

PEMANFAATAN TEKNOLOGI UNTUK MENGANALISA SENTIMEN MASYARAKATDALAM MEMBANTU PENINGKATAN EKONOMI KREATIF DI ERA NEW NORMAL

Paisal^{1*}, Therezia Naraloka²

Sistem Informasi, Fakultas Ilmu Komputer, Universitas Sjakhyakirti^{1,2}
paisal@unisti.ac.id^{1*}, therezia.naraloka@unisti.ac.id²

ABSTRACT

The use of social media is not only for communicating between friends, but is often used as a means to convey an aspiration for the community, especially the Indonesian people, regarding legal issues or problems related to the creative economy. One of the aspirations conveyed through this social media is a hashtag that is widely seen every day with a trending level, this statement comes from many sentiments from every community, some provide positive sentiments and also negative sentiments regarding responses to these hashtags which can have a good impact. or bad for everyday life in society. From this problem, the author will explain to produce a solution where the hashtag can have a good impact on society or vice versa can have an impact on the progress of the creative economy business in the new normal era. In analyzing this, the researcher uses the Naïve Bayes Classifier method which is a machine learning method that uses probability calculations, this automatic hashtag classification can be adjusted so as to minimize misclassification actions personally by obtaining positive or negative sentiment information so as to produce information that can be useful for the wider community, especially in Indonesia.

Keywords: *Social Media, Creative Economy, Hastag, Naïve Bayes Classifie*

PENDAHULUAN

Media sosial saat ini telah menjadi trend dalam komunikasi pemasaran. Media sosial adalah sebuah media online, media sosial bermanfaat bagi para penggunanya dengan berbagai cara meliputi mengenai blog, jejaring sosial, wiki, forum dan dunia virtual. Media sosial adalah sebuah kelompok aplikasi yang berbasis internet yang dibangun berbasis teknologi web yang memungkinkan pertukaran user-generated content adapun beberapa contoh media sosial yang sedang berkembang saat ini yaitu Twitter (Kurniawan, 2018)

Pemanfaatan sosial media ini tidak hanya untuk berkomunikasi antara teman saja, akan tetapi sering juga dijadikan sebuah sarana untuk menyampaikan suatu aspirasi bagi masyarakat mengenai masalah hukum ataupun masalah yang berhubungan dengan pemerintahan. Salah satu aspirasi yang disampaikan melalui sosial media ini adalah dalam bentuk sebuah hastag. Adapun bentuk hastag yang digunakan seperti hastag mengenai #newnormal dimana hastag ini merupakan sebuah penyampaian aspirasi bagi masyarakat mengenai masalah pandemi corona, dari statement hastag ini kemudian muncullah banyak sentimen dari masyarakat, dari yang memberikan sentimen positif dan juga sentimen negatif mengenai tanggapan terhadap hastag tersebut sehingga dapat berdampak baik atau buruk bagi kehidupan sehari-hari dimasyarakat seperti adanya

sentiman dan sikap yang egois dari antara setiap orang dengan orang lain yang dianggapnya tidak sejalan.

Dari permasalahan ini peneliti akan menganalisa untuk menghasilkan sebuah solusi dimana hastag tersebut dapat memberikan informasi terbaik bagi peningkatan ekonomi kreatif di era new normal. Dalam menganalisa ini akan menggunakan metode *Naïve Bayes Classifier* yang merupakan salah satu metode machine learning yang menggunakan perhitungan probabilitas. Konsep dasar yang digunakan oleh *Naïve bayes* adalah Teorema *Bayes Classifier*, yaitu teorema dalam statistika untuk menghitung peluang. *Bayes Optimal Classifier* menghitung peluang dari satu kelas dari masing-masing kelompok atribut yang ada, dan menentukan kelas mana yang paling optimal, adapun kelebihan dari penggunaan *Naïve Bayes classifier* dalam klasifikasi dokumen dapat ditinjau dari prosesnya yang mengambil aksi berdasarkan data-data yang telah ada sebelumnya. Oleh karena itu, klasifikasi dokumen dengan metode ini dapat dipersonalisasi, maksudnya adalah proses klasifikasi dokumen dapat disesuaikan dengan sifat dan kebutuhan, informasi ini dapat dilihat secara nyata melalui *hashtag filtering*, dengan menyetikkan nama filter hastagnya seperti #newnormal. Proses klasifikasi ini menggunakan *Naïve Bayes classifier*, pengklasifikasian hastag otomatis ini dapat disesuaikan sehingga meminimalisasi aksi salah pengklasifikasian secara personal dengan memperoleh informasi sentimen positif atau negatif, peneliti memilih *Naïve Bayes classifier* dikarenakan dapat menangani kuantitatif dan data diskrit, hanya memerlukan sejumlah kecil data pelatihan untuk mengestimasi parameter (rata-rata dan variansi dari variabel) yang dibutuhkan untuk klasifikasi serta dapat menangani nilai yang hilang dengan mengabaikan instansi selama perhitungan estimasi peluang cepat dan efisiensi ruang.

TINJAUAN PUSTAKA DAN PENGEMBANGAN HIPOTESIS

Berdasarkan penelitian terdahulu dari Hayuningtyas (2019), Analisis sentimen merupakan analisis teks untuk mengidentifikasi suatu opini. Sedangkan menurut penelitian terdahulu oleh Murnawan (2017), menyatakan bahwa Sentiment analysis atau sering disebut juga opinion mining adalah studi komputasional dari opini-opini orang, appraisal dan emosi melalui entitas. Sentiment analysis atau opinion mining merupakan topik riset yang penting dan sedang marak dilakukan saat ini Opinion mining merupakan cabang penelitian dari text mining (Rozi, 2012). Salah satu cara untuk mengetahui keberhasilan sesuatu adalah dengan melakukan analisis sentimen terhadap aplikasi tersebut (Giovani, 2020).

Ide dasar dari investigasi sentimen adalah untuk mendeteksi polaritas teks dokumen atau kalimat pendek dan mengklasifikasikannya dalam premis (Arief, 2019). *Sentiment analysis* atau *opinion mining* mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi linguistik dan text mining yang bertujuan menganalisa pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenaan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu (Amrullah, 2020).

Analisis dilakukan dengan melakukan klasifikasi *tweet* yang berisi sentimen masyarakat tentang tokoh tertentu. Metode klasifikasi yang digunakan dalam penelitian ini adalah *Naive Bayes Classifier*. *Naive Bayes Classifier* dikombinasikan dengan fitur untuk dapat mendeteksi negasi dan pembobotan menggunakan *term frequency* serta *TF-IDF* (Hidayatullah, 2014). Model klasifikasi digunakan untuk data testing melakukan proses klasifikasi yang menghasilkan label sentimen (positif/negatif), proses ini dinamakan dengan proses testing (Afrizal, 2019), Apabila ulasan/ulasan ini dikumpulkan

lalu diolah maka hasil tersebut akan dapat dijadikan sebuah yang memiliki sentimen terbaik. (Fitri, 2020).

Bentuk analisis pendapat atau review dapat dilakukan menggunakan analisis sentimen, di mana analisis sentimen merupakan proses yang dilakukan memahami, mengekstrak, dan mengolah data tekstual secara otomatis untuk mendapatkan informasi (Ikasari, 2020). Analisis sentimen adalah sebuah cara yang digunakan untuk mengolah komentar yang diberikan oleh pemesan atau pelanggan melalui berbagai media, mengenai sebuah produk, jasa ataupun sebuah instansi (Sipayung, 2016). Analisis sentimen digunakan untuk mengekstrak data opini, memahami serta mengolah tekstual data secara otomatis untuk melihat sentimen yang terkandung dalam sebuah opini (Sari, 2019).

Teknik pengumpulan kata yang memiliki sentimen dilakukan dengan pendekatan berdasarkan kamus (Putranti, 2014). Text processing merupakan tahap terpenting dalam analisis sentimen, proses ini menentukan kualitas data yang menjadi input komputasi algoritma, output dari masing-masing proses text processing berpengaruh pada proses selanjutnya. Adapun output text processing disajikan dalam tabel-tabel dibawah ini (Arsi, 2019). Pelabelan data dengan sentiment scoring dan manual: melakukan pelabelan data menjadi label positif atau negatif dengan menggunakan sentiment scoring dan secara manual (Fitriyah, 2020).

a) Analisis

Menurut Spradley (dalam Sugiyono, 2014) mengatakan bahwa analisis adalah sebuah kegiatan untuk mencari suatu pola selain itu analisis merupakan cara berpikir yang berkaitan dengan pengujian secara sistematis terhadap sesuatu untuk menentukan bagian, hubungan antar bagian dan hubungannya dengan keseluruhan. Analisis adalah suatu usaha untuk mengurai suatu masalah atau fokus kajian menjadi bagian-bagian (decomposition) sehingga susunan/tatanan bentuk sesuatu yang diurai itu tampak dengan jelas dan karenanya bisa secara lebih terang ditangkap maknanya atau lebih jernih dimengerti duduk perkaranya

b) Media sosial

Media sosial selalu berubah dari waktu ke waktu sehingga memungkinkan seseorang dapat menjadi kedua pihak, yaitu penerima informasi dan juga penyebar informasi. Secara garis besar media sosial bisa dikatakan sebagai sebuah media online, di mana para penggunanya (user) melalui aplikasi berbasis internet dapat berbagi, berpartisipasi, dan menciptakan konten berupa blog, wiki, forum, jejaring sosial, dan ruang dunia virtual yang disokong oleh teknologi multimedia menjadi satu kesatuan yang sulit dipisahkan serta mendorong pada hal-hal baru. Saat ini media sosial yang paling banyak digunakan dan tumbuh pesat berupa jejaring sosial, blog, dan wiki. Merebaknya situs media sosial yang muncul menguntungkan banyak orang dari berbagai belahan dunia untuk berinteraksi dengan mudah dan dengan biaya yang murah ketimbang memakai telepon. Dampak positif yang lain dari adanya situs jejaring sosial adalah percepatan penyebaran informasi. Akan tetapi ada pula dampak negatif dari media sosial, yakni berkurangnya interaksi interpersonal secara langsung atau tatap muka, munculnya kecanduan yang melebihi dosis, serta persoalan etika dan hukum karena kontennya yang melanggar moral, privasi serta aturan. (Evans, 2008),

Sebagaimana yang sudah dijelaskan di atas tentang pengertian media sosial, yakni sebuah media online artinya pengguna media online tersebut dapat berkomunikasi tanpa harus tatap muka. Awal munculnya media sosial pada abad 70-an, tepatnya tahun 1978,

saat sistem papan buletin atau bulletin board system (BBS) ditemukan oleh Ward Christensen dan Randy Suess yang merupakan pecinta dunia komputer. Sistem papan ini memungkinkan pengguna (user) untuk bisa berhubungan dengan orang lain memakai surat elektronik atau mengunggah dan mengunduh melalui perangkat lunak yang tersedia saat itu. Ini merupakan awal sebuah komunitas virtual dalam lingkup terbatas.

Dengan seiring waktu berjalan munculah berbagai media sosial dengan seabrek keunggulan, keunikan, karakteristik, dan segmentasi yang beragam. Kemunculan sejumlah situs jejaring sosial itu pada intinya bermula dari adanya inisiatif untuk menghubungkan orang-orang dari berbagai latar belakang yang ada di seluruh belahan bumi. Dalam perkembangan lebih lanjut, media sosial kini menjadi sarana aktifitas yang masuk kategori digital marketing, karena banyak dijejali kemasan dan muatan pemasaran. Salah satu mendasar yang ada pada situs-situs media sosial tersebut adalah fungsi dan layanan jejaring sosial. Layanan jejaring sosial memberikan jasa koneksi aktifitas melalui situs, platform dan dan saran yang berfungsi memfasilitasi pembentukan jaringan atau hubungan sosial di antara beragam orang yang mempunyai ketertarikan, minat (interest), kegiatan, latar belakang, maksud, kepentingan, tujuan, atau korelasi dunia nyata yang sama. Dari uraian sejarah media sosial diatas telah memunculkan berbagai karakteristik jenis media sosial. Maka dari itu pada era ini banyak macam-macam bentuk media sosial yang tercipta dan memiliki perbedaan dan karakteristik sendiri-sendiri. Menurut penelitian Overdrive, suatu lembaga riset pemasaran, jenis aplikasi media sosial saat ini sedikitnya telah mencapai 240 aplikasi yang menawarkan ratusan cara berinteraksi. Dari penelitian tersebut memperlihatkan keunikan interaksi aplikasi media sosial sangat beragam. Dalam bentuk yang simpel, media sosial adalah bentuk lintas komunikasi lewat penyebaran konten. Materi atau konten itu dapat berupa suatu percakapan diskusi, video, opini, foto dan sebagainya. Dalam hal inilah, lembaga pemerintahan harus turut berperan, bergabung dan secara maksimal dapat memanfaatkan fasilitas ini.

c) *Twitter*

Twitter adalah situs *microblogging* yang dioperasikan oleh Twitter, Inc. Disebut micro blogging karena situs ini memungkinkan penggunanya mengirim dan membaca pesan seperti blog pada umumnya. Pesan tersebut dinamakan tweets, yaitu teks tulisan sebanyak 140 karakter yang ditampilkan pada halaman profil pengguna. Perkembangan dunia teknologi semakin memudahkan para pengguna Twitter yang kian hari kian bertambah. Melalui aplikasi eksternal yang kompatibel seperti telepon seluler atau dengan SMS, penggunanya dapat mengirim dan menerima tweets langsung dari situs Twitter. Dengan kemudahan seperti ini, Twitter telah mendapatkan tempat di seluruh dunia dengan semakin bertambahnya jumlah users. *Twitter* dibentuk dan diluncurkan pada 2006 oleh Jack Dorsey. Pendirian Twitter, Inc. dibiayai oleh Obvious LLC. Markasnya terletak di San Bruno, California. Kemudian, server dan kantor Twitter yang lain tersebar di San Antonio, Texas dan Boston, Massachusetts. Twitter menjadi populer pada tahun 2007 pada festival South by Southwest. Selama acara tersebut berlangsung, pengguna Twitter meningkat dari 20.000 tweets per hari menjadi 60.000 tweets. Pada September 2010, Twitter mengganti logo dan membuat desain.

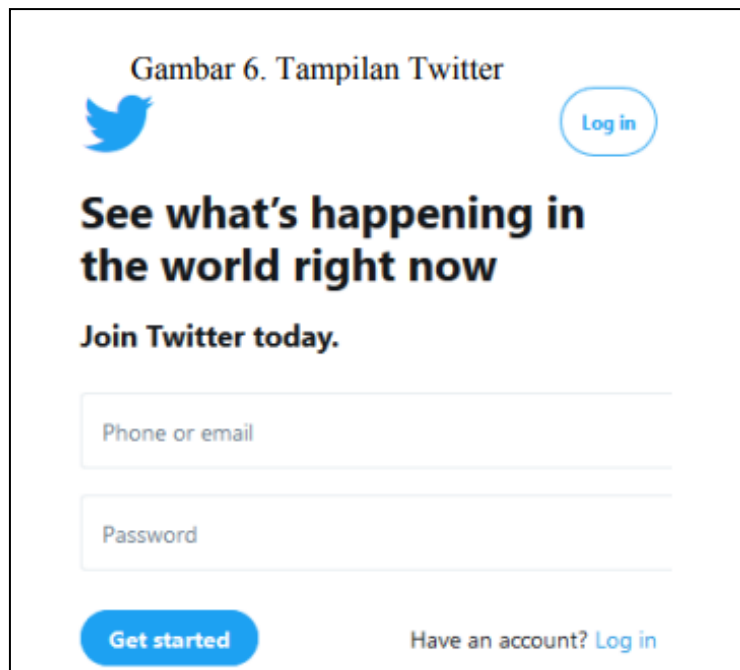


Figure 1: Tampilan Twitter

Dalam menggunakan *Twitter*, terdapat konten-konten pendukung diantaranya :

- 1) *Home*
Home adalah tampilan *Twitter* setelah melakukan log in. Tampilan ini berisi tweet dari akun-akun yang sudah *follow*. *Home* juga biasa disebut dengan timeline. Tweet yang kita tulis juga akan muncul pada timeline orang yang menjadi *follower*.
- 2) *Profile*
Profile adalah tampilan yang terdapat foto, biodata, dan juga aktifitas pengguna akun mulai dari kegiatan *tweet*, *retweet*, dsb.
- 3) *Follower*
Follower adalah pengguna *Twitter* lain yang mengikuti kita. Dengan menjadi *Follower*, pengguna tersebut akan mendapatkan update *tweet* dari kita.
- 4) *Following*
Following adalah pengguna *Twitter* lain yang kita ikuti. Dengan menjadi following, kita akan menerima update *tweet* dari akun yang diikuti.
- 5) *Mention*
Mention adalah balasan dari percakapan pengguna *Twitter* dengan menandai akun *Twitter* yang menjadi lawan bicaranya. Menandai dilakukan dengan menulis @ dan diikuti nama akun lawan bicara. Mention bersifat publik, sehingga dapat dilihat oleh pengguna lainnya.
- 6) *Favorite*
Favorite adalah *tweet* yang diberi tanda bintang. Dengan demikian, *Tweet* tersebut tidak akan hilang dari halaman *Twitter* kita.
- 7) *Replay*
Replay adalah konten yang memungkinkan kita untuk membalas pesan/ *tweet* yang ditujukan kepada kita.
- 8) *Retweet (RT)*

Retweet atau biasa disingkat RT. *Retweet* adalah konten yang memungkinkan kita untuk membagikan tweet orang lain dengan maksud sepaham dengan *tweet* yang ditulis. *Retweet* tidak menghilangkan akun penulis *tweet*.

9) *Direct Message* (DM)

Direct Message atau biasa disingkat DM adalah pesan yang dikirim oleh pengguna *Twitter* secara rahasia. Pesan ini hanya dapat dilihat oleh pengguna *Twitter* yang dituju.

10) *Hashtag* (#)

Hashtag atau tanda pagar (#) adalah simbol yang digunakan untuk memudahkan pencarian/ pengelompokan suatu topik. Semakin banyak pengguna *hashtag*, maka topik tersebut akan menjadi *trending topic*.

11) *List*

List adalah konten yang memungkinkan pengguna *Twitter* mengelompokkan *following*nya sesuai dengan kebutuhan.

12) *Trending Topics*

Trending topics adalah topik yang sedang menjadi bahan pembicaraan banyak orang. Biasanya diikuti dengan penggunaan hashtag/ tanda pagar.

d) Metode Naive Bayes Classifier

Menurut Bustami (2013), *Naive Bayes* merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas. Definisi lain mengatakan *Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. Keuntungan penggunaan *Naive Bayes* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. *Naive Bayes* sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan. Dapat dilihat pada rumus dibawah ini.

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

Sumber : (Bustami, 2013)

Di mana :

X : Data dengan class yang belum diketahui

H : Hipotesis data merupakan suatu class spesifik

P(H|X) : Probabilitas hipotesis H berdasar kondisi X (posteriori probabilitas)

P(H) : Probabilitas hipotesis H (prior probabilitas)

P(X|H) : Probabilitas X berdasarkan kondisi pada hipotesis H

P(X) : Probabilitas X

Untuk menjelaskan metode *Naive Bayes*, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel

yang dianalisis tersebut. Karena itu, metode Naive Bayes di atas disesuaikan. Dapat dilihat pada rumus dibawah ini.

$$P(C|F1 \dots Fn) = \frac{P(C)P(F1 \dots Fn|C)}{P(F1 \dots Fn)}$$

Sumber : (Bustami, 2013)

e) Data mining

Definisi umum dari data mining itu sendiri adalah proses pencarian pola-pola yang tersembunyi (hidden patern) berupa pengetahuan (knowledge) yang tidak diketahui sebelumnya dari suatu sekumpulan data yang mana data tersebut dapat berada di dalam database, data werehouse, atau media penyimpanan informasi yang lain. Hal penting yang terkait di dalam data mining adalah:

1. Data mining merupakan suatu proses otomatis terhadap data yang sudah ada.
2. Data yang akan diproses berupa data yang sangat besar.
3. Tujuan data mining adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat. (Kusrini dan Emha Taufiq, 2009)

Data mining dilakukan dengan tool khusus, yang mengeksekusi operasi data mining yang telah didefinisikan berdasarkan model analisis. Data mining merupakan proses analisis terhadap data dengan penekanan menemukan informasi yang tersembunyi pada sejumlah data besar yang disimpan ketika menjalankan bisnis perusahaan. Kemajuan luar biasa yang terus berlanjut dalam bidang data mining didorong oleh beberapa faktor antara lain :

1. Pertumbuhan yang cepat dalam kumpulan data.
2. Penyimpanan data dalam data warehouse, sehingga seluruh perusahaan memiliki akses ke dalam database yang andal.
3. Adanya peningkatan akses data melalui navigasi web dan internet.
4. Tekanan kompetisi bisnis untuk meningkatkan penguasaan pasar dalam globalisasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk data mining (ketersediaan teknologi).
6. Perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan. (Larose, 2005)

Adapun tahapan tersebut bersifat interaktif dimana pemakai terlibat langsung atau dengan perantara *knowledge base*:

1. Pembersihan data

Pembersihan data merupakan proses menghilangkan noise dan data yang tidak konsisten atau data tidak relevan. Pada umumnya data yang diperoleh, baik dari database suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa data mining yang dimiliki. Data-data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga akan mempengaruhi performansi dari teknik data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

2. Integrasi data

Integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru. Tidak jarang data yang diperlukan untuk data mining tidak hanya berasal dari satu database tetapi juga berasal dari beberapa database atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

3. Seleksi data

Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli dalam kasus market basket analysis, tidak perlu mengambil nama pelanggan, cukup dengan id pelanggan saja.

4. Transformasi data

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining. Beberapa metode data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan clustering hanya bisa menerima input data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagi-bagi menjadi beberapa interval. Proses ini sering disebut transformasi data.

5. Proses mining

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

6. Presentasi pengetahuan

Untuk mengidentifikasi pola-pola menarik kedalam knowledge based yang ditemukan. Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses data mining, mencoba metode data mining lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu :

1. Deskripsi

Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpul suara mungkin tidak menemukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun dengan record lengkap menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

5. Pengklusteran

Pengklusteran merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan record yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan record-record dalam kluster lain. Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan record dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan record dalam kelompok lain akan bernilai minimal.

6. Asosiasi

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja (market basket analysis). (Larose,2005).

Arsitektur utama dari sistem data mining, pada umumnya terdiri dari beberapa komponen sebagai berikut :

1. Database, data warehouse, atau media penyimpanan informasi, terdiri dari satu atau beberapa database, data warehouse, atau data dalam bentuk lain. Pembersihan data dan integrasi data dilakukan terhadap data tersebut.
2. Database, data warehouse, bertanggung jawab terhadap pencarian data yang relevan sesuai dengan yang diinginkan pengguna atau user.
3. Basis pengetahuan (Knowledge Base), merupakan basis pengetahuan yang digunakan sebagai panduan dalam pencarian pola.
4. Data mining engine, merupakan bagian penting dari sistem dan idealnya terdiri dari kumpulan modul-modul fungsi yang digunakan dalam proses karakteristik (characterization), klasifikasi (classification), dan analisis kluster (cluster analysis). Dan merupakan bagian dari software yang menjalankan program berdasarkan algoritma yang ada.

5. Evaluasi pola (pattern evaluation), komponen ini pada umumnya berinteraksi dengan modul-modul data mining. Dan bagian dari software yang berfungsi untuk menemukan pattern atau pola-pola yang terdapat dalam database yang diolah sehingga nantinya proses data mining dapat menemukan knowledge yang sesuai.
6. Antar muka (Graphical user interface), merupakan modul komunikasi antara pengguna atau user dengan sistem yang memungkinkan pengguna berinteraksi dengan sistem untuk menentukan proses data mining itu sendiri.

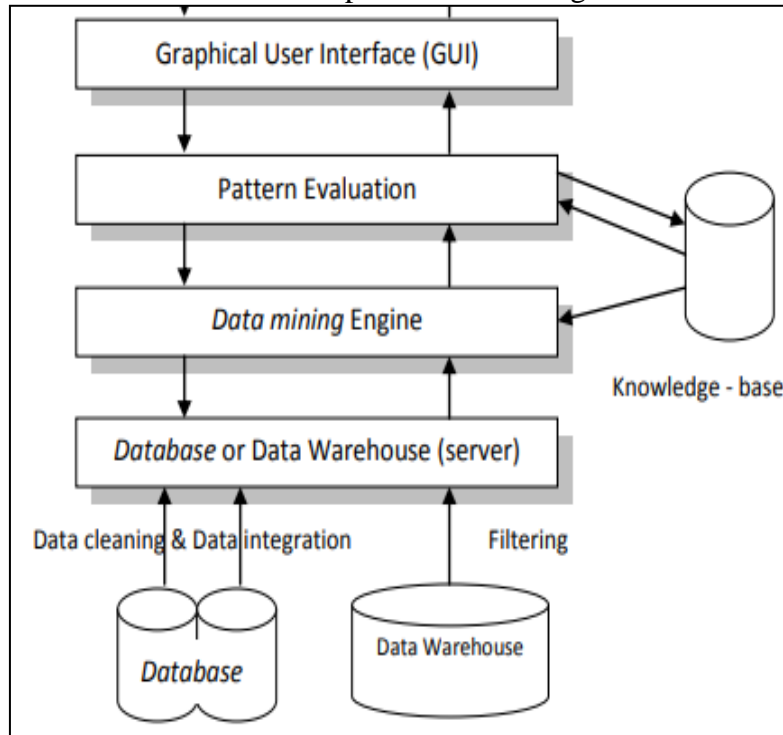


Figure 2: Arsitektur Data mining

Mengacu kepada enam fungsi data mining, yaitu :

- 1) Fungsi deskripsi (description)
- 2) Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.
- 3) Fungsi estimasi (estimation)
- 4) Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan baris data (record) lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.
- 5) Fungsi prediksi (prediction)
- 6) Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.
- 7) Fungsi klasifikasi (classification)

- 8) Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.
- 9) Fungsi pengelompokan (clustering)
- 10) *Clustering* merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas obyek-obyek yang memiliki kemiripan. Cluster adalah kumpulan record yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan record dalam cluster yang lain.
- 11) Fungsi asosiasi (*association*)
- 12) Tugas asosiasi dalam data mining adalah untuk menemukan atribut yang muncul dalam satu waktu. Data Mining sering juga disebut KDD, walaupun beberapa peneliti menganggap bahwa *Data Mining* adalah salah satu proses dari KDD.

Text mining merupakan bidang baru yang berkembang untuk mengumpulkan informasi dari Bahasa alami. Text mining ini adalah proses menganalisis text untuk mengekstrak informasi yang digunakan untuk tujuan tertentu. Text mining digunakan untuk mencari pola-pola yang ada di teks dalam bahasa natural yang tidak terstruktur seperti buku, e-mail, artikel, halaman web, dll. Berdasarkan ketidakteraturan struktur data teks, maka proses text mining memerlukan beberapa tahap awal yang pada intinya adalah mempersiapkan agar teks dapat diubah menjadi lebih terstruktur. Area penerapan text mining yang paling populer adalah :

- 1) Ekstraksi Informasi (Information Extraction): Identifikasi frase kunci dan keterkaitan di dalam teks dengan melihat urutan tertentu melalui pencocokan pola.
- 2) Pelacakan topik (Topic tracking): Penentuan dokumen lain yang menarik seorang pengguna berdasarkan profil dan dokumen yang dilihat pengguna tersebut.
- 3) Perangkuman (summarization): Pembuatan rangkuman dokumen untuk mengefisienkan proses membaca.
- 4) Kategorisasi (Categorization): Penentuan tema utama suatu teks dan pengelompokan teks berdasarkan tema tersebut ke dalam kategori yang telah ditentukan.
- 5) Penggugusan (Clustering): Pengelompokan dokumen serupa tanpa penentuan kategori sebelumnya (berbeda dengan kategorisasi di atas).
- 6) Penautan konsep (Concept Linking): Penautan dokumen terkait dengan identifikasi konsep yang dimiliki bersama sehingga membantu pengguna untuk menemukan informasi yang mungkin tidak akan ditemukan dengan hanya menggunakan metode pencarian tradisional.
- 7) Penjawaban pertanyaan (Question answering): Pemberian jawaban terbaik terhadap suatu pertanyaan dengan pencocokan pola berdasarkan pengetahuan.

f) *Weka*

Widiarsono (2005), Weka merupakan suatu program komputer (software) yang dapat membantumemecahkan berbagai masalah matematis yang terdapat dalam kehidupan sehari-hari lebih tepatnya dalam bidang teknis. Kita bisa memanfaatkan kemampuan Weka untuk menemukan solusi dari berbagai masalah numeric secara cepat dan tepat, mulai dari hal yang paling dasar, misalkan sistem 2 persamaan dengan 2 variabel: $x - 2y = 3212x + 5y = 12$ hingga yang kompleks, seperti mencari akar-akar

polinomial, interpolasi darisejumlah data, perhitungan dengan matriks, pengolahan sinyal, dan metodenumerik.

Weka merupakan kepanjangan dari Matrix Laboratory. Sesuai dengannamanya, struktur data yang terdapat dalam Weka menggunakan matriks atauarray berdimensi dua (double). Oleh karenanya penguasaan teori matriks mutlakdiperlukan bagi pengguna pemula Weka agar mudah dalam mempelajari danmemahami operasi-operasi yang ada di Weka.

METODE PENELITIAN

Metode penelitian merupakan suatu cara yang dapat digunakan untuk mencapai tujuan yang diharapkan melalui suatu penelitian dengan teknik-teknik dan alat-alat tertentu. Adapun metode yang digunakan dalam penelitian ini yaitu metode development research suatu kegiatan penelitian yang bertujuan dan berusaha mengembangkan atau melengkapi pengetahuan yang sudah ada atau diketahui. Permasalahan manusia dan lingkungan alamnya selalu berkembang yang kesemuanya ini harus memperoleh jawaban yang simbang. (Ati, 2012).

Dalam pengambilan data sosial media dengan menggunakan teknik *crawling* yang memanfaatkan sosial media *Twitter* API. Sehingga dengan memanfaatkan *Twitter* API tersebut, bisa dengan mudah memperoleh data-data sentiment masyarakat. Untuk selanjutnya dikumpulkan dan disimpan dalam sebuah file atau basis data Penelitian yang dilakukan untuk melakukan implementasi teknik.

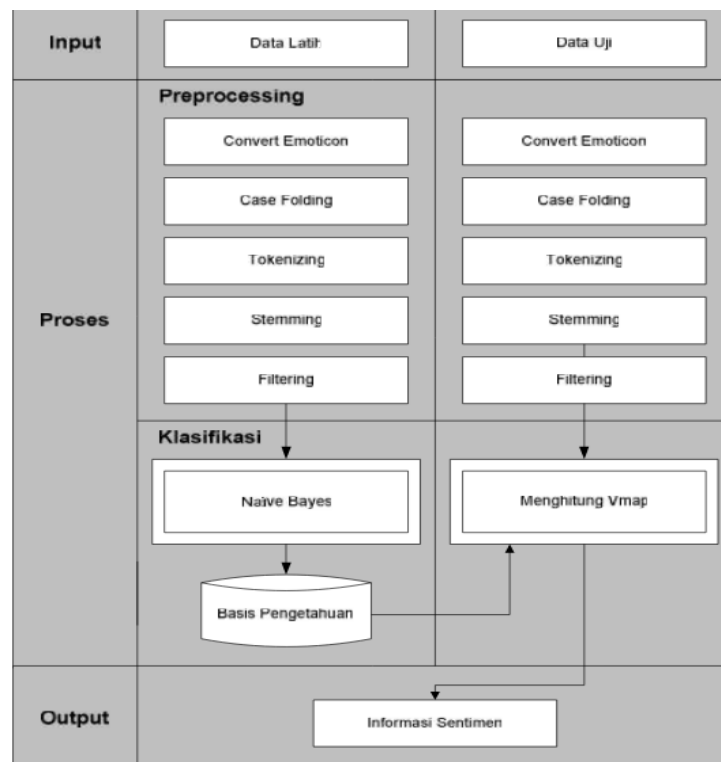


Figure 3: Alurkerja Sistem

Adapun penjelasan alurkerja dalam pengambilan data tersebut yaitu :

1. Input Data dari proses pengambilan data melalui di Twitter api yang dihubungkan dengan aplikasi TAG, kemudian data akan diolah.
2. Dilanjutkan dengan melakukan proses pengolahan data hastag
 - a. Proses *Precocessing*, yaitu proses convert emoticon dimana digunakan untuk mengkonversi lambang emosi, case folding tokenzing yaitu mengecek kata pada kamus, tokenizing yaitu pemotongan string setiap kata, stemming yaitu pemetaan dan penguraian bentuk dari suatu kata menjadi bentuk kata dasarnya dan filtering yaitu suatu proses dimana diambil sebagian kata tertentu, dan membuang kata yang tidak sesuai.
 - b. Proses Klasifikasi dengan algoritma *Naive Bayes* dan menghitung vmap adalah proses menghitung jumlah sentimen positif, negatif dan netral.
3. Menghasilkan Ouput berupa informasi persentase sentimen masyarakat baik sentimen positif, negatif dan hastag netral.

Tahap yang digunakan yaitu menyimpan data yang sudah ditemukan kedalam file atau database. Berikut ini adalah penjelasannya:

- 1 Memperoleh kunci *twitter api* Hal pertama yang harus dilakukan oleh pengguna adalah membuat akun *twitter*. Tujuannya untuk memperoleh *api key*, *api secret*, *access token* dan *access token secret*.
- 2 Menggunakan library Twitter Pada aplikasi yang dikembangkan, dilakukan *import library twitter* pada aplikasi TAGS V.6.1
- 3 Menghubungkan dengan *twitter searching* API Setelah dilakukan *import*, selanjutnya adalah menghubungkan apkikasi yang sedang dikembangkan dengan twitter, caranya adalah memasukkan *api key*, *api secret access token* dan *access token secret* pada bagian program, dari aplikasi yang sedang dikembangkan
- 4 Menyimpan data kedalam file atau database Setelah aplikasi dijalankan, maka hasil dari crawling disimpan dalam file, berupa json, xml, csv, xls atau ke dalam database.

HASIL DAN PEMBAHASAN

Adapun proses pengambilan data sebagai dalam melakukan analisa terhadap perancangan data mining untuk pemanfaatan teknologi untuk menganalisa sentimen masyarakat dalam membantu peningkatan ekonomi kreatif di era new normal.

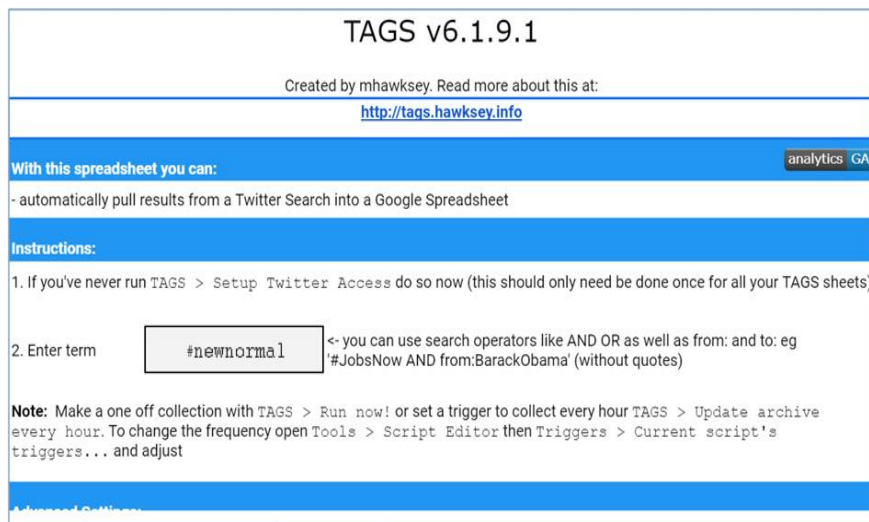


Figure 4: Aplikasi Tags V6.1.9.

Setelah pencarian informasi Ekstraksi data akan dilakukan dengan menggunakan aplikasi weka. Aplikasi ini tidak memiliki koneksi ke sebuah database karena dataset yang akan dimasukan dianggap sudah bersifat statis atau tidak ada perubahan nilai atau atribut. *Dataset* yang akan diolah diharuskan bertipe CSV (*comma separated values*). Hasil dari datamining ini adalah informasi yang nantinya dijalankan dengan menggunakan weka. Adapun informasi ini memiliki sub-sub menu sebagai berikut :

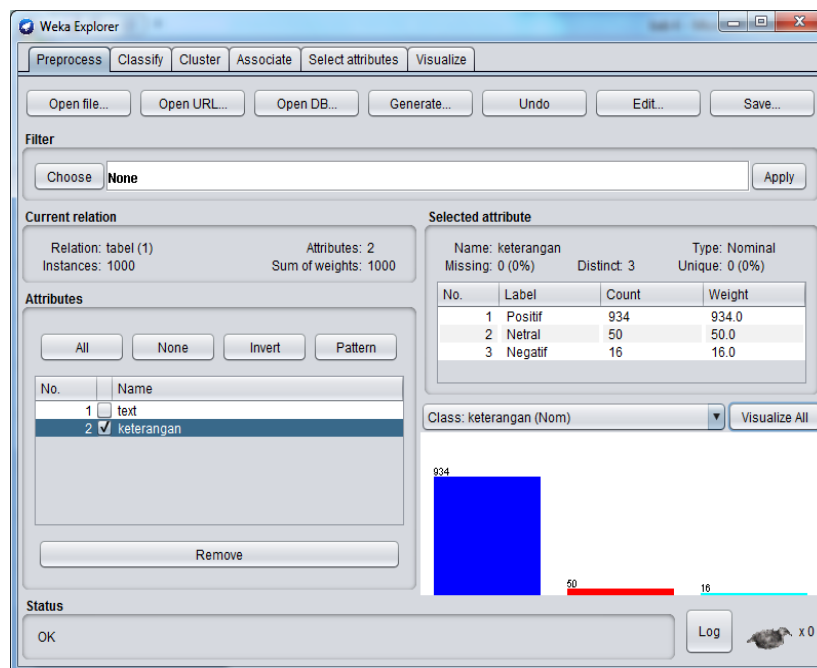


Figure 5: Proses Explorer

Tahap berikutnya adalah proses classify data yang akan dilakukan dengan menggunakan aplikasi weka. Proses ini digunakan untuk membuat class mengenai data Analisis Sentimen Masyarakat Terhadap Hastag Berdasarkan Opini Dari Twitter Menggunakan Metode *Naive Bayes Classifier* sehingga diperoleh nilai PRC Area positif 0.932, 0.054 Netral dan 0.014 Negatif.

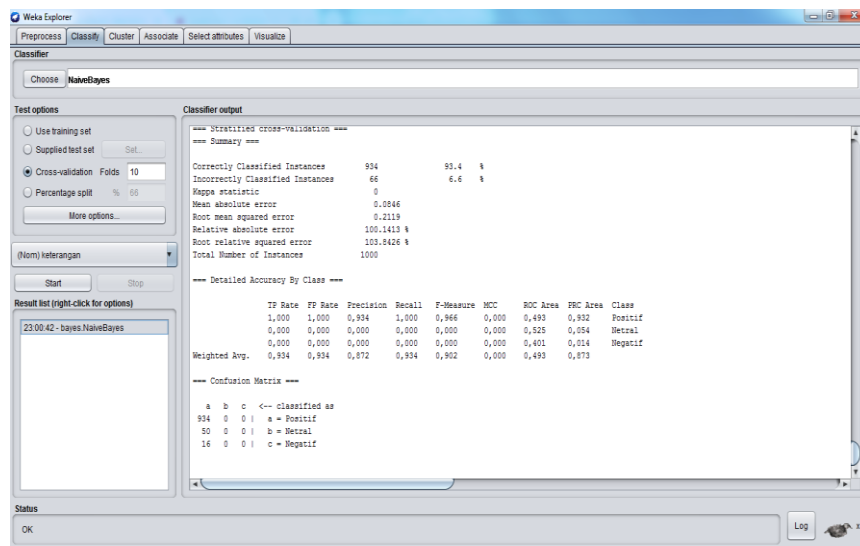


Figure 6: Proses Explorer

KESIMPULAN

Berdasarkan hasil penelitian serta pembahasan yang dilakukan maka dapat diambil kesimpulan bahwa :

1. Sesuai dengan tujuan dari pemanfaatan teknologi untuk menganalisa sentimen masyarakat dalam membantu peningkatan ekonomi kreatif di era new normal ini menggunakan Metode *Naive Bayes Classifier* diharapkan dapat memberikan informasi yang sangat berpengaruh berpengaruh besar terhadap masyarakat dalam peningkatan ekonomi kreatif.
2. Proses evaluasi diharapkan dapat memudahkan mendapatkan informasi mengenai Sentimen Masyarakat Terhadap Hastag disosial media Berdasarkan Opini dari *Twitter* yang berdasarkan kategori yaitu *netral, negatif, positif*, sehingga hasil yang didapat dari data sentimen masyarakat dapat dijadikan sebagai pendukung keputusan memberikan solusi terbaik terhadap peningkatan ekonomi kreatif di era new normal di Indonesia.

Dari kesimpulan yang telah dikemukakan, maka penulis ingin memberikan beberapa saran yang akan dijadikan sebagai bahan masukan yang bermanfaat. Untuk mendukung keberhasilan penelitian kedepan :

1. Teknik data mining mempunyai beberapa algoritma. Oleh karena itu, perlu dilakukan penelitian guna mencari *algoritma* paling efisien yang dapat diaplikasikan pada pengendalian proses.
2. Penelitian hanya dibatasi pada parameter-parameter proses dan kualitas yang dikendalikan. Oleh karena itu penelitian bisa diteruskan untuk parameter-parameter lain.
3. Semoga penelitian ini dapat dikembangkan lebih lanjut dan ucapan terima kasih kepada pihak L2Dikti Wilayah 2 dan Universitas Dehasen untuk penyelenggaraan kegiatan SNASIKOM 2021

REFERENSI

- Arief Rifiana. (2019). Analisis Sentimen Topik Viral Desa Penari Pada Media Sosial Twitter Dengan Metode Lexicon Based. Jurnal Ilmiah Matrik , Vol.21 No.3. DOI: <https://doi.org/10.33557/jurnalmatrik.v21i3.727>
- Ahmad Zuli Amrullah. (2020). Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square. Jurnal BITe : Jurnal Bumigora Information Technology. <https://journal.universitاسbumigora.ac.id/index.php/bite>
- Afrizal Sarika. (2019). Implementasi Metode Naive Bayes untuk Analisis Sentimen Warga Jakarta Terhadap Kehadiran Mass Rapid Transit. Jurnal Informatik Edisi Ke-15, Nomor 3. <https://ejournal.upnvj.ac.id/index.php/informatik/article/view/1454>
- Ati Sigit Heri S, Sri. (2012). Persepsi Pemustaka Terhadap Layanan Sirkulasi Di Badan Arsip Dan Perpustakaan Provinsi Jawa Tengah. <http://ejournal-s1.undip.ac.id/index.php/jip>
- Arsi Primandani. (2019). Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (Svm). Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK), Vol. 8, No. 1. DOI: 10.25126/jtiik.202183944
- Bustami. 2013. “Penerapan Algoritma Naive Bayes untuk Mengklasifikasi Data”. Nasabah Asuransi. Aceh
- Evans, Dave, 2008. “Social Media Marketing An Hour A Day, Wiley Publishing”, Inc : Canada
- Fitriyah Nur. (2020). Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (Svm). Jurnal Gaussian, Volume 9, Nomor 3. <https://ejournal3.undip.ac.id/index.php/gaussian/>
- Fitri Evita. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. TRANSFORMATIKA Vol. 18, No. 1. journals.usm.ac.id/index.php/transformatika
- Giovani Angelina Puput. (2020). Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi. Jurnal TEKNOINFO, Vol. 14, No. 2. DOI : 10.33365/jti.v14i2.679
- Hayuningtyas Ratih Yulia. (2019). Analisis Sentimen Opini Publik Bahasa Indonesia Terhadap Wisata Tmii Menggunakan Naive Bayes Dan Pso. Jurnal TECHNO Nusa Mandiri Vol. 16. DOI: <https://doi.org/10.33480/techno.v16i1.115>
- Hidayatullah Ahmad Fathan. (2014). Analisis Sentimen Dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter. Seminar Nasional Informatika UPN ”Veteran” Yogyakarta. <https://media.neliti.com/media/publications/174778-ID-analisis-sentimen-dan-klasifikasi-katego.pdf>
- Ikasari Diana. (2020). Analisis Sentimen Dan Klasifikasi Tweets Berbahasa Indonesia Terhadap Transportasi Umum Mrt Jakarta Menggunakan Naive Bayes Classifier. Jurnal Ilmiah Informatika Komputer Volume 25 No. 1. <http://dx.doi.org/10.35760/ik.2020.v25i1.2427>
- Kurniawan, Salamoon Daniel. (2018). Instagram, Ketika Foto Menjadi Mediator Komunikasi Lintas Budaya Di Dunia Maya”. Paper Universitas Kristen Petra Surabaya. Retrieved from http://repository.petra.ac.id/16642/1/Publikasi1_10021_1481.pdf
- Kusrini, luthfi taufiq Emha, (2009), Algoritma Data Mining, Penerbit Andi, Yogyakarta.

- Larose D, T. (2005). *Discovering knowledge in data : an introduction to data mining*. Jhon Wiley & Sons Inc
- Murnawan. (2017). Pemanfaatan Analisis Sentimen Untuk Peningkatan Popularitas Tujuan Wisata. *Jurnal Penelitian Pos dan Informatika*, Vol.07 No 02. DOI: 10.17933/jppi.2017.070203
- Putranti Noviah Dwi. (2014). Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan Maximum Entropy dan Support Vector Machine. *IJCCS*, Vol.8, No.1. <https://doi.org/10.22146/ijccs.3499>
- Rozi Imam Fahrur. (2012). Implementasi Opinion Mining (Analisis Sentimen) untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi. *Jurnal EECCIS* Vol. 6, No. 1. <https://jurnaleeccis.ub.ac.id/index.php/eccis/article/view/164>
- Sipayung vasaria M. (2016). Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier. *Jurnal Sistem Informasi (JSI)*, VOL. 8, NO. 1. <http://ejournal.unsri.ac.id/index.php/jsi/index>
- Sari Fransiska Vina. (2019). Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naive Bayes Classifier Berbasis Konversi Ikon Emosi. *Jurnal SIMETRIS*, Vol. 10 No. 2. DOI: <https://doi.org/10.24176/simet.v10i2.3487>
- Sugiono. (2014). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Widiarsono, Teguh. (2005). *Tutorial Praktis Belajar Matlab*. pdf, Jakarta