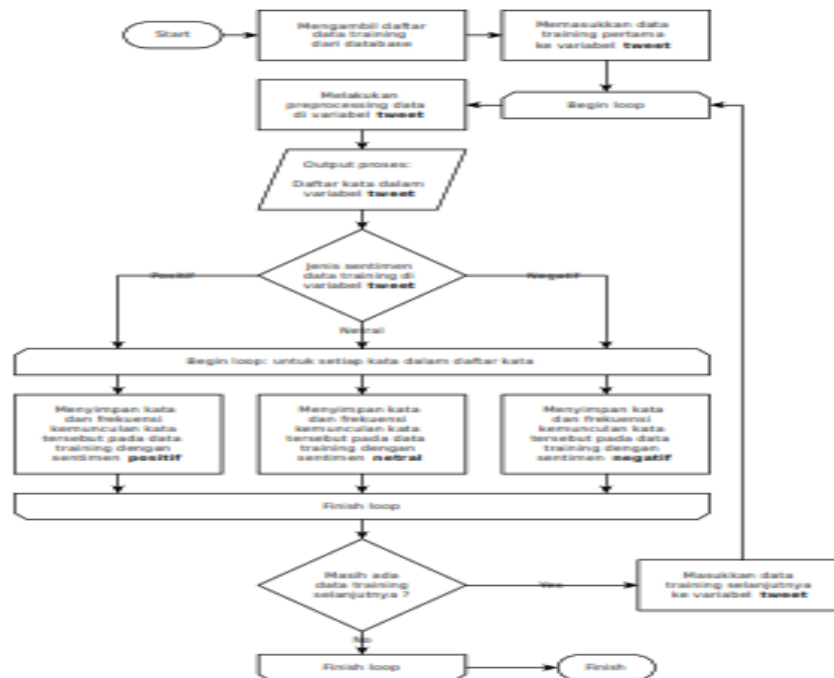
Klasifikasi *tweet* dilakukan untuk menentukan sentimen dari suatu *tweet*, apakah positif, negatif atau netral. Flowchart dari proses klasifikasi dapat dilihat pada gambar 5



Gambar 5 Proses klasifikasi

2.6. Proses training dengan naïve bayes

Proses *training* dapat dilihat pada gambar 6. Proses *training* ini dilakukan untuk mengubah data *training* menjadi daftar kata dan frekuensi kemunculan kata tersebut pada dokumen dengan sentimen positif, negatif atau netral.



Gambar 6 Proses Training Naïve Bayes

2.7. Pengujian Aplikasi

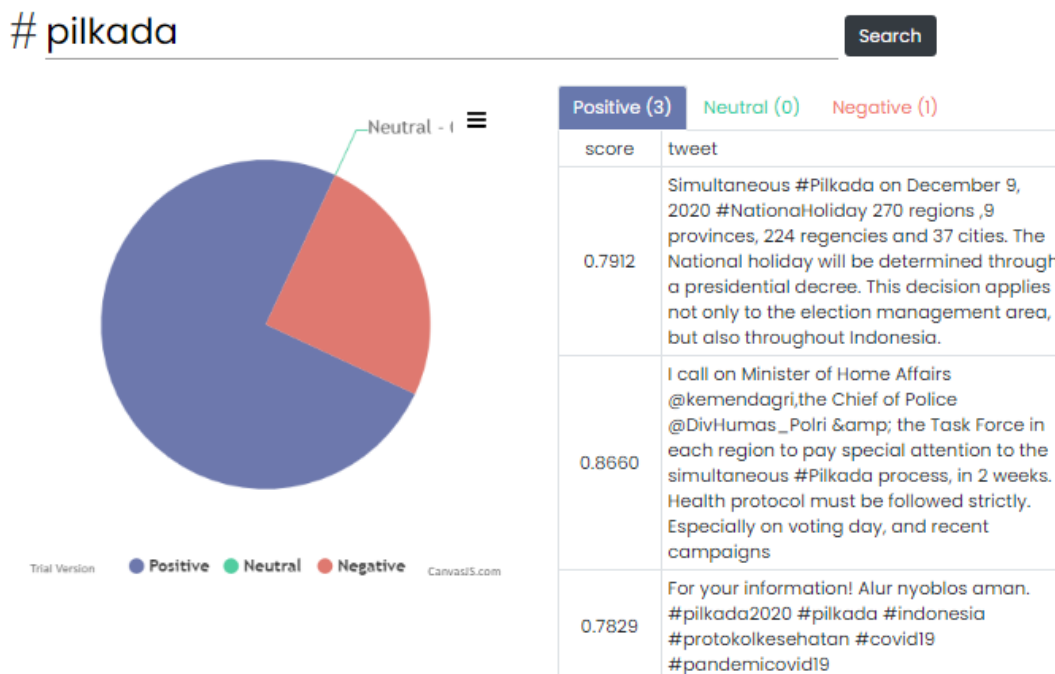
Pengujian aplikasi dilakukan dalam dua tahap, yaitu pengujian modul klasifikasi tweet dan pengujian GUI aplikasi. Pengujian modul klasifikasi dilakukan dengan dua metode. Metode yang pertama adalah dengan menguji 1 set data testing menggunakan 20 set data training, sedangkan metode yang kedua adalah dengan menggunakan k-fold cross validation. Adapun pengujian GUI dilakukan menggunakan black-box testing.

3. Hasil dan Pembahasan

Aplikasi sentiment analisis dengan klasifikasi naïve bayes dapat dilihat pada gambar 7 dibawah ini.

Analisis Sentimen Twitter dengan Klasifikasi Naive Bayes

Masukan type hastag atau keyword untuk proses klasifikasi

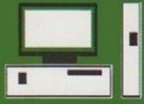


Gambar 7 Tampilan Aplikasi Analisis Sentimen dengan klasifikasi bayes

Pada halaman monitor terdapat tiga komponen utama. Yang pertama adalah penginputan hastag atau kata kunci yang dicari di tweet. Pada area kedua disebalah kanan ini dicantumkan informasi mengenai *tweet-tweet* yang terkumpul berdasarkan kata pencarian yg diinput berdasarkan sentimennya masing - masing yang terkandung di dalamnya beserata nilai score bayes. Yang ketiga adalah grafik sentiment berdasarkan sentiment positive netral dan negative. Melalui grafik ini dapat diketahui jumlah *tweet-tweet* dari masyarakat dengan sentimen tertentu setiap harinya. Pada gambar 7 diatas dicontohkan dengan kata kunci pilkada yang dihasilkan adalah ada 3 sentimen positif , 0 sentimen netral, dan 1 sentimen negative.

3.1. Pengujian system

Pengujian klasifikasi dilakukan menggunakan dua metode. Metode yang pertama adalah dengan menguji 1 set data testing menggunakan 20 set data training. Data testing yang digunakan adalah



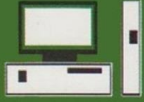
300.000 data, sedangkan data training yang digunakan adalah 20 set data dengan jumlah data yang berbeda- beda, mulai dari 45.000 data sampai 900.000 data. Setiap data set memiliki rasio jumlah sentimen positif, negatif dan netral yang sama. Hasil pengujiannya adalah sebagai berikut :

Tabel 1 Hasil pengujian 300.000 data dengan 20 set data *training*

Skenario	N Training	Akurasi (dalam persen)			
		Positif	Negatif	Netral	Keseluruhan
1	45.000	78,86	75,70	93,35	82,64
2	90.000	79,00	77,27	94,03	83,43
3	135.000	78,80	78,29	94,24	83,78
4	180.000	78,83	78,86	94,39	84,03
5	225.000	78,90	79,12	94,44	84,15
6	270.000	78,96	79,37	94,53	84,29
7	315.000	78,90	79,52	94,58	84,33
8	360.000	78,94	79,79	94,64	84,46
9	405.000	78,91	79,89	94,68	84,49
10	450.000	78,90	80,03	94,70	84,54
11	495.000	78,97	80,09	94,74	84,60
12	540.000	78,97	80,22	94,78	84,66
13	585.000	78,96	80,36	94,77	84,70
14	630.000	78,97	80,37	94,80	84,71
15	675.000	79,01	80,46	94,82	84,76
16	720.000	78,98	80,45	94,82	84,75
17	765.000	78,99	80,59	94,85	84,81
18	810.000	79,00	80,66	94,85	84,83
19	855.000	79,03	80,70	94,87	84,87
20	900.000	79,02	80,72	94,87	84,87

Dari pengujian dengan metode pertama ini, didapat hasil sebagai berikut :

1. Modul klasifikasi yang dibuat dapat mengklasifikasikan sentimen netral dengan akurasi mencapai 94,87%, sentimen positif dengan akurasi 79,03% dan sentimen negatif dengan akurasi 80,72%.
2. Jumlah data training yang digunakan tidak memiliki pengaruh yang signifikan pada tingkat akurasi klasifikasi. Hal ini terlihat antara skenario 1 dan 20 selisih akurasi cukup kecil, yaitu secara keseluruhan hanya terjadi perubahan sebanyak 2,23%.
3. Metode yang kedua adalah pengujian menggunakan k-fold cross validation dengan nilai $k = 20$. Ada tiga skenario yang digunakan, yaitu :



- Menggunakan 1.200.000 data, yang terdiri atas masing-masing 400.000 data untuk setiap jenis sentimen.
- Menggunakan 900.000 data, yang terdiri atas masing-masing 300.000 data untuk setiap jenis sentimen.
- Menggunakan 600.000 data, yang terdiri atas masing-masing 200.000 data untuk setiap jenis sentimen.

Pengujian dengan tiga skenario di atas memberikan hasil sebagai berikut :

Tabel 2 Hasil pengujian 20-fold cross validation pada 3 skenario

Skenario	Akurasi (dalam persen)			
	Positif	Negatif	Netral	Keseluruhan
1	76,58	82,92	97,42	85,64
2	78,24	81,53	97,54	85,77
3	78,07	81,10	97,44	85,54

Dari pengujian dengan metode kedua ini, didapatkan hasil :

1. Dari semua skenario pengujian, modul klasifikasi yang dibuat dapat mengklasifikasikan sentimen netral dengan akurasi lebih dari 97%, sentimen negatif antara 76-78% dan sentimen positif antara 81-83%.
2. Jumlah data training yang digunakan tidak memiliki pengaruh yang signifikan pada tingkat akurasi klasifikasi. Hal ini terlihat pada ketiga skenario tersebut selisih akurasi keseluruhan sangatlah kecil, yaitu kurang dari 0,5%.

Dari kedua metode pengujian yang telah dilakukan, dapat ditarik hasil dan kesimpulan sebagai berikut :

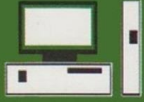
1. Jumlah data training yang digunakan tidak memiliki pengaruh yang signifikan pada tingkat akurasi klasifikasi.
2. Modul klasifikasi yang dibuat dapat mengklasifikasi sentimen dengan akurasi yang cukup tinggi, yaitu untuk sentimen netral mencapai 97%, sentimen positif dengan akurasi sekitar 78% dan sentimen negatif dengan akurasi sekitar 80%. Secara keseluruhan, tingkat akurasi klasifikasi sentimen yang didapat adalah sekitar 85%.

4. Kesimpulan

Aplikasi sentiment analisis sudah mampu menganalisis sentime negative, positif dan netral dari twitter. Dengan menggunakan dua metode pengujian yaitu Metode yang pertama adalah pengujian 1 set data testing menggunakan 20 set data training, sedangkan metode yang kedua adalah menggunakan 20-fold cross validation dengan tiga skenario jumlah data. Dari pengujian ini, diketahui bahwa modul klasifikasi yang dibuat dapat mengklasifikasi sentimen dengan akurasi yang cukup tinggi, yaitu untuk sentimen netral mencapai 97%, sentimen positif dengan akurasi sekitar 78% dan sentimen negatif dengan akurasi sekitar 80%. Secara keseluruhan, tingkat akurasi klasifikasi sentimen yang didapat adalah sekitar 85%.

Daftar Pustaka

- [1] Al Fatta, H. (2017). Analisis dan Perancangan Sistem Informasi untuk Keunggulan Bersaing Perusahaan dan Organisasi Modern. Yogyakarta: Penerbit Andi
- [2] Arifin, Z.A.; Mahendra, I.P.A.K. & Ciptaningtyas, H.T. (2009). Enhanced Confix Stripping Stemmer and Ants Algorithm for Classifying News Document in Indonesian Language. Dalam acara 5th International Conference on Information & Communication Technology and Systems (ICTS).
- [3] Ghulam Asrofi Buntoro. (2017). Sentiment Analysis to Prediction DKI Jakarta Governor 2017 on Indonesian Twitter. International Journal of Science, Engineering and Information Technology Trunojoyo.
- [4] Hidayatullah, A.F. & Azhari S.N. (2013). Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik pada Twitter. Dalam acara Seminar Nasional Informatika 2014 (semnasIF 2014) UPN "Veteran" Yogyakarta. Yogyakarta, 12 Agustus 2014.



- [5] Luhulima; Y.Y, Mardji dan Muflikhah L. (2013). “Sentiment Analysis pada Review Barang Berbahasa Indonesia dengan Metode K-Nearest Neighbor (K-NN)”. Repositori Jurnal Mahasiswa PTIIK UB. 2, (5).
- [6] Metsis, V., Androutsopoulos, I., & Paliouras, G. (2006). Spam filtering with naive bayes-which naive bayes?. In CEAS, 17, 28-69.
- [7] Sergios Theodoridis. (2015). Machine Learning: A Bayesian and Optimization Perspective, Elsevier 2015.
- [8] V. Ikoru, M. Sharmina, K. Malik, and R. Batista-Navarro, —Analyzing Sentiments Expressed on Twitter by UK Energy Company Consumers,| 2018 Fifth Int. Conf. Soc. Networks Anal. Manag. Secur., pp. 95–98, 2018.
- [9] U. R. Hodeghatta, —Sentiment Analysis of Hollywood Movies on Twitter,| Proc. 2013 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Min., pp. 1401–1404, 2013.
- [10] R. Sandoval-Almazan and D. Valle-Cruz, —Facebook Impact and Sentiment Analysis on Political Campaigns,| Proc. 19th Annu. Int. Conf. Digit. Gov. Res. Gov. Data Age, pp. 561–567, 2018
- [11] U. Yaqub, S. A. Chun, V. Atluri, and J. Vaidya, —Sentiment based Analysis of Tweets during the US Presidential Elections,| Present. Proc. 18th Annu. Int. Conf. Digit. Gov. Res., pp. 149–156, 2017
- [12] C. V. S. Araujo, R. M. Neto, F. N. G., and E. F. Nakamura, —Predicting Music Success Based on Users’ Comments on Online Social Networks,| Proc. 23rd Brazillian Symp. Multimed. Web, pp. 1–10, 2017