

Analisis Komentar Pada Twitter Terhadap Lapangan Kerja Dengan Metode Naïve Bayes

Rizky Kurnia Pratama¹, Putry Wahyu Setyaningsih²

Fakultas Teknologi Informasi/Sistem Informasi
Universitas Mercu Buana Yogyakarta
Yogyakarta, Indonesia

e-mail: ¹kurniarizkyy9@gmail.com, ²putryws@mercubuana-yogya.ac.id

Diajukan: 13 Juni 2022; Direvisi: 07 April 2022; Diterima: 09 Mei 2023

Abstrak

Pandemi covid telah memberikan dampak bagi perekonomian suatu negara, terutama Indonesia. Dampak yang terkena imbas pada sektor ekonomi yaitu sektor lapangan kerja. Isu mengenai lapangan kerja pada twitter mulai muncul ketika pada tahun 2020 bulan Desember lalu saat corona mulai masuk ke Indonesia sehingga gelombang PHK banyak terjadi. Masyarakat Indonesia lantas menggunakan media kanal yaitu twitter untuk berkomentar tentang kondisi mereka. Twitter adalah salah satu media kanal internet yang diciptakan untuk jejaring sosial dan sarana ekspresi diri sehingga pengguna yang menggunakan twitter akan mendapat kepuasan tersendiri. Penelitian ini mencoba memberikan perspektif suatu analisis komentar pada twitter terkait lapangan kerja di masa pandemi covid-19. Analisis dilakukan dengan menggunakan aplikasi Orange, proses yang dilakukan melalui tahapan preprocessing, transformation, filtering, tokenizing, dan normalization. Tahapan selanjutnya yaitu pelabelan otomatis dengan metode Vader, klasifikasi dengan metode naive bayes dan pembobotan dengan metode TF-IDF serta kalkulasi dari aplikasi orange data mining yang representasikan dengan hasil extended confusion matrix. Data yang penulis analisis menurut aplikasi orange sebanyak 3929 data tweet dari tanggal 22-30 Agustus 2021 dengan memanfaatkan Web Crawling API Twitter. Hasil penelitian dari aplikasi orange data mining memperlihatkan akurasi dengan Naive Bayes mendekati angka sempurna yaitu 99% dengan jumlah sentimen di masyarakat sebanyak 25% positif, 11% negatif dan 64% netral.

Kata kunci: Analisis Sentimen, Orange, Preprocessing, Twitter.

Abstract

The Covid pandemic has had an impact on the economy of a country, especially Indonesia. The impact on the economic sector is the employment sector. The issue of employment on Twitter began to emerge when in 2020 last December when the corona began to enter Indonesia so that a wave of layoffs occurred. Indonesian people then use the media channel, namely twitter to comment about their condition. Twitter is one of the internet media channels created for social networking and a means of self-expression so that users who use Twitter will get their own satisfaction. This study tries to provide a perspective on an analysis of comments on Twitter related to employment during the COVID-19 pandemic. The analysis is carried out using the Orange application, the process is carried out through the stages of preprocessing, transformation, filtering, tokenizing, and normalization. The next stage is automatic labeling using the Vader method, classification using the naive Bayes method and weighting using the TF-IDF method and calculations from the orange data mining application which are represented by the results of the extended confusion matrix. The data that the author analyzes according to the orange application is 3929 tweet data from 22-30 August 2021 by utilizing the Twitter Web Crawling API. The results of the orange data mining application show that the accuracy with Naive Bayes is close to the perfect number, which is 99% with the number of sentiments in the community as much as 25% positive, 11% negative and 64% neutral.

Keywords: Sentiment Analysis, Orange, Preprocessing, Twitter.

1. Pendahuluan

Arus kemajuan teknologi tidak bisa terelakan dampak yang ditimbulkan, hal terlihat perubahan struktur kehidupan masyarakat yang hari ini bergantung kepada gawai dan kemudahan yang ditawarkan sangat membantu tugas dari manusia itu sendiri. Internet merupakan wasilah dalam mencari informasi, segala bentuk data yang diperlukan ada disini, hanya bermodal kouta dan smartphone. Informasi yang

potensial di era globalisasi saat ini bersumber pada media sosial. Twitter adalah salah satu media sosial yang sangat terkenal di era internet saat ini, suatu platform untuk bercakap-cakap dan aktualisasi diri. Indonesia termasuk salah satu bagian dari pertumbuhan pengguna aktif paling besar menurut laporan finansial twitter kuartal ke-3 tahun 2019 [1]-[5]

Twitter dapat digunakan untuk beranekaragam penelitian yang dapat dilakukan, salah satunya yaitu sentimen analisis. Sentimen analisis merupakan bagian dari teknik opinion mining. Proses analisis sentiment dilakukan untuk melihat persepsi terhadap suatu problematika atau dapat digunakan sebagai identifikasi kecenderungan market. Munculnya pengaruh dan manfaat dari analisis sentimen menyebabkan lonjakan penelitian mengenai sentiment analisis secara eksponensial yang sangat beranekaragam, bahkan berkisar 20-30 perusahaan di amerika sendiri memfokuskan penelitian pada layanan sentimen analisis (Noviah Dwi Putranti, January 2014) [6]-[10].

Era milenial yang serba mudah untuk mengolah informasi, terutama yang berkaitan dengan informasi lowongan kerja telah menjadi kebutuhan yang melonjak di masyarakat yang mana akhirnya perusahaan membagikan informasi mengenai lowongan kerja di berbagai platform salah satunya adalah twitter. berbagai macam media informasi yang digunakan untuk berbagi informasi seputar lowongan kerja hal ini dapat memudahkan masyarakat umum dengan berbagai kelas pendidikan mulai dari SMA atau sederajat hingga lulusan sarjana (S1) untuk mencari lowongan kerja. Contoh studi kasus : penelitian dari Warda Ningrum Yang mengangkat tema Analisis Sentimen Pada Lowongan Pekerjaan Dengan Algoritma Naïve Bayes, dimana hasil pengujian dengan metode Naïve Bayes, Support Vector Machine, Maximum Entropy dengan hasil kinerja sistem mendapatkan nilai precision sebesar 60%, recall sebesar 60%, F-Measure sebesar 60% dan Accuracy sebesar 30%. Dari total 10 data lowongan pekerjaan yang telah dilakukan proses dan klasifikasi oleh sistem maka hasilnya 4 data yang termasuk ke dalam teknik, 1 data termasuk ke dalam ekonomi dan 2 data yang termasuk ke dalam sarjana. [11].

Penelitian berikutnya tentang mengenai analisis kecemasan *freshgraduate* dalam menghadapi dunia kerja, tujuan penelitian ini mengetahui kecemasan mahasiswa dalam menghadapi dunia kerja terutama mahasiswa *freshgraduate* di masa pandemi *covid-19*. Metode penelitiannya adalah metode survei dengan alat skala kecemasan. Subjek penelitiannya adalah mahasiswa prodi psikologi FK ULM *freshgraduate* berjumlah 57 orang. Dari hasil penelitian ini kategori terbanyak sekitar yaitu 43.4% menunjukkan memiliki kecemasan yang tinggi dan 13.2% adalah kecemasan pada level rendah. Terdapat juga sebuah teori dari penelitian ini yaitu hubungan kecemasan dengan IPK dan lama studi itu tidak ada (Zwagery, 2020). [12]

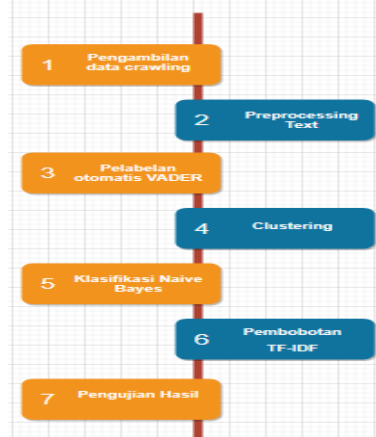
Penelitian yang lain membahas tentang sentimen analisis mengenai pembelajaran daring pada masa pandemi. Metode yang dipakai adalah Naïve Bayes karena metode ini memiliki akurasi yang baik untuk komputasi data dan analisis. Tools yang dipakai adalah Rapidminer dan dilanjutkan dengan tools Twitter publik Streaming Drone Emprit Academy. Secara manual menghasilkan 12906 data tweet yang berkaitan dengan kata kunci "pembelajaran daring", "kuliah", "belajar", "daring" dan tagar #BelajarDariRumah yang semua kata kunci itu difilter menjadi kata kunci "daring" dan "rumah" pada tweet bahasa Indonesia. Hasil dari penelitian ini adalah sentimen terbanyak jatuh kepada sentimen negatif berkisar 69%, sentimen positif 30% dan sentimen negatif 1%. Kata kunci frekuensi tinggi ada pada kata "stres" dan "malas" (Samsir, 2021). [13]

Dari penelitian sebelumnya maka penulis akan melakukan sentimen analisis pada media sosial di twitter tentang komentar masyarakat atas sentimen lapangan pekerjaan yang ada di twitter dengan menggunakan Aplikasi Orange Data Mining dengan metode *naïve bayes*. Alasan penulis memilih metode *naïve bayes* yaitu kelebihan *naïve bayes* yang sederhana dalam algoritma, kalkulasi perhitungan yang cepat, serta memiliki akurasi yang tinggi. Menurut Husin hasil akurasi algoritma NBC dengan algoritma C4.5 masih kurang baik nilai akurasi. hal ini karena dalam algoritma C4.5 semua atribut terfilter menjadi bagian kumpulan bagian lebih kecil, namun model akan terbentuk menjadi lebih rumit dan sulit untuk dipahami jika data berukuran besar dengan atribut yang banyak. sehingga perlu pemotongan yang dapat mengurangi akurasi. sedangkan *naïve bayes* cenderung ditetapkan pada kumpulan data yang besar dan dapat mengampukan data yang kurang lengkap (missing value) serta kuat terhadap atribut yang tidak bermakna dan noise pada data (Husin Muhamad, 2017). [14]

Untuk proses yang dilakukan pertama memanfaatkan tools twitter API sebagai pengambilan sample secara realtime, jumlah yang kami ambil ada 3929 data. Kedua data twitter tadi akan masuk tahap preprocessing dan cleaning data untuk filterisasi dan klasifikasi berdasarkan opini yang ada, selanjutnya Pengolahan data menurut metode yang sudah disebutkan. Proses terakhir adalah Ekstraksi fitur dimana dari opini tadi dikelompokkan menjadi 3 (tiga) yaitu sentimen positif, negatif dan netral dalam mencari pekerjaan dalam kondisi pandemi covid-19.

2. Metode Penelitian

Berikut ini proses tahapan dalam mengerjakan penelitian pada gambar berikut [14]:

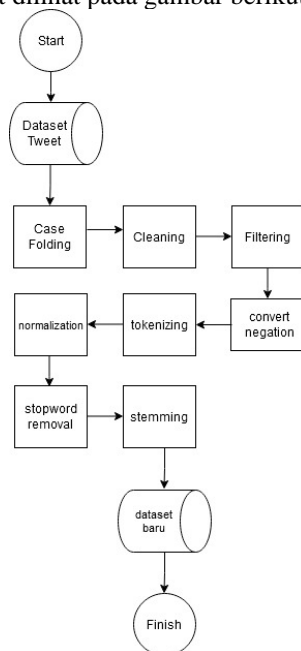


Gambar 1. Tahapan Penelitian

Tahapan dalam sentimen yaitu proses preprocessing. Pada proses preprocessing, tweet yang telah dihimpun akan melalui tahapan pemrosesan teks yang terdiri dari enam tahap yaitu:

- a. Case folding
- b. Cleaning
- c. Filtering
- d. Tokenizing
- e. Transformation
- f. Normalization

Fungsi tahapan ini yaitu agar data yang diperoleh akan cepat terstruktur dan mudah untuk melakukan proses analisis data. Proses yang kedua adalah preprocessing text, dimana proses ini melalui tahapan demi tahapan dilalui yang berguna untuk membersihkan data yang tidak bernilai. Data yang memiliki noise seperti tag html, emotikon, dan tanda baca yang tidak berguna. Untuk penerapan case folding dalam orange data mining itu sudah tersedia dalam aplikasi orange sehingga secara otomatis aplikasi orange akan mengubahnya. Tahapan preprocessing merupakan hal yang penting untuk menuju proses selanjutnya, yaitu menghapus atribut yang tidak berguna dalam tahapan klasifikasi. Data yang dimasukkan ke dalam tahap ini adalah data yang masih mentah dan belum ada proses cleaning, sehingga ketika telah dilakukan proses ini maka dapat mempermudah tahapan klasifikasi. proses preprocessing terdiri dari beberapa proses yang dapat dilihat pada gambar berikut [15]-[20]:



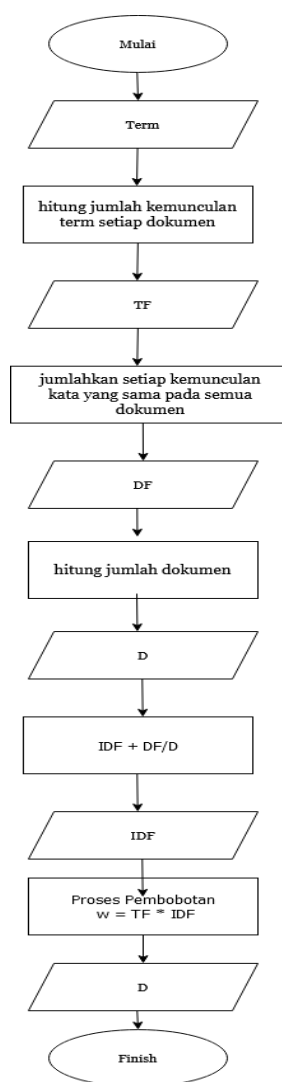
Gambar 2. Proses Preprocessing

Pembobotan term yaitu proses dimana term berubah sebagai bentuk matriks yang mewakili dari sekumpulan dokumen yang telah dilakukan ekstraksi fitur, pembobotan term dilakukan untuk mengkalkulasi frekuensi kemunculan suatu kata pada setiap dokumen. Metode yang digunakan adalah TF-IDF sebagai proses metode pembobotan, dan akan dilakukan kalkulasi pembobotan pada setiap term/kata berdasarkan nilai dalam sekelompok dokumen.

$$W_{x,n} = TfIdf_{x,n} + \alpha w_{x,n-1}$$

Keterangan:

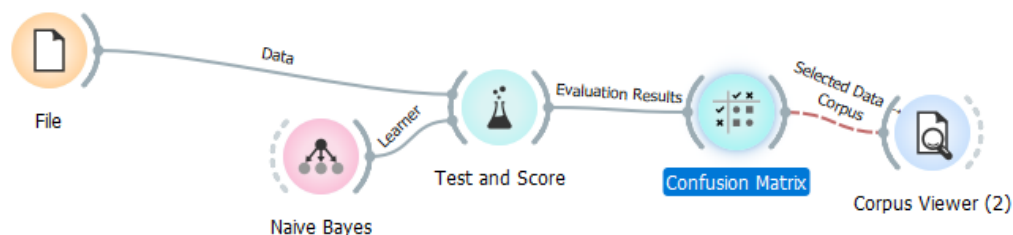
- $W_{x,n}$ = Bobot kata x pada interval ke-n
- $TfIdf_{x,n}$ = Nilai Tf-Idf kata x pada interval ke-n
- α = *Discounted comulative factor*



Gambar 3. Proses Pembobotan

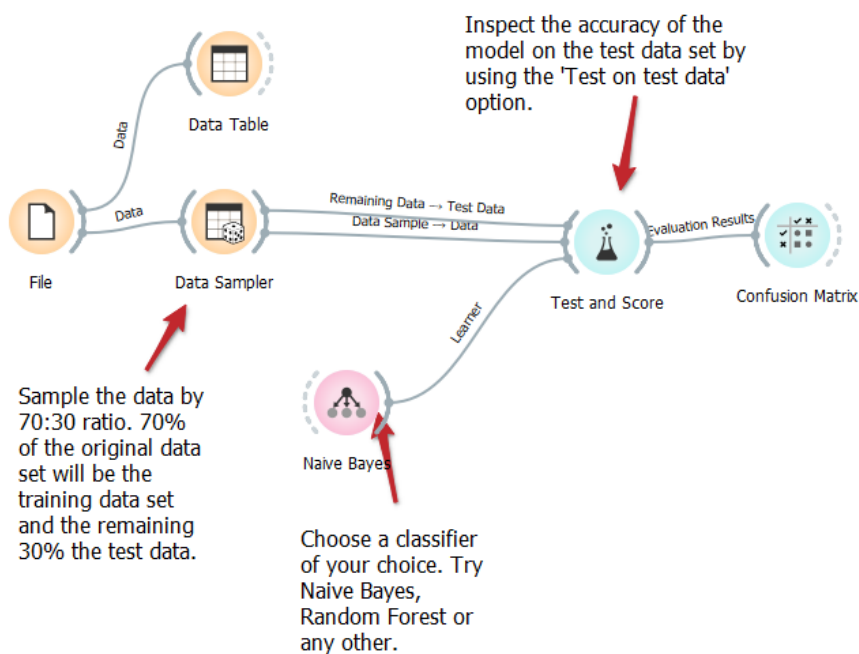
Data yang telah diproses pada prosedur pre-processing akan lanjut ke dalam proses pembobotan dan akan diberikan algoritma Naive Bayes. Proses pembelajaran akan dilakukan dengan data yaitu data latih. data latih adalah data yang sudah mempunyai label positif, negatif dan netral. sebagai pembelajaran untuk memahami label dari suatu dokumen yang dimaksud. jumlah data keseluruhan adalah 3929 data.

Untuk proses dari yang akan diterapkan saya akan melampirkan gambar flow dalam aplikasi orange data mining sebagai gambaran proses yang dilakukan untuk mencapai tujuan.



Gambar 4. Metode pada orange

Pada proses data yang sudah melewati proses pre-processing dengan baik akan dilanjutkan dengan pembobotan dan diberikan algoritma Naive Bayes. Proses pembelajaran dilakukan dengan data yaitu data latih. Data latih adalah data yang sudah memiliki label positif dan negatif, sebagai pembelajaran untuk mengetahui label dari suatu dokumen yang dimaksud. Jumlah keseluruhan data adalah 3929 data dengan pembagian data latih dan data uji. Perbandingan antara data latih dan data uji adalah sebanyak 30% untuk data latih dan 70% untuk data uji. Data latih dan data uji akan sama-sama diimplementasikan menggunakan algoritma Naive Bayes.



Gambar 5. Flow Orange Naive Bayes

Pelabelan menggunakan analisis sentimen Valance Aware Dictionary and Sentimet Reasoner (VADER).dengan pelabelan otomatis dengan bobot nilai sentimen sebagai berikut pada tabel di bawah berikut:

Tabel 1. Pelabelan

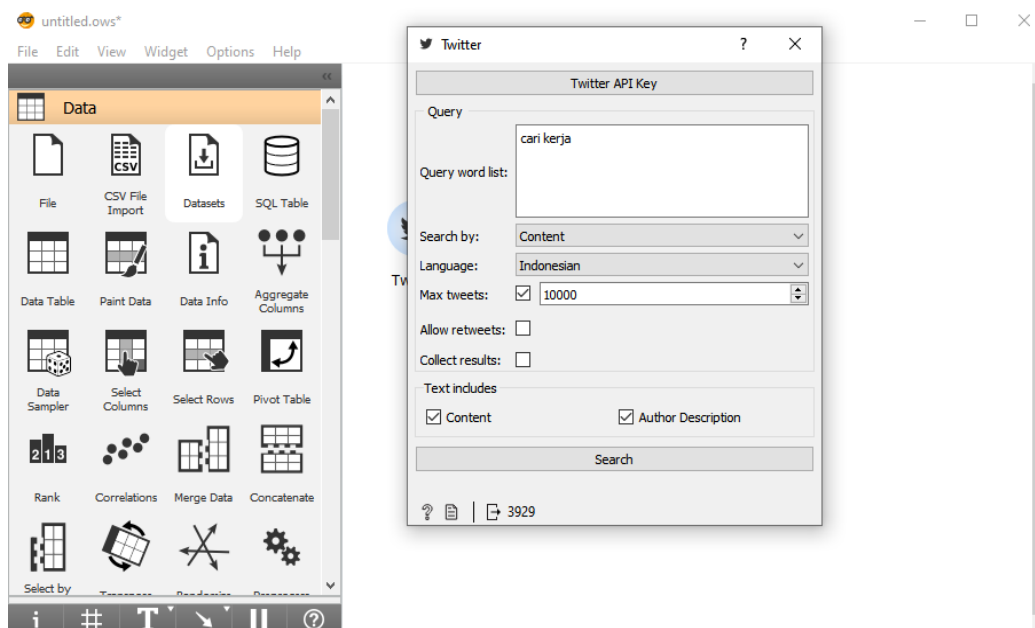
No	Sentimen	Score	Keterangan
1	Positif	≥ 0.05	Menunjukkan kata lebih mengarah pada sentimen positif
2	Netral	$> -0.05 - < 0.05$	Menunjukkan bobot kata mengarah pada sentimen netral
3	Negatif	≤ -0.05	Menunjukkan bobot kata mengarah pada sentimen negatif

3. Hasil dan Pembahasan

Bagian ini membahas hasil dari penelitian dan pada waktu yang sama juga memberikan pembahasan dan yang komprehensif. Hasil penelitian dapat disajikan menggunakan gambar, grafik, tabel, dan lainnya yang membuat pembaca dapat memahami hasil penelitian dengan mudah. Pembahasan dapat dibuat dengan menggunakan beberapa sub-bab.

3.1. Proses Data Crawling

Penelitian ini akan berfokus pada judul yang dibawa penulis yaitu komentar pada twitter mengenai lapangan pekerjaan pada masa covid-19. Tweet diambil dari aplikasi orange data mining sejumlah 10000 data tweet, namun karena dari aplikasi ditemukan hanya bisa mengambil data dari twitter sebanyak 3929 tweet maka penulis memakai data ini untuk analisa lebih lanjut. Dimana akan dianalisa menjadi cluster serta rekapitulasi sehingga melahirkan kata yang dominan timbul dari status yang disebut sebagai corpus. Widget corpus yaitu himpunan media digital yang memberi deksripsi jumlah baris kalimat, serta memastikan mana fitur untuk kepentingan riset analisis data. Query yang dipakai penulis adalah “*cari kerja*”.



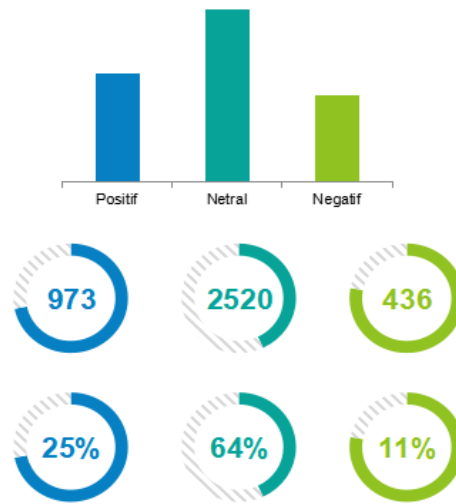
Gambar 6. Data Crawling

3.2. Analisis dan Pembahasan

Dalam penelitian ini penulis akan melakukan tahapan dari hasil proses penelitian dengan menggunakan fitur dari perangkat lunak Orange Data Mining. Ketika semua tahapan telah dilakukan maka akan mempermudah kelanjutan penelitian dan terlihat hasil dari penelitian ini.

3.2.1. Komentar Sentimen Masyarakat

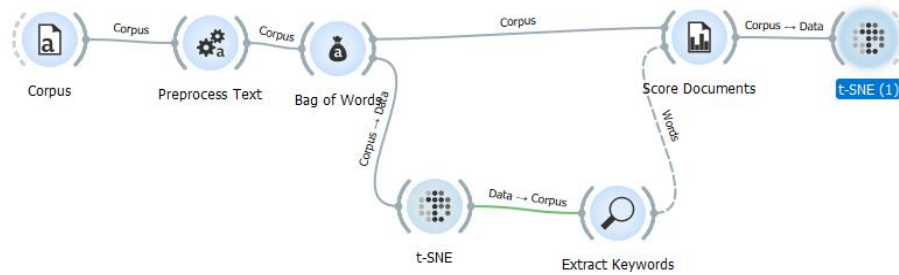
Pada gambar ini menunjukkan data persentase dimana menunjukkan sentimen positif berjumlah 973 record, netral 2520 record dan negatif 436 record.



Gambar 7. Sentimen Masyarakat

3.2.2. Pembobotan

Dalam pembobotan TF-IDF pada orange data mining untuk flow pada orange dapat dilihat pada gambar berikut:



Gambar 8. TF-IDF

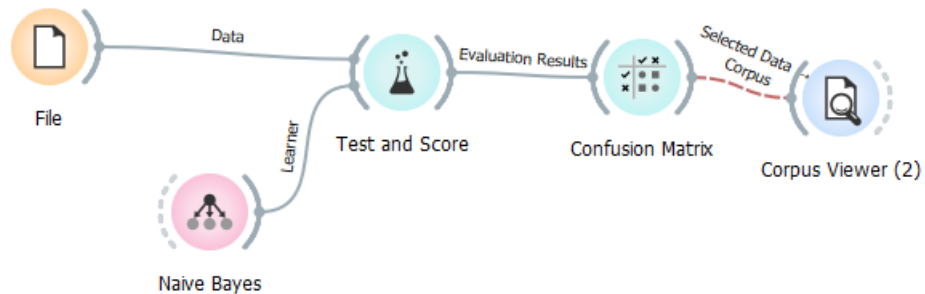
Dari hasil pemrosesan orange data mining dapat kita lihat dengan filter 100 kata yang sering muncul dapat dihasilkan pada gambar berikut:

Word	TF-IDF
kerja	0.324
cari	0.304
susah	0.098
semangat	0.083
bikin	0.075
pandemi	0.074
cv	0.070
10k	0.067
kuliah	0.026
lulu	0.023
duit	0.022
anak	0.019
rumah	0.019
semoqa	0.019
capek	0.016
uang	0.015
kerjaan	0.015
qaji	0.015
resign	0.012

Gambar 9. Hasil pembobotan orange

3.2.3. Klasifikasi

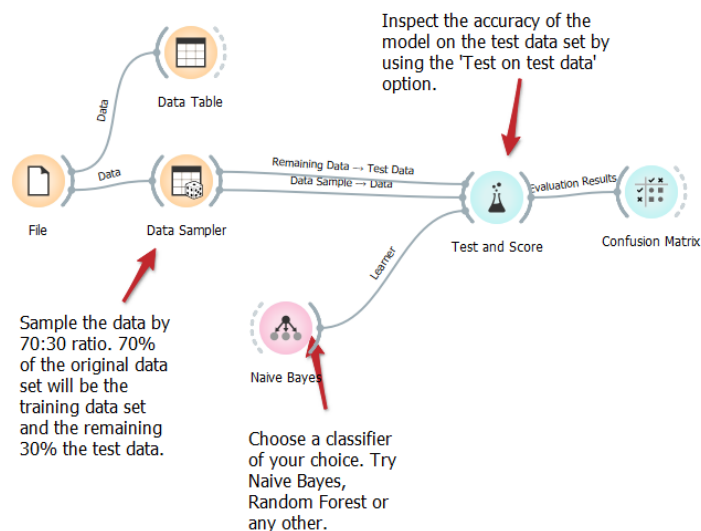
Metode yang digunakan adalah naive bayes dan untuk flow pada orange data mining bisa dilihat dari gambar di bawah berikut:



Gambar 10. Flow Klasifikasi

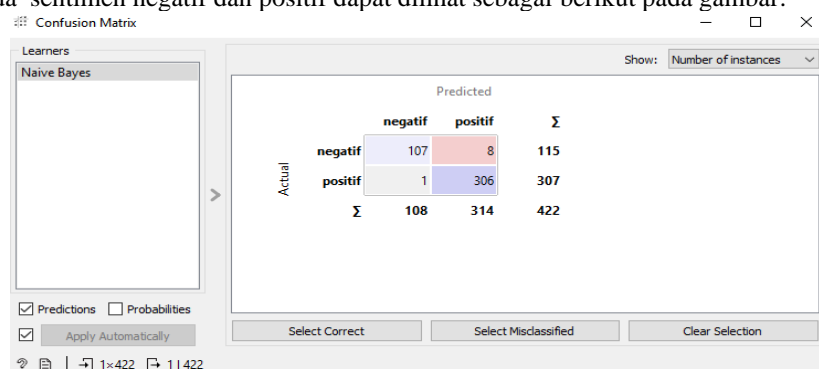
3.2.4. Pengujian Hasil dan Akurasi

Pada orange data mining untuk pengujian dan akurasi bisa dilihat pada flow gambar sebagai berikut:

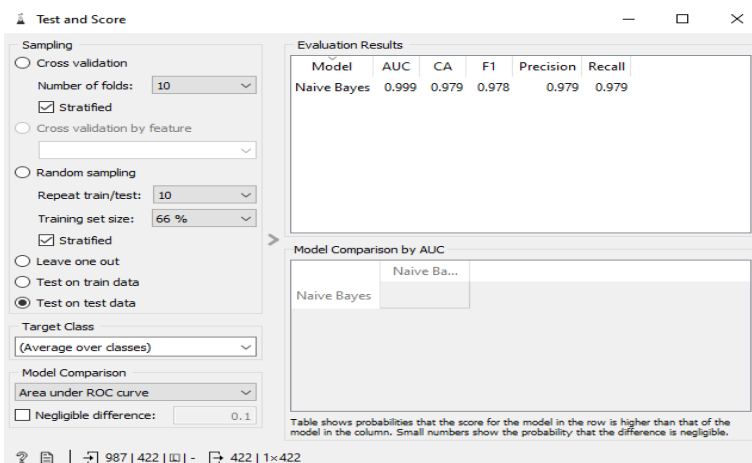


Gambar 11. Flow widget Pengujian Akurasi

Tampilan dari orange data mining dengan widget confusion matrix dan hasil evaluating result yang berfokus pada sentimen negatif dan positif dapat dilihat sebagai berikut pada gambar:



Gambar 12. Confusion Matrix



Gambar 13. Evaluation Results Confution Matrix

Karena penulis memakai rasio 70:30 maka hasil dari data keseluruhan menjadi 1178 dari total 3929 data. Untuk mengukur pengujian hasil dengan metode naive bayes, dengan masing – masing perbandingan 70 % dan 30 % yang menghasilkan nilai akurasi 99% dengan rumus sebagai berikut:

$$Recall = \frac{TP}{TP + FN} = \frac{107}{107 + 1} = 0,99$$

Setelah itu akan ada proses perhitungan nilai precision dengan menggunakan rumus pada persamaan berikut:

$$Precision = \frac{TP}{TP + FP} = \frac{107}{107 + 8} = 0,92$$

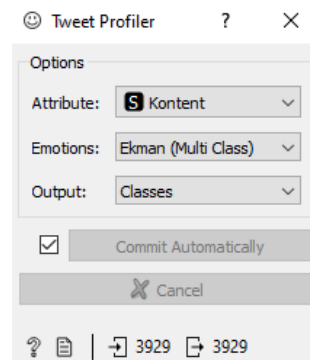
Terakhir nilai akurasi didapat melalui perhitungan menggunakan rumus pada persamaan berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{107 + 306}{107 + 306 + 8 + 1} = 0,9786$$

Dapat disimpulkan bahwa dari hasil perhitungan dengan menggunakan rumus, berdasarkan hasil recall sebanyak 99% data terklasifikasi dengan benar, yang mana data review yang berpeluang bernilai negatif adalah benar bernilai negatif dan sebaliknya. Kemudian, nilai precision berfungsi untuk mengetahui rasio prediksi benar negatif dibandingkan dengan semua hasil yang berprediksi negatif, serta nilai hasil perhitungan rumus yaitu sebesar 92%. Lalu nilai akurasi yang didapat dari perhitungan rumus sebesar 97% yang berarti bahwa nilai tersebut menunjukkan data preview diprediksi dengan betul pada semua class sentimen (class negatif dan class positif).

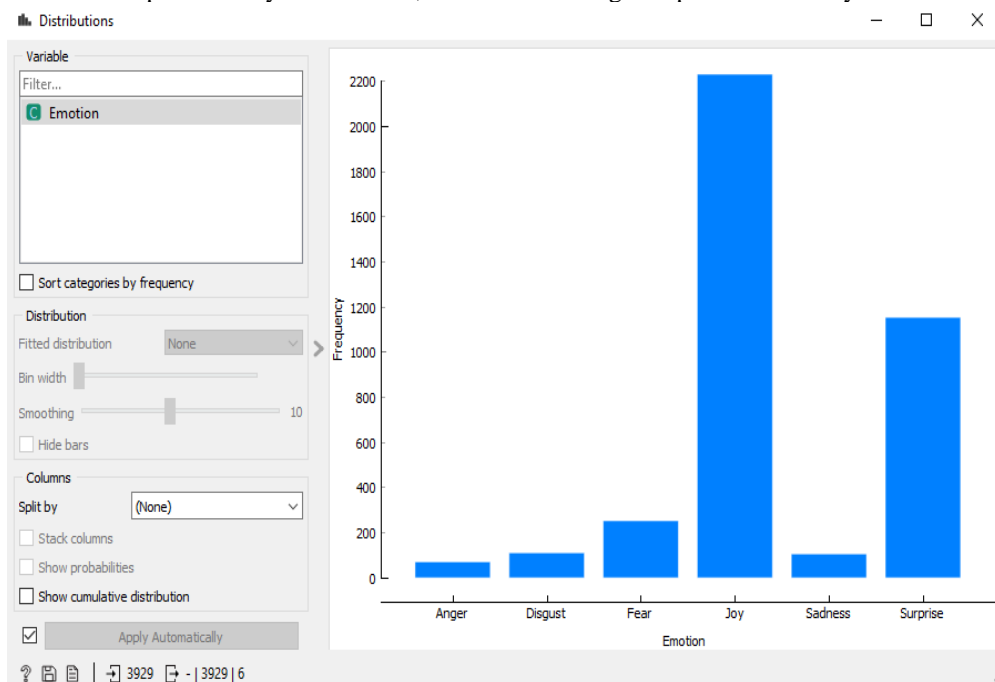
3.2.5. Tweet Profiler dan distribusi

Tweet Profiler mengambil data tentang sentimen dari server untuk setiap tweet (ataupun dokumen) yang diberikan. Widget mengirimkan informasi ke server, tempat model menghitung probabilitas ataupun skor emosi. Widget menunjang 3 klasifikasi emosi, ialah Ekman, Plutchik serta Profile of Mood States(POMS). Klasifikasi emosi terdapat 3 kategori yaitu Ekman, Plutchik ataupun Profile of Mood States. Klasifikasi kelas jamak akan menghasilkan satu emosi yang sangat bisa jadi per dokumen, sedangkan banyak label akan menciptakan nilai dalam kolom untuk tiap emosi (Wiguna, 2021). Penelitian ini akan menggunakan atribut konten untuk analisis, klasifikasi emosi Ekman dengan opsi multi-kelas dan memilih untuk mengamati variabel Emosi yang telah dikelompokkan dengan Orange Data Mining. Pada penelitian ini menggunakan data 3929 tweets mengenai lapangan pekerjaan. Data yang telah di crawling menggunakan widget dari Orange Data Mining dengan Corpus dan dihubungkan ke Tweet Profiler.



Gambar 14. Tweet Profiler

Widget Distribusi menampilkan distribusi nilai atribut diskrit atau kontinu. Jika data berisi variabel kelas, distribusi dapat dikondisikan di kelas. Setelah melakukan tweet profiler pada widget Langkah selanjutnya menghubungkan corpus pada distribution. Hasil akan terlihat 6 bentuk emosi dari data twitter yang telah di input. Pada penelitian ini widget menampilkan emosi para pengguna twitter dengan query word list "cari kerja". Dari hasil 6 emosi ini data menunjukkan bahwa respon dari twitter per tanggal 22 sampai 30 Agustus 2021 menghasilkan paling tinggi frekuensinya yaitu joy sebanyak 2233 data, diurutan kedua adalah surprise sebanyak 1155 data, dan diurutan ketiga ekspresi fear sebanyak 253 data.



Gambar 15. Distribution

4. Kesimpulan

Dari penelitian yang telah dilakukan bahwa analisis sentimen dapat digunakan untuk mengetahui sentimen masyarakat terhadap isu lapangan pekerjaan di tengah pandemik Covid-19. Setelah dilakukan analisis sentimen, terlihat hasil komentar sentimen yang terkait dengan lapangan pekerjaan di masa pandemi. Berdasarkan hasil uraian analisis yang dilakukan diatas, dapat ditarik kesimpulan yaitu:

1. Dimasa pandemi sentimen pada penelitian ini menghasilkan komentar twitter yang positif sebanyak 25% sedangkan negatif sebanyak 11% dan netral sebanyak 64% dari hasil pelabelan otomatis *Vader*.
2. Hasil akurasi dengan melihat extended confusion matrix sebesar 98% setelah diproses dengan orange data mining.

3. Hasil proses *data mining orange Tweet profiler emosi* dengan metode *Ekman* menghasilkan paling tinggi frekuensinya yaitu *joy* sebanyak 2233 data, diurutkan kedua adalah *surprise* sebanyak 1155 data, dan diurutkan ketiga ekspresi *fear* sebanyak 253 data.

Daftar Pustaka

- [1] Anief, S. (2020). Analisis Sentimen Untuk Respon Masyarakat Terhadap Universitas (Studi Kasus : Universitas Mercu Buana Yogyakarta) Sentiment Analysis for Public Response to the University (Case Study : Universitas Mercu Buana Yogyakarta), 1-7.
- [2] Arianto, B. (2021). Dampak Pandemi COVID-19 terhadap Perekonomian Dunia, 212-224.
- [3] Arifidin, S. (2016). Pembangunan Aplikasi Rekomendasi Berita Skripsi, 114.
- [4] Ariyanti, D. (2020). Teks Mining untuk Klasifikasi Keluhan Masyarakat Pada Pemkot Probolinggo Menggunakan Algoritma Naïve Bayes Jurnal IKRA-ITH Informatika, 125-132.
- [5] Harijiatno, S. D. (2019). Analisis Sentimen Pada Twitter Menggunakan Multinomial Naive Bayes Skripsi, 121.
- [6] Komputer, J. T. (2020). Naive Bayes classifier optimization for text classification on e-government using particle swarm optimization Jurnal Teknologi dan Sistem Komputer, 21-26.
- [7] KURNIAWAN, R. (2020). Analisis Sentimen Masyarakat Terhadap Virus Corona Berdasarkan Opini Dari Twitter Berbasis Web Scraper, 67.
- [8] Lengkong, N. C. (2021). Analisis Sentimen Penerapan Psbb Di Dki Jakarta Dan Dampaknya Terhadap Pergerakan Ihsg, 20.
- [9] MAULIZON, O. S. (2018). Klasifikasi Sentimen Masyarakat Terhadap Skripsi, 65.
- [10] Nadya, I. (Agustus 2017). Penerapan Data Mining untuk Menentukan Besar. JUI SI, Vol. 03, No. 02.
- [11] Nursowfa, R. F. (2020). Penanganan Pelayanan Kesehatan Di Masa Pandemi Covid-19 Dalam Perspektif Hukum Kesehatan, 17.
- [12] Zwagery, R. V. (2020). Kecemasan menghadapi dunia kerja pada mahasiswa fresh graduate pada masa pandemi Covid 19, 10-14.
- [13] Samsir. (2021). Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes Jurnal Media Informatika Budidarma, 157-163.
- [14] Husin Muhamad, C. A. (2017). Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris. Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK), Vol. 4, No. 3, hlm. 180-184.
- [15] Rahmat. (2020). Text Mining Judul Skripsi Menggunakan K-Means Clustering (Studi Kasus Dijurusan Teknik Informatika UIN Alauddin Makassar, 70.
- [16] Rahmawati, L. (2016). Analisa Clustering Menggunakan Metode K-Means Dan Hierarchical Clustering (Studi Kasus : Dokumen Skripsi Jurusan Kimia, Fmipa, Universitas Sebelas Maret) Jurnal Teknologi & Informasi ITSmart, 66.
- [17] Rahmawati, L. (2016). Analisa Clustering Menggunakan Metode K-Means Dan Hierarchical Clustering (Studi Kasus : Dokumen Skripsi Jurusan Kimia, Fmipa, Universitas Sebelas Maret) Jurnal Teknologi & Informasi ITSmart, 66.
- [18] Syarifuddin, M. (2020). Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode Naïve Bayes Dan Knn, 23-28.
- [19] Wiguna, R. A. (2021). Analisis Text Clustering Masyarakat Di Twitter Mengenai Omnibus Law Menggunakan Orange Data Mining Journal of Information Systems and Informatics, 1-12.
- [20] Akhmad Muzakki, A. W. (2 Desember 2021). Sentimen Analisis Masyarakat Di Twitter Terhadap Pilkada 2020 di Tengah Pandemi Covid-19 dengan Metode Naive Bayes Classifier. Jurnal Teknik Informatika (JUTIF) Vol. 2, No. 2, Desember 2021, hlm. 101-107.