



JK. 2020;9 (2): 138 - 147

JURNAL ILMU KOMPUTER (COMPUTER SCIENCE JOURNAL)

<http://jik.htp.ac.id>

IMPLEMENTASI TEXT MINING PADA TWITTER DENGAN ALGORITMA K-MEANS CLUSTERING SEBAGAI DASAR KEBIJAKAN MARKETING BIRO PERJALANAN WISATA

Hendry Fonda¹, Budy Mustika², Eka Sabna³, Yuda Irawan⁴¹Sistem Informasi, STMIK Hang Tuah Pekanbaru^{2,3,4}Teknik Informatika, STMIK Hang Tuah Pekanbaru**Email :**fondaanda@gmail.com¹, mustikabudy@gmail.com², es3jelita@yahoo.com³, yudairawn89@gmail.com⁴

Abstract

Nowadays in modern times, internet and social media is an important thing in life. One of them is twitter. Every tweets by users is stored on the developer's storage media and is rarely used properly. A businessmen needs information about what is trending related to the types of business products. For this reason, the authors want to explore and classify these social data to obtain information that is useful for business. The authors used a descriptive analytical approach. This method describes or provides an overview of the object under study through data or samples that have been collected as is without analyzing and making general conclusions. In terms of data collection, it begins with downloading the uploaded data from Twitter users and then processing it with the K-Means algorithm. The results obtained from grouping the data produce data in the form of words that are currently being discussed. These results are used by businessmen to determine the direction of their product marketing policies. These results can also be used by businessmen as predictors of people's tendencies in traveling.

Keywords: Internet, K-Means, Twitter, Businessmen, Marketing.

Abstrak

Pada zaman modern seperti saat ini, penggunaan internet dan sosial media adalah instrumen yang wajib dalam menjalani kehidupan sehari-hari. Salah satunya twitter. Setiap unggahan pengguna twitter tersimpan pada media penyimpanan milik pengembangnya dan jarang dimanfaatkan dengan baik. Setiap pelaku usaha membutuhkan informasi tentang apa yang sedang tren terkait dengan jenis produk usahanya. Untuk itulah penulis ingin menggali dan mengelompokkan data bagi sosial tersebut untuk mendapatkan sebuah informasi yang kiranya bermanfaat untuk pelaku usaha. Dalam penelitian ini, penulis menggunakan metode pendekatan deskriptif analitis. Metode ini mendeskripsikan atau memberi gambaran terhadap objek yang diteliti melalui data atau sampel yang telah terkumpul sebagaimana adanya tanpa melakukan analisis dan membuat kesimpulan yang berlaku untuk umum. Dalam hal pengumpulan data dimulai dengan mengunduh data unggahan para pengguna twitter dan kemudian diolah dengan algoritma K-Means. Hasil yang diperoleh dari pengelompokan data tersebut menghasilkan data berupa kata-kata yang sedang tren dibicarakan. Hasil inilah yang digunakan oleh pelaku usaha untuk menentukan arah kebijakan pemasaran produknya. Hasil tersebut juga bisa dimanfaatkan oleh pelaku usaha sebagai bahan prediksi terhadap kecenderungan masyarakat dalam berwisata.

Keywords: : Internet, K-Means, Twitter, Pelaku Usaha, Pemasaran.<https://doi.org/10.33060/JIK/2020/Vol9.Iss2.180>

E-ISSN : 2579 - 3918 | P-ISSN : 2302 - 710X

PENDAHULUAN

Zaman dimana kita sekarang jalani ditandai dengan pertumbuhan data dan informasi yang banyak dan cepat terkumpulkan, disimpan, dan disediakan di media elektronik. Sebagian besar data bisnis disimpan dalam bentuk dokumen teks yang secara virtual sama sekali tidak terstruktur. Menurut kajian dari Merrill Lynch dan Gartner, 85 hingga 90 persen dari semua data korporat diambil dan disimpan dalam bentuk data tak terstruktur (McKnight, 2005). Dalam kajian yang sama juga dituliskan bahwa data tak terstruktur ini selalu menjadi dua kali lipat dalam hal ukuran (size) setiap 18 bulan sekali. Karena knowledge adalah power dalam dunia bisnis saat ini, dan knowledge diturunkan dari data dan informasi, organisasi bisnis yang bisa secara efektif dan efisien masuk ke beragam sumber data teks mereka akan memiliki knowledge yang diperlukan untuk membuat keputusan yang lebih baik, yang membawa ke keuntungan kompetitif atas berbagai bisnis yang sedang ketinggalan di belakang. Inilah yang mengakibatkan kebutuhan terhadap text mining cocok dengan gambaran besar bisnis hari ini.

PT. Hika Raya Berkah adalah perusahaan yang bergerak dalam bidang jasa biro perjalanan wisata yang melayani perjalanan wisata dengan tujuan domestik maupun internasional. Paket perjalanan yang ditawarkan pun beragam, mulai dari keberangkatan dengan kelompok (bersama-sama) maupun secara pribadi maupun keluarga (private). Dilihat dari banyaknya calon konsumen yang datang dan konsultasi tentang rencana perjalanan wisatanya, maka dirasa perlu sebuah prediksi untuk penjualan paket perjalanan wisata yang sesuai dengan selera konsumen.

Dalam menjalankan bisnisnya terutama dalam hal pemasaran produk paket perjalanan wisata, perusahaan ini

METODE PENELITIAN

Pada penelitian ini, penulis menggunakan pendekatan deskriptif analitis. Menurut (Sugiyono 2009:29) penelitian deskriptif analitis adalah suatu metode yang berfungsi untuk mendeskripsikan atau memberi gambaran terhadap objek yang diteliti melalui data atau sampel yang telah terkumpul sebagaimana adanya tanpa melakukan analisis dan membuat kesimpulan yang berlaku untuk umum.

Dengan kata lain penelitian deskriptif analitis mengambil masalah atau memusatkan perhatian kepada masalah-masalah sebagaimana adanya saat penelitian dilaksanakan, hasil penelitian yang kemudian diolah dan dianalisis untuk diambil kesimpulannya.

Secara garis besar, langkah-langkah yang perlu dilakukan pada pendekatan deskriptif analitis dapat digambarkan sebagai berikut :



Gambar 1 Langkah-langkah Pendekatan Deskriptif Analitis.

Langkah-langkah penelitian deskriptif analitis sebagaimana gambar 1 diatas dapat diuraikan sebagai berikut :

1. Merumuskan tujuan penelitian.

Dilakukan langkah-langkah sebagai berikut :

- Proses identifikasi dan merumuskan masalah yang hendak diteliti dan dapat dilaksanakan dengan sumber yang ada.
- Menentukan tujuan dari penelitian yang akan dikerjakan. Tujuan dari penelitian harus konsisten dengan rumusan dan definisi dari masalah.
- Menentukan batas masalah, sejauh mana penelitian deskriptif analitis akan dilaksanakan.
- Menelusuri sumber-sumber kepustakaan maupun jurnal yang ada hubungannya dengan masalah yang ingin dipecahkan.
- Merumuskan hipotesis.

2. Mengumpulkan dan mengorganisasi data.

Pada tahap ini diawali dengan autentikasi pada twitter (API), dilanjutkan dengan pengumpulan data berupa teks dengan cara diunduh dari twitter terkait dengan Destinasi Wisata paling banyak 250 (dua ratus lima puluh) tweet. Setelah teks diunduh, dilanjutkan dengan tahap Preprocessing data teks (Case folding, Tokenizing, Filtering dan Stemming).

3. Proses K-means Clustering menggunakan RapidMiner Studio. Pada tahap ini, data yang telah diunduh dan diolah tersebut, akan dilanjutkan dengan proses clustering menggunakan aplikasi RapidMiner Studio.

4. Menarik kesimpulan dan membuat laporan.

Setelah semua tahap diatas selesai dan telah diperoleh hasil, maka dilanjutkan dengan menarik kesimpulan dan membuat laporan akhir.

1. Analisis Sistem

Analisis sistem yaitu suatu tindakan mengumpulkan, mencari dan meneliti masalah-masalah yang akan dibahas dengan jelas

sehingga lebih dalam memecahkan suatu masalah, kebutuhan sistem yang diharapkan sehingga dapat di implementasikan sesuai dengan harapan.

Analisis sistem merupakan tahap yang penting dalam membangun sebuah sistem. Pada tahap ini dilakukan proses pengidentifikasian masalah yang mungkin bisa terjadi yang meliputi pengguna, perangkat keras maupun perangkat lunak.

4.1.1 Analisis Masalah

Analisis masalah implementasi text mining pada twitter dengan algoritma K-means clustering ini menggunakan metode PIECES (Performance, Information, Efeciency, Control, Economy.Service). Kerangka atau metode tersebut dapat di lihat pada uraian dibawah ini.

a. Performance (Kinerja)

Sistem Lama	Sistem Baru
Efisiensi kinerja pada sistem lama belum adanya suatu sistem sebagai pengambil keputusan untuk pemasaran paket perjalanan wisata.	kinerja perusahaan sangat terbantu, karena sistem dapat menyediakan data tren wisata yang dibutuhkan perusahaan.

b. Information (Informasi)

Sistem Lama	Sistem Baru
Informasi hanya diperoleh dari biro perjalanan lain, maskapai penerbangan maupun informasi dari media.	Adanya sistem <i>text mining</i> dapat menggali informasi pada <i>timeline</i> media sosial <i>twitter</i> , informasi tersebut yang dijadikan acuan untuk paket wisata. Dengan sistem <i>text mining</i> semua informasi yang di dapat akan di olah oleh sistem kemudian sistem memberikan informasi wisata mana yang ramai di bincangkan di media sosial <i>twitter</i> .

c. Economy (Ekonomi)

Sistem Lama	Sistem Baru
Perusahaan dinilai kurang ekonomis dalam menjalankan pemasarannya, misalnya karena masih menggunakan anggaran operasional yang kurang efektif.	Dari segi ekonomi pada sistem baru sangat memberikan manfaat bagi perusahaan, seperti meminimalisir penggunaan kertas, biaya operasional dan lain-lain, karena aplikasi dibangun secara <i>digital</i> .

d. Control (Kendali)

Sistem Lama	Sistem Baru
Sistem pemasaran produk yang sedang berjalan saat ini dari segi kendali belum ada.	Pada sistem baru sistem kendali lebih terjamin karena setiap pengambilan keputusan terkait pemasaran telah diambil berdasarkan data yang dihasilkan oleh sistem.

e. Efeciency (Efesien)

Sistem Lama	Sistem Baru
Masih menggunakan metode pemasaran yang lama, seperti brosur sehingga dianggap kurang efektif dan efisien dari segi waktu dan biaya.	Data yang dihasilkan oleh sistem, digunakan sebagai acuan pengambilan keputusan, sehingga dapat mengurangi biaya operasional maupun waktu.

f. Service (Layanan)

Sistem Lama	Sistem Baru
Sistem lama masih manual, belum terstruktur. Sehingga layanan yang diberikan oleh perusahaan belum tentu sesuai dengan selera konsumen.	Perusahaan dapat meningkatkan layanannya, karena produk wisata yang ditawarkan dianggap sesuai dengan selera konsumen.

2. Analisis Kebutuhan Sistem

Pada dasarnya sistem yang dibangun tidak jauh berbeda dengan sistem yang sudah ada. Sistem yang dibangun hanyalah peralihan dari sistem manual ke sistem digital. Dengan adanya implementasi text mining pada twitter dengan algoritma K-means clustering ini diharapkan dapat membantu dan mempermudah dalam menyelesaikan permasalahan yang terjadi.

Kebutuhan sistem terbagi menjadi dua, kebutuhan sistem fungsional dan kebutuhan sistem non-fungsional.

1. Kebutuhan Fungsional

Kebutuhan fungsional adalah kebutuhan-kebutuhan yang memiliki keterkaitan langsung dengan sistem. Kebutuhan fungsional juga berisi informasi-informasi apa saja yang harus ada dan dihasilkan oleh sistem. Dalam sistem ini pegawai atau admin dapat melakukan hal sebagai berikut :

1. Mengelola informasi trend twitter.
2. Mencari trending topic twitter pada sistem text mining.
3. Mendapatkan informasi tweet yang trend mengenai hasil pencarian.

2. Kebutuhan Non-fungsional

Kebutuhan non-fungsional adalah kebutuhan yang tidak secara langsung terkait dengan fitur tertentu di dalam sistem.

a. Kebutuhan perangkat keras

Adapun perangkat keras yang digunakan dalam membangun

aplikasi ini adalah sebagai berikut :

1. Prosesor Intel® Core i5-450M 2.30 GHz
2. RAM 4 Gigabyte
3. Harddisk 500 Gigabyte
4. Monitor LG 14 inch
5. Smartphone Android

b. Kebutuhan perangkat lunak

Adapun perangkat lunak yang digunakan dalam membangun aplikasi ini adalah sebagai berikut :

1. Sistem operasi Microsoft Windows 10
2. XAMPP Versi 3.2.4.
3. Notepad++
4. Web Browser Google Chrome/Mozilla Firefox.

3. Analisis Data

Dalam mengembangkan sistem yang sedang dirancang, tentunya penulis terlebih dahulu mempelajari dengan seksama data- data yang ada dan yang akan dibutuhkan nantinya oleh perusahaan. Hal ini sangat penting karena sistem yang dirancang sepenuhnya berdasar pada data yang ada serta data yang sekiranya akan dibutuhkan. Data yang akan dijadikan masukan adalah berupa kata kunci yang menjadi dasar oleh sistem dalam melakukan pencarian. Misal menggunakan kata kunci "wisata pantai", maka sistem akan melakukan pencarian di twitter terkait kata kunci tersebut yang sedang ramai dibicarakan. Hasilnya bisa bermacam-macam, mulai dari tujuan yang paling banyak dibicarakan hingga perkiraan biaya yang dibutuhkan dalam sebuah paket perjalanan wisata.

4. Analisis Perangkat Keras

Berikut ini perangkat keras yang digunakan dalam analisis dan perancangan sistem yang dapat di lihat pada tabel 1 dibawah ini.

Tabel 1 Analisa Perangkat Keras

No	Nama Hardware	Spesifikasi	Jml	Fungsi
1	Laptop	Processor intel core i5, Ram 8 Gb, Harddisk 500 Gb	1	Sebagai alat antarmuka.
2	Printer	Canon MP237	1	Sebagai alat pencetak
Total				

5. Analisis Perangkat Lunak

Berikut ini perangkat lunak yang digunakan dalam analisis dan perancangan sistem yang dapat di lihat pada tabel 2.

Tabel 2. Analisa Perangkat Lunak

No	Kebutuhan	Keterangan	Fungsi
1	Sistem Operasi	Windows 10 64-bit	Mengontrol operasi-operasi dasar sistem, termasuk menjalankan aplikasi.
2	Aplikasi	Notepad++	Pembuatan coding atau sebagai text editor
		Xampp v.3.2	Sebagai server web untuk komputer lokal
		Chrome Browser	Web browser
		Star UML	Tools untuk membuat usecase diagram
		PHP	Bahasa pemrograman
		Rapidminer Studio v. 9.7	Tools untuk olah data text mining dan proses K-Means Clustering.

1.3.1 Analisis Biaya

Perancangan implementasi text mining pada twitter dengan algoritma K-means clustering ini diperlukan biaya yang digunakan dalam tahap perancangan sistem hingga tahap implementasi sistem, berikut analisis biaya yang dibutuhkan.

Tabel 3. Analisa Biaya

No	Kegiatan	Analisa Biaya
1	Proposal skripsi	Rp. 800.000,-
2	Analisa sistem	Rp. 1.000.000,-
3	Desain sistem	Rp. 1.000.000,-
4	Pembuatan program	Rp. 5.000.000,-
5	Implementasi sistem	Rp. 500.000,-
6	Instalasi sistem	Rp. 100.000,-
7	Pelatihan user	Rp. 1.000.000,-
	Total	Rp. 9.400.000

PERANCANGAN SISTEM

A. Usecase Diagram

Perancangan dimulai dengan mengidentifikasi actor dan use case. Use case digunakan untuk memberikan gambaran secara umum dari sistem yang akan dirancang. Actor yang terdapat pada sistem ini hanya satu yaitu pengguna dari aplikasi. Gambaran keseluruhan interaksi antara actor dengan use case terdapat pada Gambar 2 dibawah ini :



B. Skenario Usecase Diagram

Dari gambar usecase diagram diatas, maka dapat dibuat skenario usecase untuk memperjelas proses pada usecase diagram sebagai berikut :

- Nama Use Case : Cari kata trending
- Pre Kondisi : Berada di menu admin
- Post Kondisi : Berhasil mencari data

Aksi Aktor "Admin"	Reaksi Aplikasi
Skenario Normal	
1. Meng-klik menu pencarian	
	2. Menampilkan form pencarian data
3. Input kata pencarian	
	4. Data hasil pencarian berhasil ditampilkan

- Nama Use Case : Melihat informasi data API twitter
- Pre Kondisi : Berada di menu hasil pencarian
- Post Kondisi : Berhasil download data

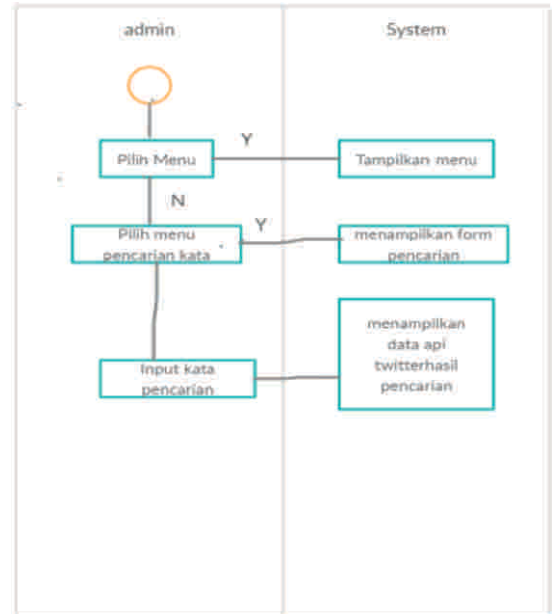
Aksi Aktor "Admin"	Reaksi Aplikasi
Skenario Normal	
1. Mengklik menu hasil pencarian	
	2. Menampilkan data hasil pencarian
3. download data hasil pencarian dalam bentuk excel atau pdf	
	4. Data berhasil didownload

- Nama Use Case : olah data API twitter
- Pre Kondisi : masuk ke aplikasi rapidminer
- Post Kondisi : Berhasil mengolah data

Aksi Aktor "Admin"	Reaksi Aplikasi
Skenario Normal	
1. Masuk ke aplikasi rapidminer	
	2. Menampilkan form olah data
3. melakukan upload data API twitter	
	4. Data berhasil diolah dan di jadikan informasi untuk paket wisata

C. Activity Diagram

Activity diagram menggambarkan aktifitas-aktifitas yang terjadi dalam sistem. Activity diagram untuk admin dapat dilihat pada gambar 4. dibawah ini.



Gambar 4. Activiy Diagram Admin

A. Rancangan User Interface

Adapun beberapa rancangan user interface implementasi text mining pada twitter dengan algoritma K-means yang tertera sebagai berikut :

1. User Interface Halaman Pencarian Kata

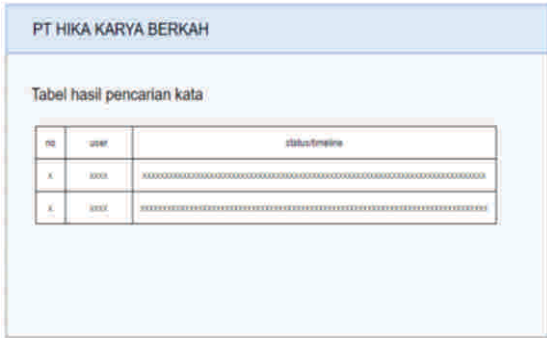


Gambar 4. Halaman Web

User interface pada Gambar 4.3 merupakan gambaran kasar untuk halaman Web untuk admin disini admin dapat melakukan pencarian kata trending topic pada data API twitter.

1. User Interface Halaman Hasil Pencarian

Setelah admin melakukan pencarian, maka sistem akan menampilkan halaman hasil pencarian dalam bentuk tabel seperti pada gambar 4.4 dibawah ini.



Gambar 5. Halaman Utama

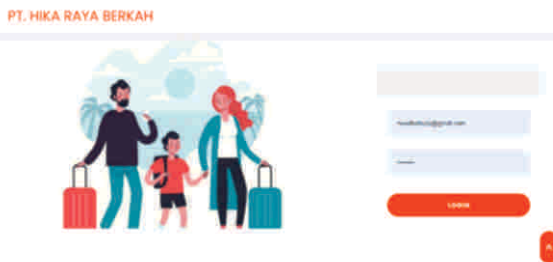
Prototipe pada gambar 4.4 merupakan rancangan untuk halaman hasil pencarian data, data yang dihasilkan didapat dari data API twitter seperti nama pengguna dan status atau timeline pengguna.

HASIL & PEMBAHASAN

Implementasi text mining pada twitter dengan algoritma K-means clustering di PT Hika Raya Berkah ini menghasilkan kemudahan dan efektifitas dalam mengelola paket wisata karena sistem ini dibangun menggunakan text mining untuk memproses data sehingga data dan informasi yang di dapat lebih akurat dengan pengelompokan (clustering) data berupa teks di media sosial twitter dengan menggunakan algoritma K-means. Metode ini merupakan metode yang populer digunakan untuk mendapatkan deskripsi dari sekumpulan data dengan cara mengungkapkan kecenderungan setiap individu.

a. Halaman utama dan login

Adalah halaman awal yang meminta user untuk masuk kedalam sistem. Halaman ini dapat dilihat pada gambar 6 dibawah ini.



Gambar 6. Halaman utama dan login

b. Halaman Pencarian

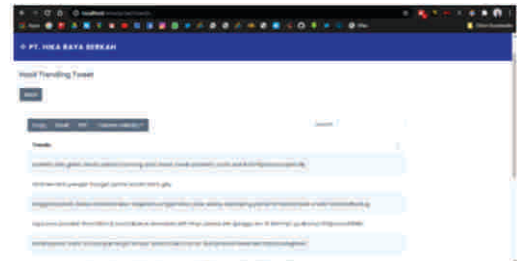
Pada halaman ini diketikkan kata kunci untuk mendapatkan data yang diinginkan. Kata kunci yang digunakan adalah "wisata pantai". Halaman pencarian dapat dilihat pada Gambar 7.



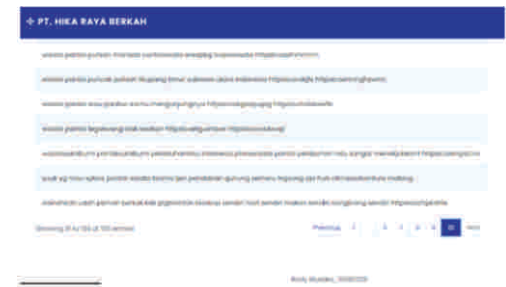
Gambar 7. Halaman pencarian

c. Halaman hasil pencarian

Pada halaman ini akan ditampilkan data tweet dari pengguna twitter yang berkaitan dengan kata kunci "wisata pantai" tersebut. data ini dapat di unduh dengan klik tombol "excel" yang terdapat di bagian atas sebelah kiri halaman Halaman ini dapat dilihat pada Gambar 8 dan 9 dibawah ini.



Gambar 8. Halaman awal hasil pencarian



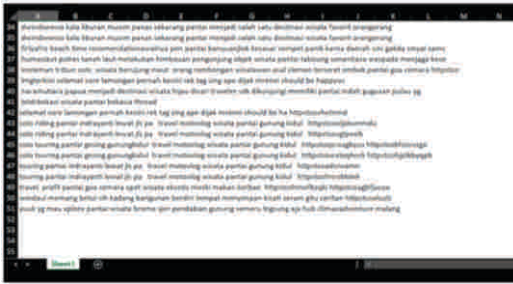
Gambar 9. Halaman akhir hasil pencarian

d. Hasil unduhan

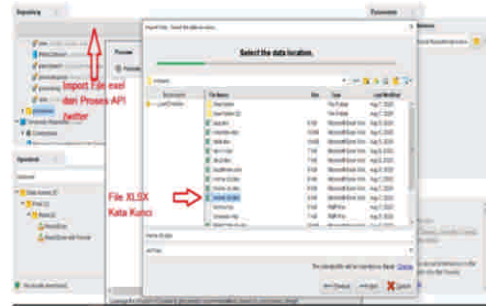
Hasil unduhan berupa data dengan format "xlsx" inilah yang dijadikan data awal hasil pencarian terpopuler pada twitter. Halaman excel dapat dilihat pada Gambar 10 dan 11 dibawah ini.



Gambar 10. Halaman awal Microsoft Excel



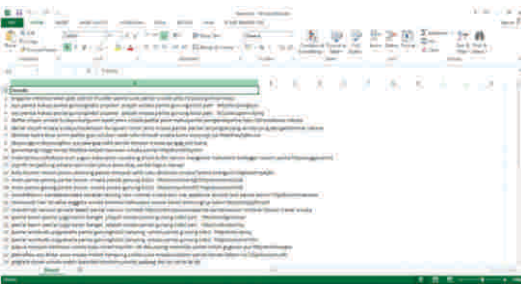
Gambar 11 Halaman akhir Microsoft Excel



Gambar 14. Halaman Import data

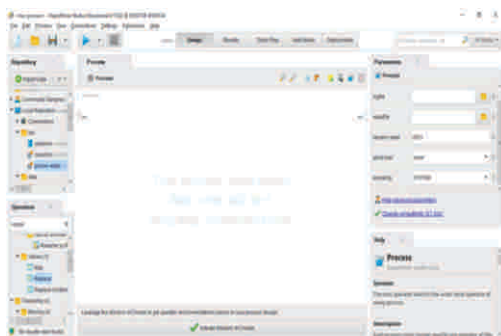
e. Halaman Hasil Case Folding

Case folding adalah mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf 'a' sampai dengan 'z' yang diterima. Karakter selain huruf dihilangkan dan dianggap delimiter. Hasil ini dapat dilihat pada gambar 12 dibawah ini.



Gambar 12. Halaman Hasil Case Folding

f. Halaman Software Rapidminer



Gambar 13. Halaman Software Rapidminer

g. Halaman memasukan data Tokenize

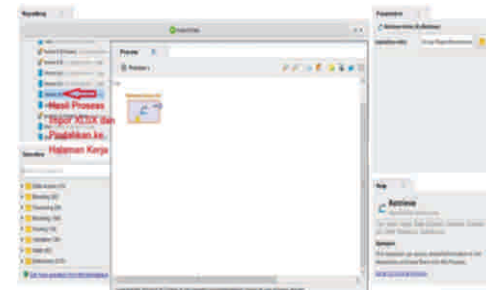
Halaman ini merupakan halaman Tahap Tokenize, dimana tahap pemotongan string input berdasarkan tiap kata yang menyusunnya setelah hasil dari Case folding

1. Import Data XLSX hasil dari pencarian kata kunci.
2. Pilih File dan tekan Next – Finish

Contoh dari tahap ini dapat dilihat pada gambar 14 dibawah ini.

h. Halaman memasukan hasil import data ke proses Tokenize

Pada tahapan ini bertujuan untuk memasukan data excel yang sudah di import tadi untuk di proses. Contoh dari tahap ini dapat dilihat gambar 15 dibawah ini.



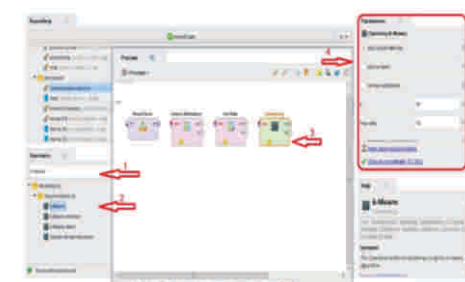
Gambar 15. Halaman memasukan hasil import data ke proses Tokenizing

i. Proses memasukan tools clustering K-means

Ini adalah tools penting untuk menentukan metode K-means sebagai pencari hasil dari kata kunci populer tentang wisata di twitter. Langkah-langkahnya sebagai berikut :

1. Cari tools K-means pada kolom pencarian di operator.
2. Tambahkan tools K-means pada halaman proses K-means.
3. Klik tools K-means.
4. Atur pada bagian parameters.

Contoh pada tahapan ini dapat dilihat pada gambar 16 dibawah ini,

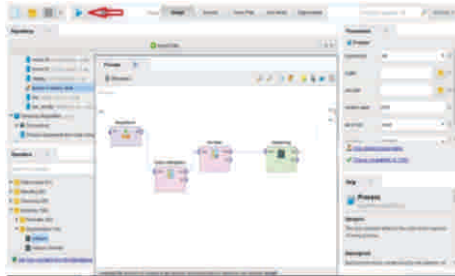


Gambar 16. Proses Memasukan Tools Clustering K-means

j. Halaman proses Clustering

Pada halaman ini semua tools yang telah terhubung siap untuk dijalankan, dengan cara klik tombol play pada menu rapidminer.

Contoh tahapan ini dapat dilihat pada gambar 17 dibawah ini.

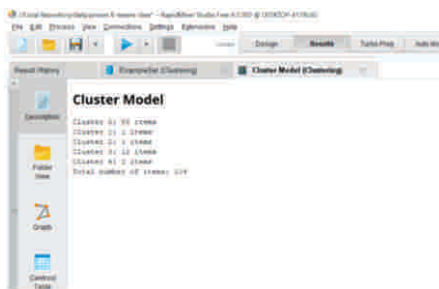


Gambar 17. Halaman Proses Clustering

k. Halaman hasil dari K-means Cluster Model

Halaman ini berisikan informasi tentang jumlah cluster yang ada beserta dengan nilai perkelompok cluster.

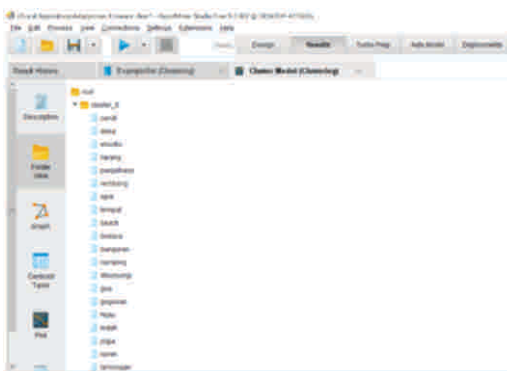
Contoh tahapan ini dapat dilihat pada gambar 18 dibawah ini.



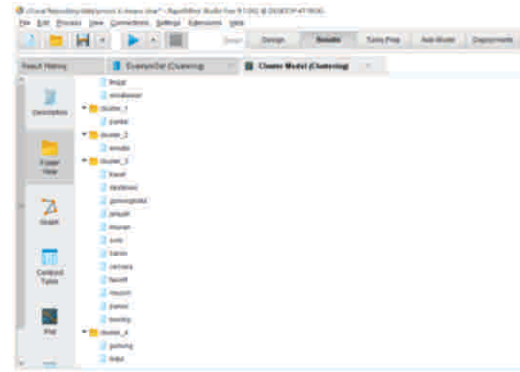
Gambar 18. Halaman hasil K-means Cluster Model

l. Halaman isi dari setiap cluster

Contoh pada tahapan ini dapat dilihat pada gambar 19 dan 20 dibawah ini.



Gambar 19. Halaman isi dari setiap Cluster



Gambar 20. Halaman isi dari setiap Cluster

Jika diuraikan, masing-masing cluster memuat kata-kata sebagaimana dijelaskan pada tabel 4 dibawah ini.

Tabel 4. Isi dari tiap cluster

No	Cluster	Penjelasan
1	1	Memuat kata : pantai
2	2	Memuat kata : wisata
3	3	Memuat kata : travel, destinasi, gunungkidul, jelajah, liburan, solo, baron, cemara, favorit, musim, panas, touring.
	4	Memuat kata : gunung dan kidul
	0	Memuat kata : candi, desa, eksotis, karang, punjulharjo, rembang, spot, tempat, beach, budaya, bunguran, camping, dikunjungi, gua, gugusan, hijau, indah, jogja, keren, lombok, lamongan, natuna, objek, papua, populer, selamat, suka, taman, traveller, jogyakarta, ancol, area, bagus, bangunan, barat, batu, baturaden, berujung, betul, blitar, bromo, coklat, daerah, dufan, gelombang, hutan, ijen, inisiemen, kakao, kampung, kawasan, kecamatan, laut, lokasi, malang, malioboro, mangrove, maruspantai, menikmati, merapi, naramutiara, objek, ombak, outdoor, padang, pantaimawun, pasir, pendakian, penganakpantai, pengunjung, pinus, point, pulau, rombongan, sanitanjung, semeru, sempet, senantiasa, sensasi, seram, sleman, sluban, tanah, tanjungtanjung, terseret, timur, tinggi, wisatawan.

m. Halaman hasil K-means dalam bentuk Line

Pada halaman ini kita bisa mengetahui topik apa yang paling sering dibicarakan di twitter pada line yang paling tinggi, pada proses ini orang di twitter banyak membicarakan tentang pantai.

Contoh pada tahapan ini biasa dilihat pada gambar 21 dibawah ini.



Gambar 21. Halaman Hasil K-means dalam bentuk line

n. Hasil Analisis

Berdasarkan analisis data yang telah dilakukan terhadap tweets dengan kata pencarian “wisata pantai” dan dilanjutkan dengan clustering menggunakan RapidMiner Studio, menghasilkan kata “pantai” dan “kidul” menempati posisi paling tinggi. Sedangkan jika dilihat berdasarkan nama pantai, diperoleh hasil kata “gunung kidul” dan “lombok”. Hasil ini dapat memberikan informasi kepada perusahaan bahwa objek wisata pantai yang populer pada saat ini adalah pantai Gunung Kidul di Yogyakarta dan pantai Lombok di Nusa Tenggara Barat.

Hasil analisis dapat diuraikan pada tabel 5.2 dibawah ini,

No	Kata Populer Hasil Analisis	Cluster
1	Pantai	1
2	Kidul	3 dan 4
3	Lombok	0

Tabel 5.2. Kata Populer Hasil Analisis.

Hasil analisis data inilah yang dijadikan dasar oleh perusahaan dalam menentukan arah kebijakan pemasaran produknya. Langkah-langkah yang bisa diambil antara lain :

1. Perusahaan memulai kerjasama dengan penyedia jasa wisata di Yogyakarta maupun di Lombok;
2. Perusahaan juga melakukan kerjasama dengan penyedia jasa lainnya seperti maskapai penerbangan, hotel, wahana permainan, restaurant maupun jasa angkutan penumpang pada lokasi-lokasi tersebut.
3. Setelah adanya kerjasama dengan berbagai pihak sebagaimana diuraikan pada dua poin diatas, perusahaan dapat menyediakan paket perjalanan ke daerah tersebut dan memulai promosi. Promosi yang

dimaksud dapat berupa konten atau iklan pada media sosial maupun elektronik.

KESIMPULAN DAN SARAN

Kesimpulan

Setelah melakukan analisis dan pengujian pada bab-bab sebelumnya, dapat disimpulkan sebagai berikut :

1. Text Mining dengan algoritma K-means Clustering yang diimplementasikan pada media sosial twitter, dapat dijadikan salah satu cara oleh perusahaan dalam menentukan arah kebijakan pemasarannya. Hasil yang diperoleh dari sistem ini menunjukkan informasi wisata yang populer dan sedang ramai dibicarakan ditengah masyarakat.
2. Hasil yang diperoleh tersebut adalah kata-kata yang populer pada waktu pencarian dilakukan.
3. Hasil analisis data yang dilakukan terhadap tweet dengan kata pencarian “wisata pantai” menghasilkan kata “gunung kidul” dan “lombok” sebagai tren.
4. Dengan sistem text mining yang memanfaatkan pencarian terpopuler pada twitter dapat dijadikan pertimbangan oleh perusahaan dalam menentukan arah kebijakan pemasaran produk wisatanya.

Saran

Berdasarkan kesimpulan diatas masih ada hal lain yang bisa dikembangkan untuk penelitian selanjutnya, antara lain :

1. Menyediakan tempat (space) khusus pada sistem untuk mengelompokkan pencarian berdasarkan jenis, sifat maupun biaya yang dibutuhkan untuk mengunjungi suatu objek wisata.
2. Perlu dikembangkan aplikasi berbasis android untuk mempermudah perusahaan biro perjalanan wisata dalam mengetahui tentang tempat wisata yang sedang populer, sehingga perusahaan mendapatkan informasi yang akurat terkait wisata yang sedang populer.
3. Untuk kedepan perlu adanya implementasi text mining pada aplikasi media sosial yang lain untuk meningkatkan informasi yang lebih banyak seperti Facebook dan Instagram.

DAFTAR PUSTAKA

- Budiyono. 2003. Metodologi Penelitian Pendidikan. Surakarta : Universitas Sebelas Maret.
- Daniel Riano Kaparang dan Eko Sedyono. (2013). “Penentuan Alih Fungsi Lahan Marginal Menjadi Lahan Pangan Berbasis Algoritma K-means di Wilayah Kabupaten Boyolali.” JdC. 2. 20.
- Irawan, Y. (2019). Sistem Informasi Pemasaran Busana Syar’i dengan Penerapan Customer Relationship Management (CRM) Berbasis Web. INTECOMS: Journal of Information Technology and Computer Science, 2(1), 1-9.



- ROFIQOH, Ummi; IRAWAN, Yuda; MELYANTI, Rika. APLIKASI PENDAFTARAN DAN PENERIMAAN DATA PERSYARATAN CALON ANGGOTA LEGISLATIF SEMENTARA SECARA ONLINE PADA PARTAI KEADILAN SEJAHTERA DI KABUPATEN BENGKALIS. *Riau Journal Of Computer Science*, 2020, 6.2: 145-153.
- Eriyanto, Ocki. 2012. Analisis Peramalan Penjualan Handphone Balckberry pada PT Selular Shop Mall.
- Han, J., Kamber, M., Data Mining Concept and Technique, 2nd Ed, Elsevier, 2006.
- Hermawati, Fajar Astuti. 2013. Data mining. Yogyakarta : Penerbit Andi
- <http://jagocoding.com>, yang diakses pada Minggu, 01 Maret 2020 pukul 20.40 WIB
- Michael, B., "Automatic Discovery of Similar Words," Survey of Text Mining: Clustering, Classification and Retrieval, LLC, pp. 24-43, 2004.
- Nazir. M. 2003. Metode Penelitian. Jakarta : Ghalia Indonesia
- Irawan, Y., Wahyuni, R., & Herianto, H. (2019). Media Pembelajaran Bahasa Inggris Dasar Menggunakan Macromedia Flash 8 di Tk Kartika 1.50 Kecamatan Sail Kota Pekanbaru. *Informatika*, 11(2), 1-7.
- Mardeni, M. (2020). Information System For Monitoring Of Pregnant Mother In Riau Province Health Office. *Journal of Applied Engineering and Technological Science (JAETS)*, 1(2), 85-90.
- Sabna, Eka. "ANALISIS DATA MAHASISWA DENGAN ALGORITMA K-MEAN UNTUK Mendukung STRATEGI PROMOSI STMIK HANG TUAH PEKANBARU." *Jurnal Ilmu Komputer* 8.1 (2019): 113-117.
- Irawan, Y. (2019). Aplikasi E-Commerce Untuk Pemasaran Kerajinan Tangan Usaha Kecil Menengah (UKM) di Riau Menggunakan Teknik Dropshipping. *Jurnal Ilmiah Core IT: Community Research Information Technology*, 7(1).
- Wahyuni, R., & Irawan, Y. (2019). Web-Based Heart Disease Diagnosis System With Forward Chaining Method (Case Study Of Ibnu Sina Islamic Hospital). *Journal of Applied Engineering and Technological Science (JAETS)*, 1(1), 43-50.
- Sabna, Eka, and Muhardi Muhardi. "Penerapan Data Mining Untuk Memprediksi Prestasi Akademik Mahasiswa Berdasarkan Dosen, Motivasi, Kedisiplinan, Ekonomi, Dan Hasil Belajar." *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi* 2.2 (2016): 41-44.
- Irawan, Y. (2019). Implementation Of Data Mining For Determining Majors Using K-Means Algorithm In Students Of SMA Negeri 1 Pangkalan Kerinci. *Journal of Applied Engineering and Technological Science (JAETS)*, 1(1), 17-29.
- Sabna, Eka. "PEMANFAATAN DATA MINING UNTUK PENEMPATAN BUKU DI PERPUSTAKAAN MENGGUNAKAN METODE ASSOCIATION RULE." *Jurnal Ilmu Komputer* 8.2 (2019): 59-63.
- Santoso R., "Aplikasi Katalog Online untuk Pencarian Konten Buku dengan Metode Text mining pada Perpustakaan STIKOM Surabaya," Skripsi, STIKOM, Surabaya, 2012.
- Sosial Media. <http://www.unpas.ac.id/apa-itu-sosial-media/>, diakses pada hari Jumat, 21 Februari 2020, pukul 20.15 WIB
- Lenardo, G. C., & Irawan, Y. (2020). Pemanfaatan Bot Telegram sebagai Media Informasi Akademik di STMIK Hang Tuah Pekanbaru. *JTIM: Jurnal Teknologi Informasi dan Multimedia*, 1(4), 351-357.
- Sugiyono, "Metode Penelitian Kuantitatif Kualitatif dan R&D", Bandung: Alfabeta, 2010.
- Irawan, Y. (2019). Penerapan Data Mining Untuk Evaluasi Data Penjualan Menggunakan Metode Clustering Dan Algoritma Hirarki Divisive Di Perusahaan Media World Pekanbaru. *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM)*, 4(1), 13-20.
- Fonda, H. (2020). KLASIFIKASI BATIK RIAU DENGAN MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORKS (CNN). *Jurnal Ilmu Komputer*, 9(1), 7-10.
- MUHARDI, Muhardi, et al. Design Of Web Based LMS (Learning Management System) in SMAN 1 Kampar Kiri Hilir. *Journal of Applied Engineering and Technological Science (JAETS)*, 2020, 1.2: 70-76.
- Hussein, O. S., Wahyuni, R., & Mukhtar, H. (2018). Sistem informasi deteksi kehadiran dan media penyampaian pengumuman dosen dengan menggunakan teknik pengenalan QR code. *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, 3(2), 85-92.