



Analisis Sentimen Komentar Pengunjung Air Terjun Waikelo Sawa Desa Tema Tana Wewewa Timur Kabupaten Sumba Barat Dengan Metode Naïve Baiyes Classifier

¹Amatus Simson Lende, ²Gregorius Kopong Pati, ³Alexander Adis
^{1,2,3}STIMIKOM Stella Maris Sumba

Alamat Surat:

Email: amatussimsonlende06@gmail.com*, gregkopong80@gmail.com

Article History:

Diajukan: 17 September 2024; Direvisi: 15 Oktober 2024; Diterima: 28 Oktober 2024

ABSTRAK

Analisis Sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif. Analisis sentimen diberikan oleh pengguna internet pada media sosial untuk memberikan suatu penilaian atau opini pribadi. Air Terjun Wekelo Sawa yang sering mendapatkan sentimen pengguna melalui media sosial adalah Air Terjun Wekelo Sawa. Adanya sentimen opini dari konsumen tentang wisata Air Terjun Wekelo Sawa dapat dianalisis dan dimanfaatkan untuk mendapatkan informasi yang berguna bagi pelanggan lain maupun pihak pengolah Air Terjun Wekelo Sawa. Dengan menggunakan teknik Text Mining metode klasifikasi, akan diketahui suatu sentimen bernilai positif, netral atau negatif. Salah satu algoritme yang banyak digunakan dalam analisis sentimen adalah metode klasifikasi Naïve Bayes. Penelitian ini menggunakan metode Naïve Bayes Classifier (NBC) dengan pembobotan tf-idf disertai penambahan fitur konversi ikon emosi (emoticon) untuk mengetahui kelas sentimen yang ada dari tweet tentang wisata Air Terjun Wekelo Sawa. Hasil penelitian menunjukkan bahwa metode Naïve Bayes tanpa penambahan fitur mampu mengklasifikasi sentimen dengan nilai akurasi sebesar 88,46%, sementara jika ditambahkan fitur pembobotan tf-idf disertai konversi ikon emosi mampu meningkatkan nilai akurasi menjadi 88,46%.

Kata Kunci: Naïve Bayes Classifier, Air Terjun

ABSTRACT

Sentiment Analysis is a technique for extracting text data to obtain information about positive, neutral or negative sentiment. Sentiment analysis is provided by internet users on social media to provide a personal assessment or opinion. Wekelo Sawa Waterfall which often gets user sentiment through social media is Wekelo Sawa Waterfall. The sentiment of opinions from consumers regarding Wekelo Sawa Waterfall tourism can be analyzed and utilized to obtain useful information for other customers and the Wekelo Sawa Waterfall processor. By using the Text Mining technique, the classification method will determine whether a sentiment is positive, neutral or negative. One algorithm that is widely used in sentiment analysis is the Naïve Bayes classification method. This research uses the Naïve Bayes Classifier (NBC) method with tf-idf weighting accompanied by the addition of an

emotional icon (emoticon) conversion feature to determine the existing sentiment classes from tweets about Wekelo Sawa Waterfall tourism. The research results show that the Naïve Bayes method without adding features is able to classify sentiment with an accuracy value of 88.46%, while if the tf-idf weighting feature is added along with emotional icon conversion it is able to increase the accuracy value to 88.46%.

Keywords: *Naïve Bayes Classifier, Waterfall*

1. PENDAHULUAN

Wisata alam Indonesia memiliki potensi untuk dapat dikembangkan. Pengembangan tersebut antara lain, tempat wisata dapat dijadikan sebagai ciri khas suatu daerah. Selain itu, dapat pula di tambahkan sebuah sarana informasi bagi wisatawan khususnya wisatawan luar daerah, sarana tersebut berupa sebuah web yang berisikan informasi pendukung keputusan dalam memilih sebuah tempat wisata. Hampir diseluruh wilayah Indonesia memiliki keindahan alam yang cukup baik untuk dapat dijadikan daerah tujuan wisata. Iklim tropis Indonesia yang bersifat panas membuat Indonesia memiliki curah hujan yang cukup tinggi. Kondisi tropis semacam inilah yang menyebabkan Indonesia secara keseluruhan memiliki pemandangan alam yang cukup bervariasi jumlahnya layaknya negara-negara tropis lainnya (Ridaini, 2014).

Sumba Barat Daya merupakan salah satu kabupaten yang ada di pulau Sumba Nusa Tenggara Timur dengan julukan pulau Cendana atau pulau Sandelwood, Sumba memiliki pesona wisata alam yang dapat dikunjungi. Salah satunya Kabupaten Sumba Barat Daya menyimpan banyak sekali tempat wisata yang sangat indah, sehingga sulit bagi para wisatawan untuk menentukan tempat wisata yang harus di kunjungi. Dalam memilih tempat wisata tidaklah mudah karena terdapat beberapa faktor-faktor yang perlu dipertimbangkan dalam melakukan pemilihan 2 objek wisata yang ingin dikunjungi. Faktor-faktor tersebut antara lain akses jalan, jarak tempuh, transportasi umum dan tingkat keramaian ditempat wisata.

Salah satu jenis sistem informasi adalah sistem pendukung keputusan. Alter dalam Kadir (2003) menyatakan bahwa sistem pendukung keputusan adalah sistem informasi interaktif yang menyediakan informasi, pemodelan, dan pemanipulasian data yang digunakan untuk membantu pengambilan keputusan pada situasi semiterstruktur dan situasi yang tidak terstruktur. Sistem dapat membantu pengambilan keputusan dalam pemilihan obyek wisata terbaik menurut nilai prioritas terhadap aspek pariwisata dimana nilai tersebut ditentukan sendiri oleh calon wisatawan.

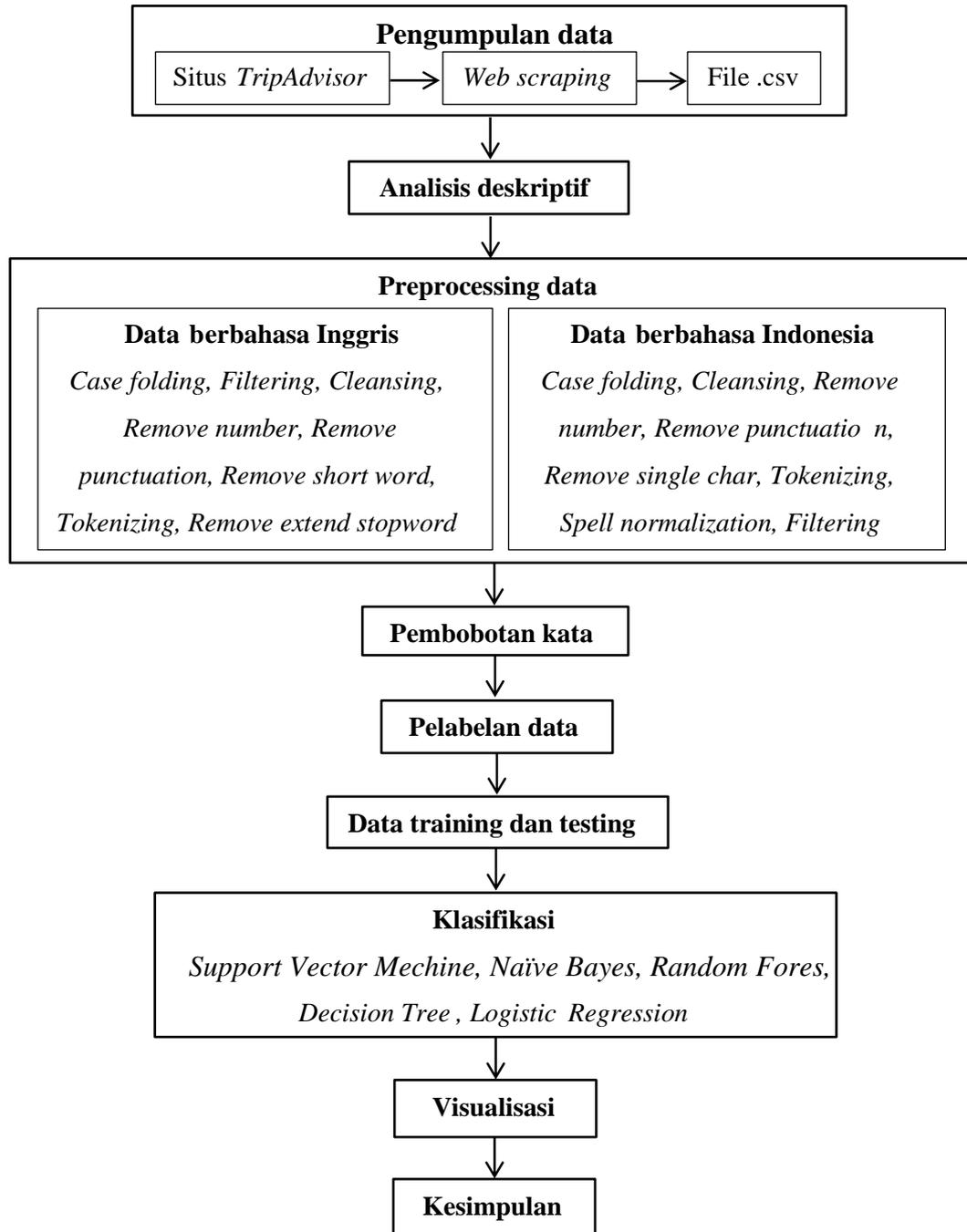
Problematika yang dihadapi masyarakat di Desa Tema Tana dalam bidang kepariwisataan masih banyak pekerjaan yang belum dilaksanakan dengan tepat, di antaranya yaitu implementasi kebijakan pengembangan objek wisatanya masih belum optimal. wisata menjadi salah satu faktor menyulitkan. Kebudayaan masyarakat di Desa Tema Tana untuk memantau perkembangan wisata yang ada di wilayahnya. Tanggapan dan komentar dari pengunjung wisata sangat diperlukan untuk pengembangan objek wisata. Dengan adanya feedback dari pengunjung diharapkan mempermudah masyarakat di Desa Tema Tana menentukan langkah apa yang harus dilakukan dalam pengembangan objek wisata yang ada.

Dengan majunya perkembangan teknologi informasi saat ini membawa banyak manfaat dan keuntungan bagi orang lain. Perkembangan teknologi informasi juga berbanding lurus dengan berkembangnya pariwisata. Pengunjung wisata dapat memberikan komentar terhadap wisata yang dikunjungi. Salah satu aplikasi dan media sosial yang sering digunakan untuk melihat dan memberikan komentar wisata yaitu Tripadvisor dan Facebook. Dalam aplikasi Tripadvisor dan Facebook terdapat banyak informasi wisata dan komentar pengunjung yang sangat membantu untuk mengetahui kualitas dari wisata yang akan dikunjungi. Dengan demikian pengguna lain dapat mudah menemukan informasi apa yang sedang diinginkan. Analisis Sentimen adalah jenis natural bahasa yang dapat dilakukan untuk mendapatkan Opinion Public. Dengan adanya analisis sentimen dapat dilakukan pengolahan kata untuk 2 melacak mood pengunjung wisata dari komentar pengunjung yang didapatkan dalam aplikasi Tripadvisor dan Facebook.

Oleh karena itu, penelitian ini akan menawarkan solusi dalam melakukan analisis sentimen terhadap komentar pengunjung wisata di Air Terjun Waikelosawa. Analisis sentimen ini digunakan untuk mengetahui performa Naive Bayes Classifier dalam melakukan klasifikasi berdasarkan komentar pengunjung pariwisata yang ada di air terjun waikelosawa, sehingga mempermudah Kebudayaan di Desa Tema Tana dalam memantau perkembangan wisata dan membantu dalam pengambilan keputusan untuk langkah apa yang tepat dalam mengembangkan wisata yang ada. Untuk membuat sebuah analisis sentimen perlu banyak hal dilakukan, di antaranya memilih classifier dan sumber data yang akan digunakan untuk analisis.

2. METODE

Pendekatan dan Desain



Lokasi Penelitian

Menurut Lokasi penelitian yang dilakukan penulis pada Waikelo Sawa Kabupaten Sumba Barat Daya Provinsi Nusa Tenggara Timur.

Data dan Sumber Data

Sumber informasi adalah subjek dari mana informasi itu diperoleh. Penelitian kualitatif memiliki sumber data primer dari ucapan, Bahasa, tindakan, serta data tambahan seperti dokumen dan sebagainya. Sumber data dalam penelitian ini antara lain

(Sugiyono, 2016).

Data Primer

Sumber data primer dikumpulkan dari observasi langsung dan wawancara dengan informasi tertentu. Data awal diambil dari informasi utama Dinas Pariwisata Kabupaten Sumba Barat Daya.

Data Sekunder

Sumber informasi skunder adalah sumber melingkar yang cocok untuk memberikan tambahan dan membangun informasi penelitian. Sumber informasi skunder diperoleh baik melalui individu kedua maupun dari data yang tidak dating secara langsung, data ini mendukung pembahasan dan penelitian karena berbagai sumber buku atau data yang diperoleh akan membantu dan mengkritisi penelitian. Untuk memperoleh data tersebut penelitian melakukan sejumlah wawancara, buku, website, apa lagi contoh eksplorasi masa lalu yang berhubungan dengan penelitian yang sedang diteliti.

Teknik Pengumpulan Data

Pengumpulan data dilakukan dengan cara mengambil data ulasan pengunjung objek wisata Lawang Sewu dari kolom komentar situs TripAdvisor dengan cara web scraping menggunakan package Selenium dengan bahasa pemrograman Python.

Teknik Analisis Data

Analisis data merupakan proses mencari dan menyusun secara sistematis data yang telah diperoleh dari wawancara, dokumentasi, dokumen pribadi, observasi, catatan lapangan gambar foto dan sebagainya, dengan cara mengorganisasikan data tersebut kedalam kategori, menjabarkan ke dalam unit –unit, melakukan sintesa, menyusun ke dalam pola, memilih mana yang penting dan mana yang akan dipelajari dan kemudian membuat kesimpulan agar dapat dipahami diri sendiri ataupun orang lain.

3. HASIL DAN PEMBAHASAN

Sumber Data

Proses data pada penelitian ini membagi data-data yang telah dilakukan proses preprocessing menjadi dua bagian yaitu data latih dan data uji. Pembuatan data latih ini bertujuan sebagai pedoman dalam menentukan hasil data uji. Data komentar yang disimpan dalam excel bisa dilihat pada gambar dibawah ini:

	text	sentimen
1		
2	Airnya jernih	positif
3	tidak jauh dari desa	positif
4	Aman saat berenang	positif
5	Harganya terjangkau	positif
6	Tempatnya strategis	positif
7	selalu dijaga keamanannya	positif
8	Tempatnya sangat menakutkan	positif
9	Airnya jernih, Harganya bisa dijangkau	positif
10	Tempatnya sangat bagus untuk dikunjungi	positif
11	penduduknya ramah dalam memberikan pelayanan	positif
12	saya puas dengan keadaan lingkungan air terjun	positif
13	Penduduknya selalu membawa perang pada saat memasuki air terjun	negatif
14	Takut berenang karena terlalu dalam air terjunnya	negatif
15	Harga tiket masuk terlalu tinggi	negatif
16	Jernih air dan pemandangan sangat indah	positif
17	Puas dengan jernihnya air terjun	positif
18	Tempatnya bisa dijangkau	positif

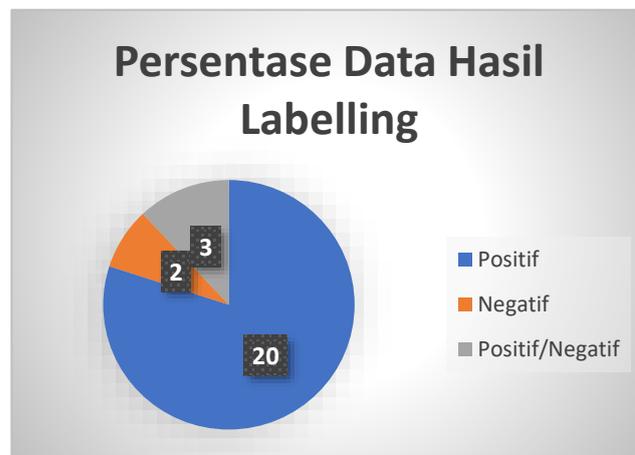
Gambar 1 Contoh Data Latih

Persentase Data Hasil Labelling

Pada proses pengambilan data yang dilakukan berdasarkan koisioner yang dibagikan kepengunjung dengan menghasilkan data dengan jumlah 26 komentar.

Setelah seluruh komentar berhasil disimpan selanjutnya dilakukan proses pelabelan. Dalam proses pelabelan yang dilakukan seharusnya untuk menentukan komentar yang didapatkan apakah bernilai positif, negative. amun pada proses pelabelan pada penelitian ini terdapat beberapa komentar yang tidak sesuai dengan sentimen komentar yang sebenarnya.

Dan hasil akhir dari proses pelabelan data menghasilkan data sebanyak 24 komentr bernilai positif, 2 komentar bernilai negatif. Persentase data bernilai positif, negatif dan positif negatif dan dapat dilihat pada gambar:



Gambar 2 Persentase Data Hasil Labelling

Ekstraksi Fitur

Setelah terbentuknya file yang akan dijadikan dataset, maka selanjutnya data tersebut akan dibentuk menjadi sebuah model klasifikasi. Namun sebelum membentuk model, ada beberapa tahapan yang harus dilakukan agar terbentuknya suatu model yang baik. Yang pertama dilakukan adalah membaca file xlsx dan kemudian dilakukan tokenisasi terhadap seluruh

dokumen dalam file tersebut. Berdasarkan hasil tokenisasi yang dilakukan, maka penulis juga ingin mengetahui frekuensi kata yang banyak diperbincangkan oleh pengunjung air terjun, untuk itu penulis memvisualisasikannya dalam bentuk wordcloud pada gambar 3.



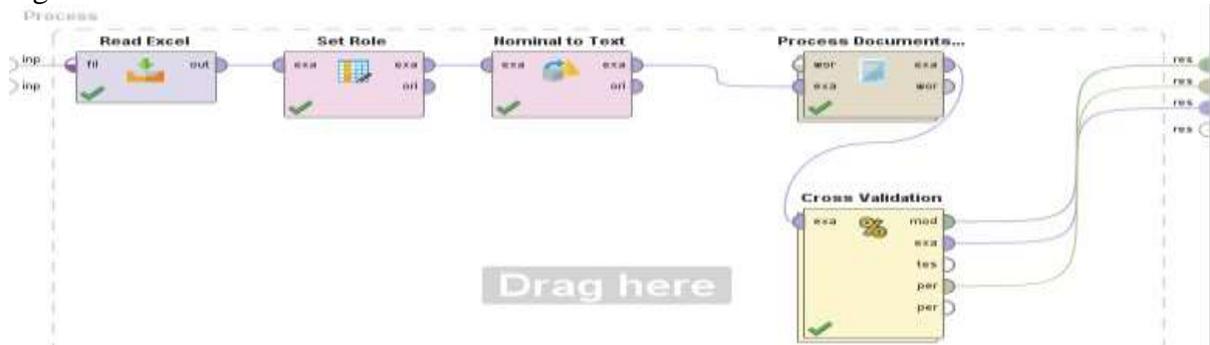
Gambar 3 Proses Document From Data

Perhitungan Akurasi Dengan RapidMiner

Penelitian ini menggunakan tools RapidMiner versi 10.1. perhitungan akurasi akan proses dengan menggunakan metode Naive Bayes Classifier dan Support Vektor Machine. Untuk mengetahui tingkat akurasi dari kedua metode tersebut.

Proses Dokumen dengan metode Naïve Bayes Classifier

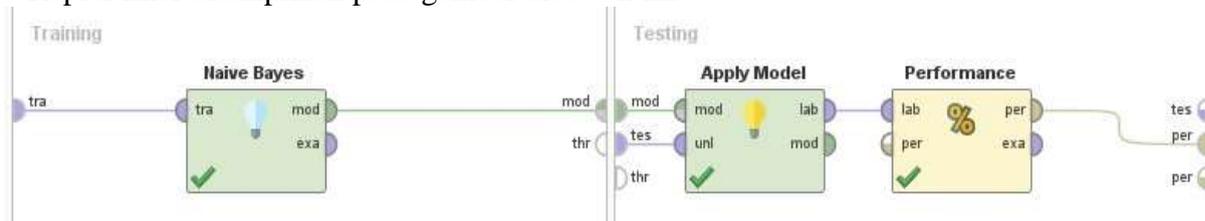
Pada proses ini diuji dengan data latih dari excel. Seperti yang terlihat pada gambar dibawah ini.



Gambar 4 Proses Dokumen Data Latih

Perhitungan Akurasi dengan Naïve Bayes Classifier pada RapidMiner

Proses *Training* dan *Testing* dengan metode Naive Bayes Classifier pada RapidMiner ditampilkan pada gambar dibawah ini:



Gambar 5 Proses *Training* dan *Testing* dengan Naive Bayes Classifier pada RapidMiner

Hasil perhitungan klasifikasi menggunakan metode Metode Naive Bayes Classifier diperoleh sebesar 88.46% seperti yang terlihat pada gambar dibawah.

Table View Plot View

accuracy: 88.46% ±1.3258% (micro average: 88.46%)

	true positif	true negatif	class precision
pred positif	23	3	88.46%
pred negatif	0	0	0.00%
class recall	100.00%	0.00%	

Gambar 6 Akurasi dengan Metode Naive Bayes Classifier pada RapidMiner

Hasil Pengujian Naïve Bayes

Akurasi yang diuji pada tools RapidMiner dengan menggunakan metode Naive Bayes Classifier. Hasil pengujian akurasi menggunakan metode Naive Bayes Classifier sebesar 88,46%. oleh karena itu tingkat akurasi yang diperoleh dengan menggunakan metode yang diuji pada tools RapidMiner.

Tabel Hasil Perhitungan Akurasi

Metode	Tingkat Predikat	True Positif	True Negatif
Naive Bayes Classifier	Pred. Positif	23	3
	Pred. Negatif	0	0
Akurasi pada tools RapidMiner			
Naive Bayes Classifier		88,46%	

4. SIMPULAN DAN SARAN

Berdasarkan Hasil pengujian akurasi menggunakan metode Naive Bayes Classifier adalah: tingkat akurasi dengan metode Naive Bayes Classifier sebesar 88,46%. memiliki hasil akurasi yang cukup baik untuk menganalisis sentimen komentar atau tweet dari pengunjung. Hasil tingkat akurasi juga menunjukkan adanya kepuasan tentang air terjun. Saran yang hendak disampaikan untuk peneliti selanjutnya adalah menggunakan metode yang berbeda sehingga bisa diketahui tingkat perbandingan kepuasan pengunjung

5. DAFTAR PUSTAKA

- G. K., & E. U. (2022). Analisis Sentimen Komentar Pengunjung Terhadap Tempat Wisata Danau Weekuri Menggunakan Metode Naive Bayes Classifier Dan K-Nearest Neighbor. *JURNAL MEDIA INFORMATIKA BUDIDARMA*.
- J. C. (2022). Penerapan Algoritma Sentiment Analysis Dan Naïve Bayes Terhadap Opini Pengunjung Di Tempat Wisata Pantai Pintu Kota, Kota Ambon. *Jurnal Teknik Informatika dan Sistem Informasi*.
- L. W., & dkk. (2017). Analisis Sentimen Terhadap Tempat Wisata Dari Komentar Pengunjung Dengan Menggunakan Metode Naïve Bayes Classifier Studi Kasus Jawa Barat. *Jurnal Prosiding SNATIF*.
- R. F., & dkk. (2019). Analisis Sentimen Wisata Jawa Tengah Menggunakan Naïve Bayes. *Jurnal INFORMA Politeknik Indonusa Surakarta*.
- Rozi, I. F., S. H. Pramono, dan E. A. Dahlan. 2012. Implementasi Opinion Mining (Analisis Sentimen) untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi. *Jurnal EECCIS 6(1): 37-43*.
- Kumar, E. (2011). *NATURAL LANGUAGE PROCESSING*. New Delhi: LK. Intrnational Publishing House Pvt, Ltd.
- Prasetyo, E. (2012). *Data Mining Konsep dan Aplikasi menggunakan MATLAB*. Yogyakarta: Penerbit Andi.
- Hidayatullah, A. F. (2014). *ANALISIS SENTIMEN DAN KLASIFIKASI KATEGORI TERHADAP TOKOH PUBLIK PADA TWITTER*. Seminar Nasional Informatika 2014 (semnasIF 2014). Yogyakarta.