



**FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO SEMARANG**

JL. IMAM BONJOL NO. 207 SEMARANG TELP. 024-3575915, 024-3575916

JAWABAN UJIAN TENGAH SEMESTER GANJIL 2022/2023

Mata Kuliah : Analitika Media Sosial	Sifat : Take Home
Hari/tanggal : Jumat, 3 November 2023	Waktu : 10.20 – 12.00
Kelompok : A12.6503	Dosen : Ika Novita Dewi, MCS
NIM : A12.2021.06612	Nama : Maulida Aristia Tsani

A. Judul

Analisis Sentimen Terhadap Komentar Youtube Tentang Diet Sehat Menggunakan Algoritma Naïve Bayes

B. Pendahuluan

GERMAS (Gerakan Masyarakat Hidup Sehat) merupakan suatu gerakan yang mempunyai tujuan untuk meningkatkan kesadaran, dan pemahaman masyarakat terhadap perilaku gaya hidup sehat serta meninggalkan kebiasaan dan perilaku masyarakat yang kurang sehat. Salah satu contoh GERMAS yang sering dilakukan yaitu diet dengan sehat. Diet pada dasarnya adalah pola makan yang cara dan jenis makanannya diatur dan bertujuan untuk menjaga kesehatan tubuh secara keseluruhan serta mencapai berat badan yang diinginkan. Tak sedikit masyarakat Indonesia yang melakukan program diet. Berdasarkan survey Kurious-Katadata Insight Center (KIC) yang menunjukkan bahwa 69,9% dari total responden mengaku melakukan diet. Ditinjau dari alasan, mayoritas atau sebanyak 75,7% responden menyebutkan bahwa menjaga kesehatan adalah alasan utama melakukan diet. Sementara diet dengan alasan menurunkan berat badan mempunyai persentase 63,1%.

Zaman sekarang muncul makanan-makanan cepat saji yang tidak mempunyai khasiat bagi tubuh. Masyarakat juga sering mengonsumsi makanan cepat saji tersebut, bahkan sebagian masyarakat makan makanan cepat saji bisa sampai 3-5 kali sehari. Jika, kebiasaan tersebut sering dilakukan dan tidak disertai dengan olahraga bisa jadi menimbulkan penyakit seperti obesitas, kanker dan sebagainya. Kegemukan atau obesitas menjadi salah satu persoalan yang dapat mengakibatkan



FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO SEMARANG

Jl. IMAM BONJOL NO. 207 SEMARANG TELP. 024-3575915, 024-3575916

depresi, cemas dan stress. Hal tersebut terjadi karena banyak masyarakat yang menganggap obesitas itu dinilai tidak sesuai dengan standar berat badan. Dengan demikian, mereka berlomba-lomba untuk melakukan diet sehat. Tak jarang, masyarakat melakukan diet sehat namun hanya dilakukan beberapa kali saja dengan alasan malas untuk melakukan diet sehat karena prosesnya yang lama. Namun, masih juga terdapat sebagian masyarakat yang melakukan program diet secara ekstrem, padahal diet ekstrem itu tidak sehat bagi tubuh. Diet sehat dapat dilakukan dengan cara mengonsumsi buah-buahan dan sayuran secara teratur dan disertai juga dengan olahraga.

Penelitian ini bertema Gerakan Masyarakat Hidup Sehat dengan topik Diet Sehat dan mengambil salah satu video Youtube akun SKWAD Fitness. Tujuan penelitian ini untuk menganalisis word cloud dan menganalisis sentimen. Analisis word cloud bertujuan untuk mengetahui kata-kata yang sering muncul dan yang sering dibahas oleh masyarakat. Sedangkan analisis sentimen bertujuan untuk mengetahui opini masyarakat tentang bagaimana cara melakukan diet sehat yang baik dan benar.

C. Metode Penelitian

1. Pengumpulan Data

Data yang digunakan untuk dilakukan analisis sentimen dan word cloud diambil dari media sosial Youtube dengan link sebagai berikut : <https://www.youtube.com/watch?v=pDcFSHH-4yU&t=448s> yang kemudian data tersebut di upload di netlytic dan menghasilkan berupa file dataset netlytic.

2. Text Preprocessing

Langkah-langkah membersihkan dataset dari netlytic :

- Download dataset dari netlytic berupa excel, kemudian pisahkan atribut description
- Copy dan Paste atribut description tersebut ke file excel yang baru, kemudian save as ke .csv dan rename menjadi Deskripsi_Diet_Sehat
- Buka aplikasi rapid miner di laptop masing-masing



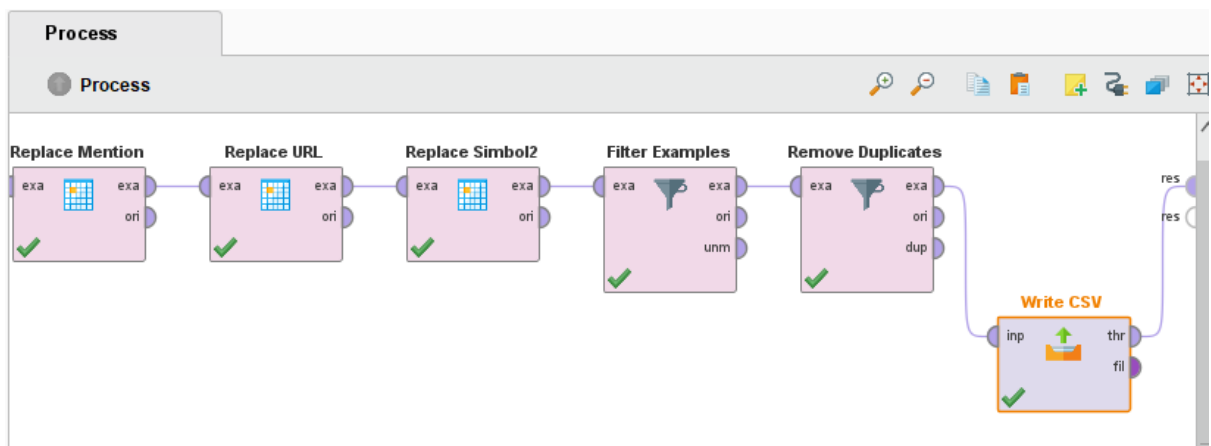
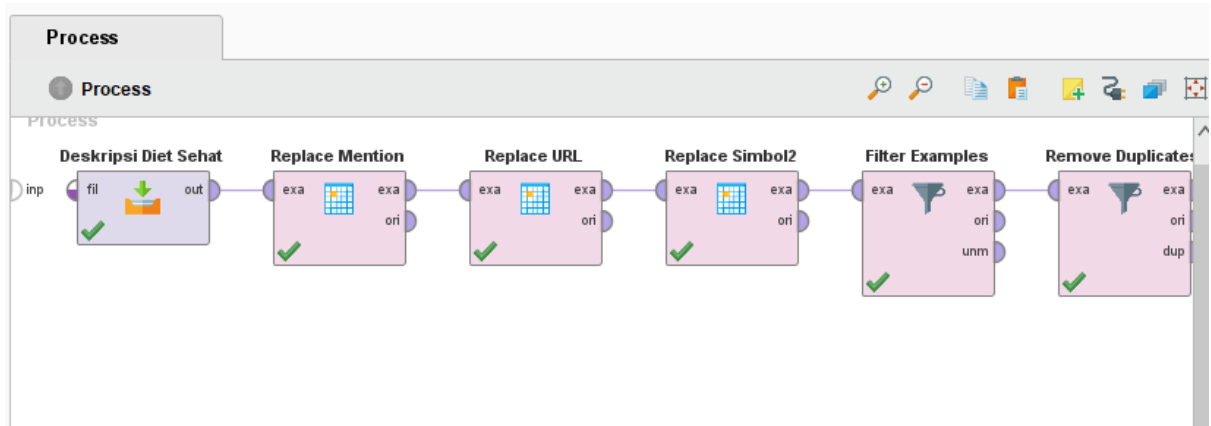
**FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO SEMARANG**

Jl. IMAM BONJOL NO. 207 SEMARANG TELP. 024-3575915, 024-3575916

- Carilah operator read CSV, setelah berhasil di drag ke layar rapid miner operator tersebut kemudian unduh file csv (Deskripsi_Diet_sehat) ke operator read csv
- Pilih operator replace, operator ini berfungsi untuk mengubah atau menghapus karakter atau kata tertentu yang diinput.
- Pertama, replace mention yang berarti menghapus seluruh mention pada teks. Di pojok kanan terdapat kolom replace what, dalam kolom tersebut diisi regular expression berupa @.* (spasi) kemudian apply, tarik garis ke hasil dan jalankan. Setelah berhasil dijalankan dicek terlebih dahulu, apabila mention sudah tidak ada maka replace mention berhasil.
- Kedua, replace URL yang berarti menghapus seluruh link (https) pada teks. Di pojok kanan terdapat kolom replace what, dalam kolom tersebut diisi regular expression berupa https.*? (spasi) kemudian apply, tarik garis ke hasil dan jalankan. Setelah berhasil dijalankan, dicek terlebih dahulu, apabila link url tersebut sudah tidak ada maka replace URL berhasil.
- Ketiga, replace symbol-simbol yang berarti menghapus seluruh simbol pada teks. Di pojok kanan terdapat kolom replace what, dalam kolom tersebut diisi regular expression berupa [!~?.,;:'"##*@\$&()-] kemudian apply, tarik garis ke hasil dan jalankan. Setelah berhasil dijalankan, dicek terlebih dahulu, apabila link url tersebut sudah tidak ada maka replace URL berhasil.
- Keempat, drag operator trim ke canvas rapidminer. Operator ini berfungsi untuk menghapus whitespace atau spasi-spasi yang tidak perlu. Kemudian tarik garis ke hasil dan jalankan. Setelah berhasil dijalankan bandingkan dengan data sebelumnya yaitu data ori yang ditarik garis ke hasil.
- Kelima, drag operator filter examples. Operator ini berfungsi untuk menghapus data yang kosong. Pada pojok kanan bagian add filters, diklik kemudian pilih atribut text dan value is not missing yang berarti data yang tidak dihapus adalah data yang tidak kosong, dan yang dihapus adalah data yang kosong. Kemudian tarik garis ke hasil dan jalankan, setelah di cek

bandingkan dengan data sebelumnya, jika data yang kosong telah hilang maka operator ini berhasil.

- Keenam, drag operator remove duplicates. Operator ini berfungsi untuk menghapus kalimat yang sama. Pada pojok kanan bagian attribut diisi single dan value diisi text. Kemudian tarik garis ke hasil dan jalankan setelah di cek bandingkan dengan data sebelumnya, jika data yang sama telah hilang maka operator ini berhasil.
- Selanjutnya, klik operator write csv. Operator ini berfungsi untuk mengexport data dalam bentuk csv. Tarik garis dari operator remove duplicates ke write csv dan dari thr write csv ditarik garis ke hasil. Selanjutnya, di pojok kanan bagian csv file klik dan cari lokasi penyimpanan file tersebut lalu rename.





**FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO SEMARANG**

Jl. IMAM BONJOL NO. 207 SEMARANG TELP. 024-3575915, 024-3575916

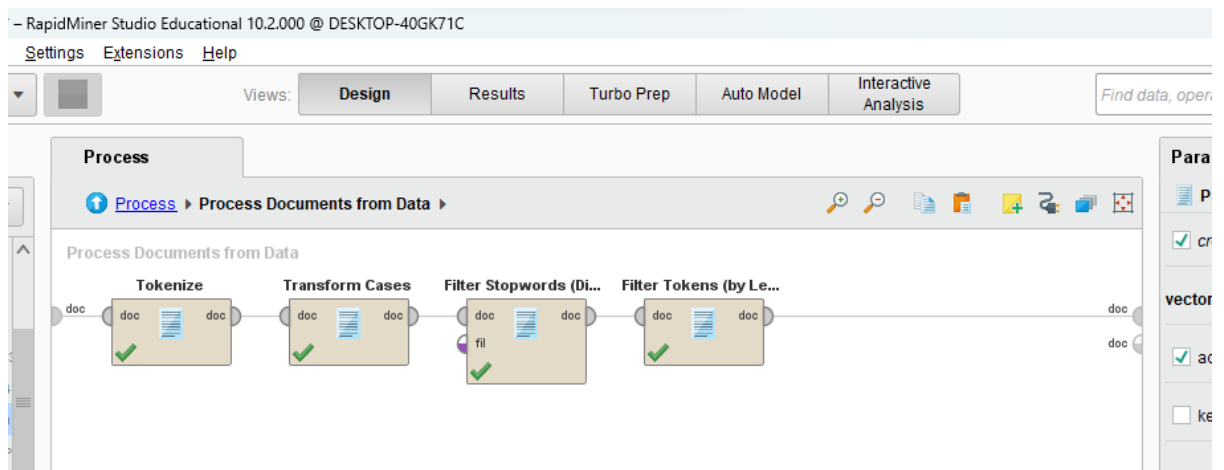
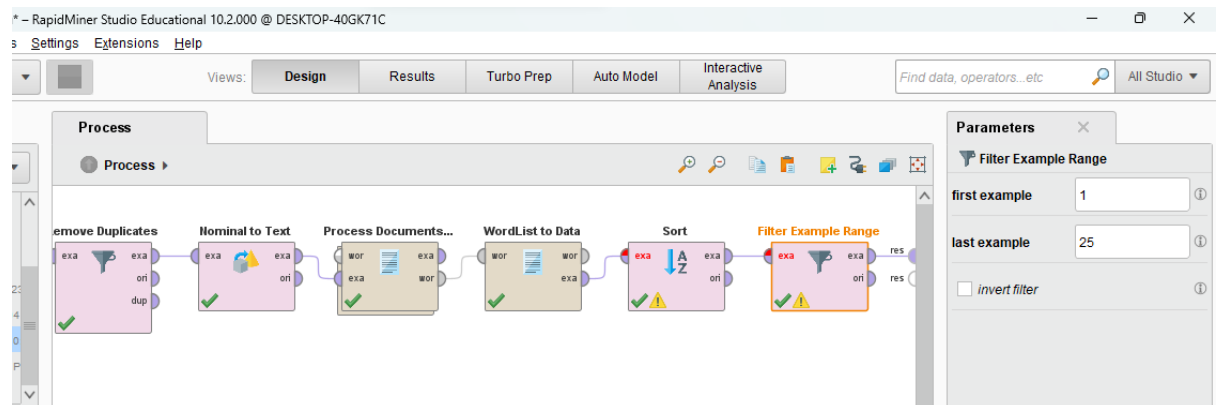
3. Document Processing

Langkah-langkah document processing

- Data masih lanjut dari data preprocessing, lalu pilih operator nominal to text. Operator ini berfungsi untuk mengubah data nominal ke data text supaya bisa diproses di document processing.
- Selanjutnya pilih operator proses document from data, jika di klik 2x terdapat subprocess yang terdapat operator untuk pemrosesan data.
- Pertama, subprocess operator tokenize. Operator ini berfungsi untuk memisahkan satu kata dengan kata lainnya, jadi satu kata akan menjadi satu atribut tersendiri. Kemudian tarik garis ke hasil dan jalankan, selanjutnya cek apakah kata-kata tersebut sudah terpecah atau belum.
- Kedua subprocess operator transform cases. Operator ini berfungsi untuk mengubah font. Jadi, semua font data tersebut sama yaitu menggunakan lower case (huruf kecil semua).
- Ketiga, subprocess operator filter stopword(dictionary). Menggunakan operator ini karena operator yang lain bukan Bahasa Indonesia, sedangkan data kita menggunakan Bahasa Indonesia. Operator ini berfungsi untuk menghapus kata-kata yang tidak penting seperti dan, yang, dari, di, adalah. Tidak perlu list satu-satu tinggal download file stopwordsbahasa di youtube.
- Keempat, subprocess operator filter tokens (by length). Operator ini berfungsi untuk menghapus kata-kata yang hanya mengandung sedikit huruf atau kata-kata yang mengandung banyak huruf. Ini dapat diatur di bagian pojok kanan atas yaitu mengatur min char 4 dan max char 25 yang berarti minimal huruf 4 dan maksimal huruf 25
- Selanjutnya memakai operator wordlist to data. Sebelumnya, output dari operator process document from data merupakan wordlist bukan example, sehingga perlu dikonversi ke example supaya berupa data set. Yaitu dengan menggunakan operator wordlist to data.
- Setelah itu, coba untuk memvisualisasikan data. Namun, tidak bisa di visualisasikan karena kebanyakan data. Sehingga perlu operator sort untuk

menyortir data yang kata-katanya dibutuhkan. Kemudian atur pilih atribut total dan ganti dengan descending yaitu mensortir kata dari yang besar ke kecil.

- Selanjutnya, pilih operator filter example range. Operator ini bertujuan untuk memfilter berdasarkan urutan rentang tertentu, output dari sort tarik ke operator ini. Dalam panel parameter first example diisi 1 dan last example diisi 25. Yang artinya visualisasi data (wordcloud) akan ditampilkan 25 kata teratas.
- Setelah operator tersebut dijalankan, pilih bagian visualization. Atur bagian plot 1, untuk plot type diganti wordcloud, value column pilih word.



4. Text labelling

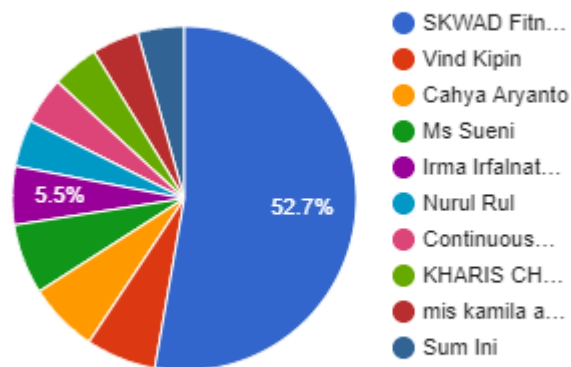
Pada proses ini merupakan lanjutan dari poin C.2 yaitu dengan memberi label positif dan negatif di 300 baris teratas. Tujuan proses ini untuk menganalisis komentar masyarakat tentang pembahasan diet sehat apakah setuju dengan diet

sehat atau tidak setuju dengan diet sehat. Hasil dari text labelling ini digunakan untuk proses selanjutnya yaitu analisis sentiment.

D. Hasil dan Pembahasan

1. Analisis Top User

Berdasarkan hasil temuan yang didapatkan dari netlytic, terdapat 10 aktor yang paling banyak berkontribusi pada topik diet sehat. Berikut 10 aktor yang menjadi top users dalam pembicaraan topik diet sehat, antara lain SKWAD Fitness (52,7%), Vind Kipin (6,6%), cahya aryanto (6,6%), Ms Sueni (6,6%), Irma Ifalnatul (5,5%), Nurul Rul (4,4%), Continuous (4,4%), Kharis Channel (4,4%), Mis Kamila Aziz (4,4%), dan Sum Ini (4,4%). Berikut lampiran top 10 user dari netlytic :



Gambar Top User

Akun SKWAD Fitness merupakan salah satu top users karena akun ini selalu membalas komentar aktor lainnya. Akun ini memposting 1 komentar dan membalas komentar aktor lainnya sebanyak 48 kali. Selain itu, akun Vind Kipin, Cahya Aryanto dan Ms Sueni merupakan top users urutan ke 2,3,4. Tiga akun tersebut sama sama sering membalas komentar aktor lainnya sebanyak 6 kali.

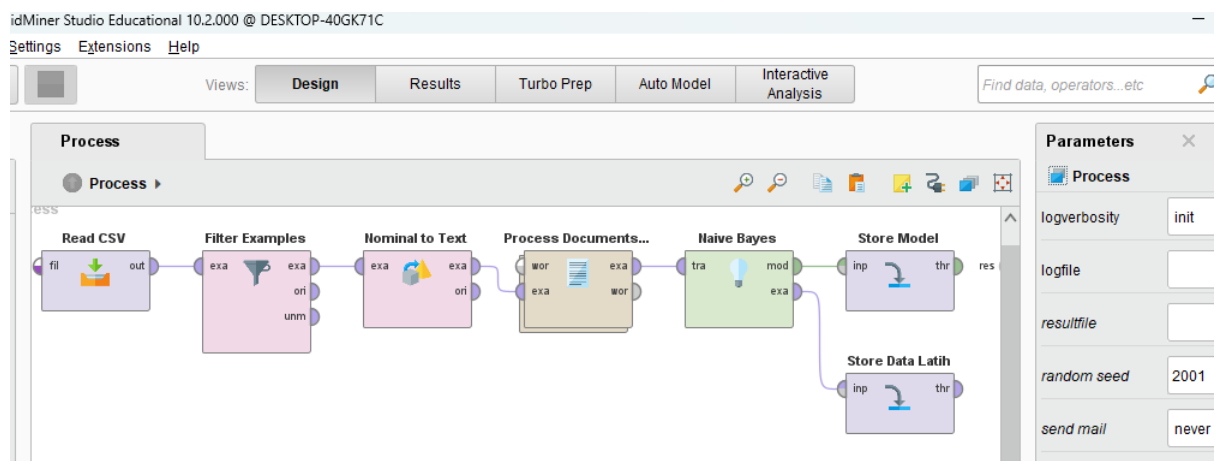
2. Word Cloud

Berikut merupakan hasil word cloud yang ditela di proses melalui rapid miner dan menghasilkan seperti gambar dibawah ini



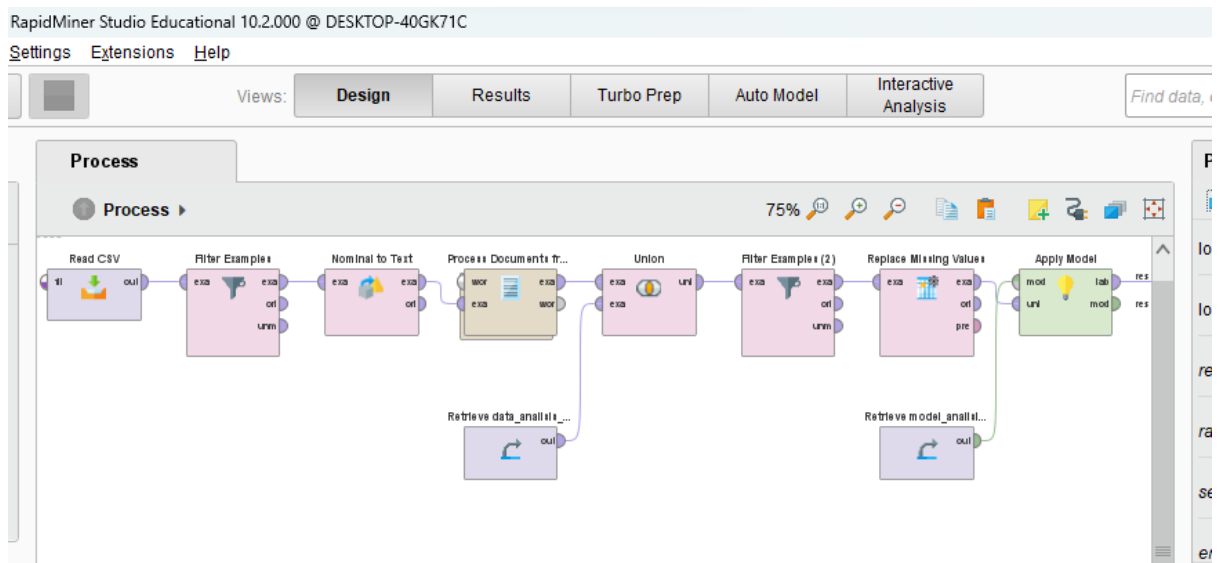
Gambar tersebut merupakan 25 kata teratas yang sering dibicarakan pada komentar youtube dengan topik diet sehat. Kata “diet” merupakan kata teratas yang sering muncul yaitu sebanyak 152 kali. Sedangkan kata turun sering muncul sebanyak 82 kali. Dengan demikian, dapat ditarik kesimpulan bahwa diet banyak dukungan dari masyarakat luas dan para pejuang diet semangat melakukan diet untuk menurunkan berat badan, mengidealkan badan, serta membuat badan menjadi sehat.

3. Sentiment Analysis



Gambar tersebut merupakan proses analisis sentiment menggunakan algoritma Naïve Bayes. Dokumen csv yang digunakan adalah dokumen csv yang sudah di labeli positif dan negatif sebanyak 300 baris. Langkah selanjutnya seperti

dengan proses-proses sebelumnya dan di training dengan operator Naïve Bayes karena untuk pengolahan analisis sentiment ini menggunakan algoritma Naïve Bayes. Setelah di training akan menghasilkan dua data yaitu data model serta data latih. Kemudian 2 data tersebut disimpan karena akan digunakan untuk proses selanjutnya.



Gambar tersebut merupakan langkah terakhir dalam pemrosesan analisis sentiment menggunakan algoritma Naïve Bayes. 4 proses masih sama dengan proses sebelumnya yaitu menggunakan operator read csv yang berisi data yang sudah di labeli 300 baris, operator filter examples, operator nominal to text dan process documents from data. Selanjutnya data latih (data_analisis_sentimen) akan digabung dengan data yang sudah dilabeli 300 baris menggunakan operator union. Setelah berhasil digabungkan, data tersebut akan difilter sentiment is missing dengan operator filter examples yang bertujuan untuk menampilkan data yang sentimennya kosong sehingga data yang berisi sentiment akan dibuang. Selanjutnya, data-data yang kosong akan dibuang menggunakan replace missing values. Setelah berhasil dibuang, data model akan di apply model menggunakan operator apply model, operator retrieve model ditarik garis ke model dan operator replace missing values ditarik ke unlabel.

y/hasil_analisis_sentimen – RapidMiner Studio Educational 10.2.000 @ DESKTOP-40GK71C

is View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Interactive Analysis

ExampleSet (Apply Model)

Open in Turbo Prep Auto Model

Filter (400 / 400 examples): all

Row No.	Sentimen	prediction(S...	confidence(...	confidence(...	confidence(...	confidence(...	text	aamiin	aam
19	?	negatif	0	0	0	1	ngmong kasi...	0	0
20	?	Negatif	0	1	0	0	mkan shariq ...	0	0
21	?	Positif	1	0	0	0	nemu chann...	0	0
22	?	negatif	0	0	0	1	besok bljar di...	0	0
23	?	positif	0	0	1	0	tiktok	0	0
24	?	Positif	1	0	0	0	terberat work ...	0	0
25	?	Positif	1	0	0	0	bismillahirra...	0	0
26	?	positif	0	0	1	0	halo	0	0
27	?	Negatif	0	1	0	0	hebat coba	0	0
28	?	Positif	1	0	0	0	woow bener...	0	0
29	?	Positif	1	0	0	0	memotivasi s...	0	0
30	?	Positif	1	0	0	0	happy banget...	0	0
31	?	positif	0	0	1	0	yesss lanjutin	0	0
32	?	positif	0	0	1	0	girl	0	0

ExampleSet (400 examples, 7 special attributes, 1,582 regular attributes)

Gambar tersebut merupakan hasil prediksi dari olahan analisis sentiment menggunakan algoritma Naïve Bayes di rapid miner. Untuk kalimat positif seperti pada kalimat 21 yaitu “nemu channel bagus banget sebulan turun”, sedangkan kalimat yang sentiment negatif pada kalimat 19 yaitu “ngomong kasih solusi curhat”. Namun, sentiment dari data tersebut tidak semuanya akurat, terdapat beberapa data yang seharusnya sentimen positif namun di olahan tersebut sentiment negatif seperti pada kalimat 27 yaitu “hebat coba”. Pada kalimat 27 tersebut seharusnya mengandung sentiment positif.

E. Kesimpulan

Berdasarkan hasil analisis sentimen terhadap komentar youtube tentang diet sehat yaitu masyarakat sangat mendukung adanya gerakan masyarakat sehat berupa diet sehat. Para masyarakat semangat untuk melakukan diet sehat supaya badannya sehat dan tidak menyebabkan obesitas. Diet sehat bisa dilakukan dengan cara mengonsumsi buah dan



FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO SEMARANG

JL. IMAM BONJOL NO. 207 SEMARANG TELP. 024-3575915, 024-3575916

sayuran secara teratur, sering melakukan olahraga, menghindari makanan yang berminyak, melakukan work out secara teratur, dan mengurangi minuman yang manis maupun bersoda.



**FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO SEMARANG**

Jl. IMAM BONJOL NO. 207 SEMARANG TELP. 024-3575915, 024-3575916

Ketentuan:

1. Judul dituliskan dengan jumlah maksimal 12 kata
2. Bagian pendahuluan terdiri atas tiga paragraf, meliputi:
 - a. Paragraf 1 merupakan penjelasan tentang **fenomena / fakta / latar belakang** yang berkaitan dengan topik yang telah ditentukan
 - b. Paragraf 2 menjelaskan tentang **masalah dan akibat** / efek yang muncul karena masalah
 - c. Paragraf 3 menjelaskan tentang **tujuan** dilakukannya analisis word cloud dan sentiment analysis
3. Dataset hasil dari Netlytic, dataset hasil processing dan dataset labelling, file RapidMiner di upload di Google Drive
4. Algoritma yang digunakan untuk sentiment analysis adalah Naive Bayes

Catatan:

1. File dikumpulkan dalam format **.pdf** dan disubmit melalui **Kulino**
2. Penamaan file adalah **UTS_AMS_NIMxxxxx**. Contoh: **UTS_AMS_02237**
3. Tidak perlu ditambahkan cover halaman atau halaman yang sejenis
4. Dikumpulkan di Kulino paling lambat **Jumat, 3 November 2023 pukul 12.00 WIB**
5. *Plagiarism will not be tolerated*