

ANALISIS SENTIMEN TRANDING NEWS JOE BIDEN MENGGUNAKAN METODE MONKEYLEARN DAN ALGORITMA NAÏVE BAYES

Dhafi Yarist Kusnaedi¹⁾, Dimas Yusril Faidil Hisyam²⁾

^{1,2}Program Studi Informatika Fakultas Teknik Universitas Siliwangi
e-mail: 177006003@student.unsil.ac.id¹, 177006060@student.unsil.ac.id²

Abstrak

Saat ini media sosial merupakan salah satu media *online* yang banyak di akses di seluruh dunia. Dengan adanya layanan microblogging seperti twitter, memungkinkan pengguna menulis dengan tujuan untuk dipublikasikan, menulis opini tokoh publik, jasa layanan, dan lain sebagainya. Itu semua menjadi daya tarik bagi pengguna social media untuk menggunakan twitter. Hal tersebut dapat dimanfaatkan untuk pengambilan opini atau sentimen terhadap suatu entitas yang sedang dibicarakan di media social seperti twitter. Penelitian ini memanfaatkan data tersebut untuk mengetahui opini atau sentimen dari pengguna twitter mengenai tokoh publik yang sedang trading yaitu Joe Biden. Pengambilan opini yang akan diambil berdasarkan tiga kelas yaitu positif, negatif, dan netral. Pada penelitian ini penulis menerapkan text-processing MonkeyLearn untuk mengatasi masalah dalam mendapatkan data yang dapat ditindak lanjuti dari teks mentah. Metode klasifikasi yang digunakan adalah algoritma naïve bayes classifier. Dari hasil penelitian yang telah dilaksanakan dapat diketahui bahwa persentase sentimen pada social media twitter mengenai tanggapan berita terkait Joe Biden menghasilkan sentimen netral yang lebih dominan dari pada sentimen positif dan negatif. Selain itu dari hasil pengujian algoritma naïve bayes diketahui bahwa tingkat akurasi mencapai 100%. Penerapan metode MonkeyLearn dapat membantu mendapatkan data untuk diolah pada Rapidminer menggunakan algoritma naïve bayes.

Kata Kunci : Analisis Sentimen, MonkeyLearn, Naive Bayes, Rapidminer, Twitter

Abstract

Currently, social media is one of the most widely accessed online media throughout the world. With the existence of microblogging services such as Twitter, it allows users to write with the aim of being published, writing public opinion opinions, services, and so on. It all becomes an attraction for social media users to use twitter. This can be used to take opinions or sentiments against an entity that is being discussed on social media such as Twitter. This study utilizes this data to find out the opinions or sentiments of twitter users regarding a public figure who is currently trading, namely Joe Biden. Opinions will be taken based on three classes, namely positive, negative, and neutral. In this study, the author applies text-processing MonkeyLearn to overcome the problem of obtaining follow-up data from raw text. The classification method used is the nave Bayes classifier algorithm. From the results of the research that has been carried out, it can be seen that the percentage of sentiment on Twitter social media regarding news responses related to Joe Biden produces neutral sentiments that are more dominant than positive and negative sentiments. In addition, from the results of testing the nave Bayes algorithm, it is known that the accuracy rate reaches 100%. The application of the MonkeyLearn method can help get data to be processed in Rapidminer using the nave Bayes algorithm.

Keywords: Sentiment Analysis, MonkeyLearn, Naive Bayes, Rapidminer, Twitter

I. PENDAHULUAN

Perkembangan teknologi ke arah serba digital saat ini semakin pesat [1]. Manusia secara umum memiliki gaya hidup baru yang tidak bisa dilepaskan dari perangkat elektronik. Salah satunya penggunaan media sosial yang merupakan sebuah perangkat teknologi informasi yang memiliki fitur berupa interaksi. Adapun jenis fiturnya yaitu mikroblog yang memungkinkan setiap penggunanya dapat

mengekspresikan diri dan menyampaikan apa yang mereka pikirkan dan Twitter merupakan salah satu contohnya. Twitter merupakan media sosial yang di dalamnya terjadi suatu interaksi antar penggunanya untuk berbagi berbagai informasi yang di dalamnya terjadi suatu interaksi antar penggunaannya dengan menggunakan perangkat elektronik berupa handphone ataupun laptop.

Perkembangan media sosial telah mengubah gaya

hidup manusia saat ini. Dengan media sosial pengguna dari media sosial akan selalu update dan dapat berbagi informasi secara real-time dengan frekuensi tinggi. Maka media sosial saat ini telah menjadi media alternatif yang digunakan untuk melihat perkembangan yang sedang hangat diperbincangkan, serta menjadi wadah untuk berinteraksi dari pengguna satu dengan yang lainnya.[2]

Hal tersebut selaras dengan fenomena saat ini di mana berita mengenai Joe Biden menjadi trending topic setelah memenangkan pemilu menjadi Presiden Amerika. Ini menjadi berita yang sangat terkenal di berbagai negara karena di sisi lain dengan terpilihnya Joe Biden maka akan berdampak terhadap peta perekonomian dunia. Sehingga bermunculan pendapat positif maupun negatif dari pengguna twitter terhadap isu tersebut. Pendapat-pendapat tersebut bisa diolah dan dianalisa agar menghasilkan data atau informasi yang dapat membantu banyak pihak untuk mendukung suatu keputusan atau pilihan. Salah satu teknik pemrosesan teks yang tepat yaitu dengan analisa sentimen. Analisa sentimen adalah studi komputasi mengenai pendapat, perilaku, dan emosi seseorang terhadap entitas. Entitas tersebut bisa berupa gambaran individu, kejadian atau topik.[3]

Untuk memudahkan penulis dalam penelitian mengenai berita Joe Biden di twitter penulis menggunakan metode MonkeyLearn. Kemudian data hasil dari metode MonkeyLearn di uji pada Rapidminer menggunakan algoritma Naïve Bayes, yang bertujuan untuk mengcross check data yang dihasilkan metode MonkeyLearn khususnya pada sentimen positif, negatif, dan netral.

II. BAHAN DAN METODE/METODOLOGI

Penelitian diawali dengan pengumpulan data kemudian memilih metode, algoritma, dan tools yang akan digunakan.

A. Data Twitter

Data tweet yang digunakan pada penelitian ini yaitu tweet pada twitter terhadap berita mengenai Joe Biden yang berupa pendapat positif, negatif, dan netral dari pengguna social media twitter.

B. MonkeyLearn

MonkeyLearn adalah sebuah layanan teks klasifikasi online yang menyediakan pengguna untuk mengunggah kumpulan data mereka untuk di uji dalam sebuah algoritma. MonkeyLearn juga menyediakan sebuah pengaturan seperti jenis klasifikasi yang berbeda, kata henti, dan n-gram range [4].

C. Algoritma Naïve Bayes

Algoritma Naïve Bayes merupakan suatu algoritma yang dapat mengklasifikasikan suatu variabel tertentu dengan menggunakan metode probabilitas dan statistik. Secara garis besar algoritma Naïve Bayes dapat dijelaskan seperti ditampilkam pada persamaan 1.

$$P(R|S) = \frac{P(R)P(S|R)}{P(S)} \quad (1)$$

Keterangan :

R : Data yang belum diketahui kelasnya

S : Hipotesis pada data R yang merupakan class khusus

$P(R|S)$: Nilai probabilitas pada hipotesis R yang berdasarkan kondisi S

$P(R)$: Nilai probabilitas pada hipotesis R

$P(S|R)$: Nilai probabilitas S yang berdasarkan dengan kondisi hipotesis R

$P(S)$: Nilai Probabilitas S

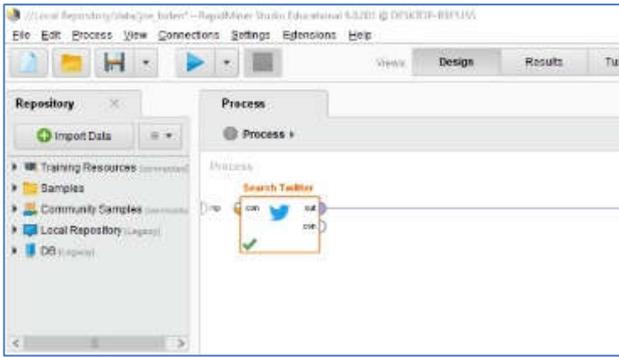
Dengan menggunakan persamaan diatas, data yang telah diperoleh dapat diproses dengan algoritma *Naïve Bayes* untuk penilaian data yang akan diklasifikasikan.[5]

D. Rapidminer

Rapid miner adalah sebuah software untuk pengolahan data. Dengan menggunakan prinsip dan algoritma data mining, Rapid miner mengekstrak pola-pola dari data set yang besar dengan mengkombinasikan metode statistika, kecerdasan buatan dan database. Rapid miner membuat mudah dalam penelitian dalam melakukan perhitungannya dalam melakukan perhitungan operator-operator. Operator ini berfungsi untuk memodifikasi data. Data dihubungkan dengan node-node operator kemudian penulis hanya tinggal menghubungkannya ke node hasil untuk melihat hasilnya. Hasil yang diperlihatkan Rapid miner pun dapat ditampilkan secara visual dengan grafik.[6]

III. HASIL DAN PEMBAHASAN

Langkah awal dalam penelitian ini penulis mengawali dengan mencari data source yang berasal dari Twitter dengan topik Joe Biden. Cara mendapatkan datanya yaitu dengan cara crawling data pada Rapidminer seperti ditampilkan pada gambar 1.



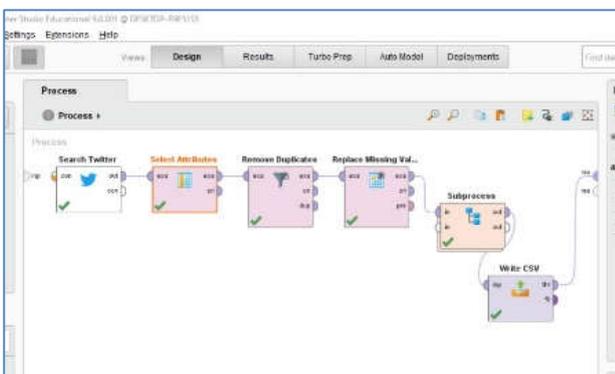
Gambar 1 Proses Crawling Data

Tahapan tweet crawling ini adalah melakukan pengambilan data tweet dari twitter.com dengan menggunakan API yang disediakan pada twitter.



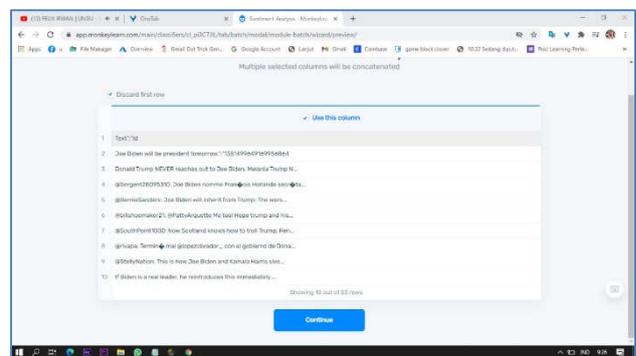
Gambar 2 Langkah-langkah tweet crawling

Pada tahap pengumpulan data dilakukan dengan cara scraping halaman pencarian twitter. Proses scraping dimulai dengan user memasukkan query pencarian twitter yang dapat berupa keyword, mention, hastag, tanggal, dan tahun[7].



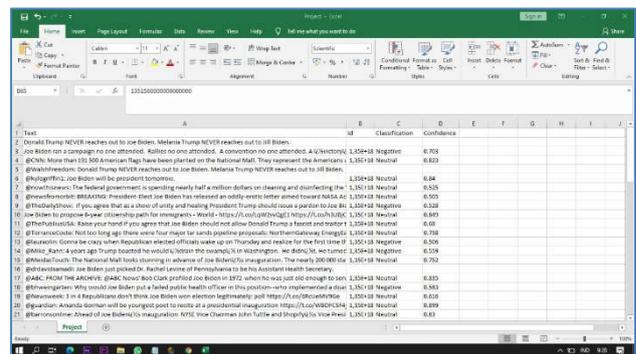
Gambar 3 Merapihkan Data Crawling

Langkah selanjutnya bisa dilihat pada Gambar 3. Fungsi dari langkah di atas yaitu untuk merapihkan data crawling dari twitter yang masih bertakan, operasi yang digunakan dalam langkah pada proses merapihkan data Crawling yaitu select attribute, remove duplicate, replace missing value, subproses, dan write CSV. Semua operator memiliki fungsi tersendiri untuk proses merapihkan data, sebagai contoh ada remove duplicate memiliki fungsi untuk menghilangkan data yang sama. Kemudian pada operator yang terakhir ada operator write CSV yang berfungsi untuk menyimpan data yang sudah di crawling ke dalam format CSV. Yang nantinya untuk di koneksikan pada MonkeyLearn



Gambar 4 MonkeyLearn

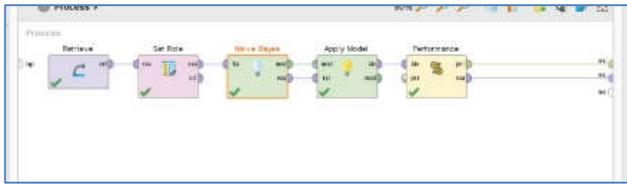
Gambar 4 adalah proses memasukan data CSV hasil dari proses gambar 3 untuk di proses pada metode MonkeyLearn. Pada gambar 4 proses yang terjadi yaitu data data akan diolah untuk menghasilkan pengelompokan data yang hasilnya memiliki respon positif, neutral, dan negatif.



Gambar 5 Data Hasil Metode MonkeyLearn

Data hasil dari proses menggunakan MonkeyLearn akan berupa CSV yang ada pada gambar 5. File CSV yang dihasilkan perlu ada perapihan sedikit agar bisa rapih seperti gambar 5 yang berguna untuk proses kepada tahap selanjutnya agar file CSV bisa terbaca. Dan pada file CSV hasil

dari proses menggunakan MonkeyLearn sudah ada pengelompokan data dalam bentuk respon positif, netral, dan negatif.



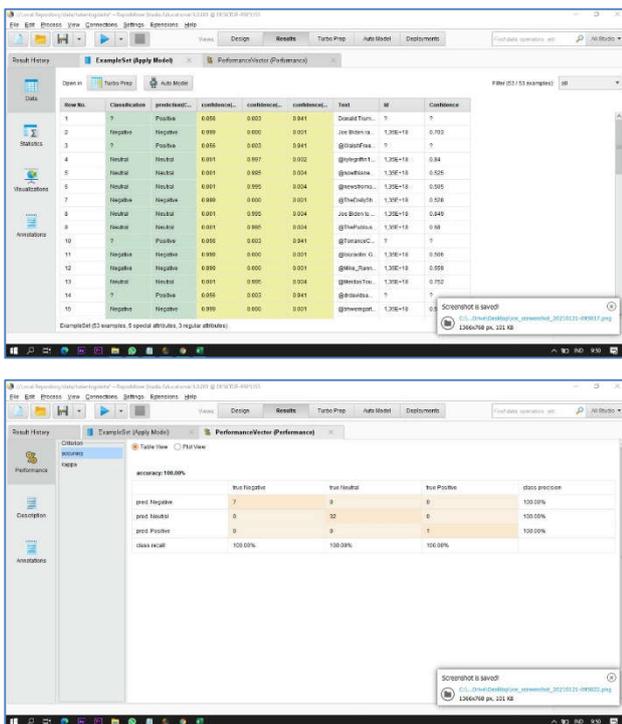
Gambar 6 Proses dengan Algoritma Naïve Bayes

Pada gambar 6 yaitu proses memasukan data CSV yang bisa dilihat pada gambar 5, untuk diproses pada Rapidminer menggunakan Algoritma Naïve Bayes. Tujuan dari penggunaan Algoritma Naïve Baye yaitu untuk mengcross check ulang data hasil dari metode MonkeyLearn terutama dalam hal pengelompokan respon positif, netral, dan negatif yang ada pada data CSV hasil dari metode MonkeyLearn.

IV. KESIMPULAN

Dalam analisis sentimen yang sudah penulis lakukan mengenai trading news Joe Biden dengan data yang di crawling dari Twitter menggunakan metode MonkeyLearn yang memiliki fungsi memberi kemudahan dalam mendapatkan data walaupun mentah agar bisa mudah di proses, dan Algoritma Naïve Bayes untuk mengcross check hasil dari metode MonkeyLearn. Bagi penulis metode MonkeyLearn dan Algoritma Naïve Bayes adalah sebuah kombinasi metode dan algoritma yang memudahkan dalam proses analisis sentimen. Kemudian dalam kategori hasil yang sudah di proses pada Algoritma Naïve Bayes sesuai ekspektasi, tanpa error atau trouble lainnya. Dimana hasil dari analisis sentimen yang penulis dapat dari trading news Joe Biden menghasilkan accuracy 100%.

DAFTAR PUSTAKA



Gambar 7 Hasil Proses Algoritma Naïve Bayes

Hasil dari pengolahan data yang di proses menggunakan algoritma Naïve Bayes ada pada gambar 7. Hasil dari data yang sudah di proses oleh algoritma Naïve Bayes yang ada pada gambar 7, pada respon positif, netral, dan negatif menghasilkan accuracy 100%.