

Implementasi Algoritma *Random Forest* dan Model *Bag of Words* Dalam Analisis Sentimen Mengenai E-Materai

Fenilinas Adi Artanto

Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Muhammadiyah Pekajangan Pekalongan,
Kabupaten Pekalongan, Indonesia

Email: fenilinasadi@gmail.com

ABSTRACT

For CPNS registration in 2024, a digital stamp purchase method will be implemented via the website <https://meterai-elektronik.com> as a requirement for complete administration. However, the page that provides e-stamps experienced an error because many visitors entered and tried to buy e-stamps at the same time. This disruption has resulted in several people having opinions about e-stamps. To find out people's sentiments about e-stamps after this incident, sentiment analysis was carried out. In analyzing these sentiments, the *Random Forest* algorithm is used with the *Bag of Words* model, where *Random Forest* is a development method of *Classification and Regression Trees (CART)* which is said to be more precise in predicting and the *Bag of Words* model is widely used with good results for predicting language modeling and classification. documents because *Bag of Words* is simple and flexible. In sentiment analysis, 86% of neutral opinions were produced, 9% of negative opinions and 5% of positive opinions. Then the *Random Forest* algorithm with the *Bag of Words* model got an accuracy value of 70.1%, precision of 50.7%, recall of 54.1%, and F1-Score of 47.6%.

Keywords: *Bag of Words, E-Materai, Random Forest, Sentiment Analysis, Text Mining.*

ABSTRAK

Dalam pendaftaran CPNS pada tahun 2024 diberlakukan metode pembelian materai digital melalui website <https://meterai-elektronik.com> sebagai syarat kelengkapan admistrasi. Tetapi halaman tersebut yang menjadi penyedia e-materai mengalami error karena banyaknya pengunjung yang masuk dan berusaha membeli e-materai di saat yang sama. Dengan adanya gangguan tersebut mengakibatkan beberapa masyarakat beropini tentang e-materai. Untuk mengetahui sentiment masyarakat tentang e-materai setelah kejadian tersebut maka dilakukan analisis sentimen. Dalam menganalisis sentimen tersebut digunakan algoritma *Random Forest* dengan model *Bag of Words* dimana *Random Forest* merupakan metode pengembangan dari *Classification and Regression Trees (CART)* yang dikatakan lebih presisi dalam memprediksi dan model *Bag of Words* banyak digunakan dengan hasil yang baik untuk memprediksi pemodelan Bahasa dan klasifikasi dokumen karena *Bag of Words* sederhana dan fleksibel. Dalam analisis sentimen dihasilkan 86% opini netral, 9% opini negatif dan 5% opini positif. Lalu algoritma *Random Forest* dengan model *Bag of Words* mendapatkan nilai *Accuracy* sebesar 70.1%, *precision* sebesar 50.7%, *recall* 54.1%, dan *F1-Score* 47.6%.

Kata Kunci: *Bag of Words, E-Materai, Random Forest, Analisis Sentimen, Text Mining.*

1. Pendahuluan

Dalam perkembangan Teknologi saat ini telah memberikan banyak perubahan, terutama pada perkembangan digitalisasi. Seperti dalam hanya materai. Bea Materai merupakan pengenaan pajak atas dokumen dan memberikan kekuatan yang sempurna, dalam artian lainnya adalah apabila sebuah akta telah dibayarkan bea materainya maka akta tersebut terhindar dari sanksi administrasi yang diatur dalam Undang-Undang bea Materai [1]. Tetapi budaya atau anggapan yang beredar adalah jika suatu dokumen atau surat perjanjian tanpa

materai dianggap tidak sah, namun sah atau tidaknya suatu perjanjian tidak ditentukan oleh pelunasan bea materai tetapi ditentukan oleh pasal 1320 Kitab Undang-Undang Hukum Perdata [2].

Berdasarkan ketentuan pasal 5 ayat 1 Undang-Undang No. 19 Tahun 2016 tentang perubahan atas UU No. 11 Tahun 2008 tentang informasi dan Transaksi Elektronik, disebutkan bahwa informasi Elektronik dan/atau Dokumen Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah. Tetapi sekarang bukan hanya dokumen fisik saja yang dikenakan bea

materai dengan materai tempel tetapi dokumen elektronik atau dokumen non fisik juga dikenakan bea materai yaitu dengan bea materai elektronik atau yang disebut e-materai [3].

Saat ini hampir semua dokumen sudah beralih menjadi dokumen digital, sehingga diperlukan e-materai dalam melakukan pembayaran pajak pada dokumen tersebut. Tetapi pada Agustus 2024 terjadi fenomena yang menjadi banyak perbincangan tentang e-materai. Di saat itu sedang dibuka pendaftaran CPNS dan dokumen dalam syarat CPNS harus di bubuhi oleh e-materai. Dalam proses pembelian e-materai saat pendaftaran CPNS 2024 menjadi viral, karena website <https://materai-elektronik.com> yang menjadi halaman untuk membeli e-materai mengalami error dengan menampilkan “error1200” [4]. Dengan terkendalanya sistem pembelian e-materai banyak menyebabkan banyak sentiment yang beredar dimasyarakat tentang e-materai. Untuk dapat menganalisa sentiment Masyarakat tentang e-materai dapat digunakan metode analisis sentiment.

Analisis sentiment merupakan sebuah cabang dari *text mining* yang dapat digunakan untuk memproses data yang berupa *text* untuk dapat di klasifikasikan maupun di klusterkan [5]. Dalam memperoleh data *text* digunakan pengambilan data pada media social. Media social merupakan sebuah situs yang digunakan untuk berkomunikasi secara online [6]. Media social yang memiliki data mayoritas *text* adalah media social X atau dulu disebut dengan twitter. Media social X atau twitter merupakan media social yang digunakan sebagai sarana protes ataupun berkomentar, sehingga untuk memperoleh persepsi sentiment masyarakat akan lebih mudah dilakukan [7].

Dalam mengklasifikasi sentiment masyarakat digunakan metode *random forest*. Algoritma *random forest* merupakan metode dalam pohon Keputusan sebagai *base classifier* yang akan dibangun dan juga dikombinasikan untuk melakukan sampling terpadu untuk membangun pohon prediksi [8]. Selain itu digunakan juga metode *Bag of Words* yang digunakan untuk menemukan dan menghitung kata yang sering di keluarkan atau di tweet oleh pengguna twitter terhadap topik e-materai. Model *Bag of Words* adalah sebuah representasi sederhana dari *text* yang digunakan pada *natural language processing* untuk menghitung kemunculan dari sebuah kata pada *text mining* [9].

Dengan adanya analisis sentiment pada e-materai akan memberikan pengetahuan baru tentang opini masyarakat dari peristiwa saat tidak dapat diaksesnya situs penjualan e-materai. Sehingga hal tersebut dapat memberikan Gambaran pada pemerintah tentang penggunaan e-materai dan juga kesiapan dari situs yang akan digunakan dalam menjual e-materai, sehingga dengan adanya analisis sentiment ini akan dapat memberikan dampak pada layanan dari e-materai nantinya.

2. Metode Penelitian

2.1. Tinjauan pPustaka

Rahman (2023) melakukan penelitian dalam menganalisis Program Merdeka Belajar Kampus Merdeka pada media sosial Twitter. Digunakan pendekatan Algoritma *Random Forest* dengan pembobotan fitur *Bag of Words* didapatkan hasil klasifikasi yang terdiri dari 31% sentimen positif dan 69% sentiment negatif dengan tingkat akurasi sebesar 87.01%.

Rajs Farhan. R. P., et all (2022) melakukan analisis sentiment pada pengguna twitter tentang PILKADA 2020. Digunakan algoritma *Support Vector Machine* dan Model *Bag of Words* dari 400 tweet yang didapatkan. Hasilmenunjukkan bahwa terdapat 187 tweet positif dan 213 tweet negatif dengan nilai akurasi sebesar 87.5%.

Fanka Angelina, F., et all (2022) melakukan penelitian tentang analisis sentiment dengan mengambil data dari komentar pada *Google Play Store* tentang aplikasi Dana. Dalam menganalisis sentiment digunakan algoritma *Random Forest* dan menghasilkan nilai akurasi sebesar 84%.

M. R. Adrian, et all (2021) dalam penelitiannya membandingkan algoritma *Random Forest* dengan algoritma *Support Vector Machine* (SVM) untuk menganalisis sentiment dari media sosial Twitter tentang PSBB saat Pandemi Covid. Hasil dari penelitian tersebut menunjukkan bahwa algoritma *Random Forest* lebih baik dengan menghasilkan nilai akurasi sebesar 58% dibandingkan algoritma *Support Vector Machine* yang hanya menghasilkan nilai akurasi sebesar 55%.

Dari penelitian terdahulu sudah banyak yang menggunakan algoritma *Random Forest*, bahkan ada yang menunjukkan bahwa Algoritma *random Forest* lebih unggul dan juga menunjukkan metode *Bag of Words* juga memberikan hasil yang baik, oleh karena itu dalam penelitian ini digunakan Algoritma *Random Forest* dan metode *Bag of Words* dalam menganalisis sentiment E-materai.

2.2. Landasan Teori

2.2.1. Media Sosial

Media Sosial merupakan sebuah platform online yang digunakan para penggunanya untuk berinteraksi satu sama lain [6]. Media Sosial terdiri dari 2 kata baku, yaitu media yang dapat diartikan sebagai suatu perantara, dan sosial yang dapat diartikan sebagai hubungan masyarakat, sehingga media sosial secara keseluruhan dapat diartikan bahwa sebagai perantara untuk menghubungkan dan berkomunikasi antar individu masyarakat [5].

2.2.2. X (Twitter)

Media sosial X sebelumnya dikenal dengan Twitter merupakan sebuah media sosial yang digunakan sebagai alat komunikasi dimana sering digunakan sebagai sarana berekspresi menggunakan text yang berupa komentar yang sering disebut dengan tweet [13]. Postingan pada media sosial X atau biasa disebut dengan tweet merupakan kegiatan pengguna twitter untuk menyampaikan pendapatnya, selain itu dalam memberikan tweet pengguna twitter juga dapat meneruskan postingan dan menambahkan akun orang lain dengan memberikan @username [14].

2.2.3. Text Mining

Text mining merupakan bagian dari *data mining* dimana jenis data yang diolah adalah data dalam bentuk *text*, dalam *text mining* kumpulan data berupa teks yang tergabung dalam kalimat ataupun paragraf akan diolah untuk mendapatkan informasi atau pengetahuan baru [15] [16] [17] [18].

2.2.4. Analisis Sentimen

Analisis sentimen merupakan metode dalam memberikan nilai sentimen pada sebuah objek. Analisis sentimen digunakan untuk membedakan teks menjadi kategori opini dan fakta kemudian diklasifikasikan menjadi tiga golongan yaitu negatif, netral dan positif [19].

2.2.5. Random Forest

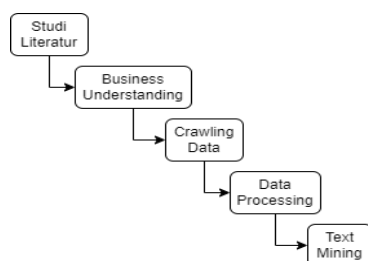
Random Forest merupakan metode pengembangan dari *Classification and Regression Trees (CART)* yang merupakan metode algoritma dari teknik pohon keputusan [12]. Banyak pohon yang akan ditumbuhkan pada *Random Forest* hingga mampu membentuk hutan (*forest*) dimana kumpulan pohon inilah yang nantinya akan di analisis [20].

2.2.6. Bag of Words (BoW)

Pendekatan *Bag of Words (BoW)* merupakan salah satu model *machine learning* yang digunakan untuk mengekstraksi fitur teks [21]. Model *Bag of Words* banyak digunakan dengan hasil yang baik untuk memprediksi pemodelan Bahasa dan klasifikasi dokumen karena *Bag of Words* sederhana dan fleksibel terhadap data teks yang spesifik [22].

3. Metode Penelitian

Kerangka umum dalam penelitian ini terlihat pada gambar 1 dimana memperlihatkan kerangka keseluruhan studi yang peneliti lakukan.



Gambar 1 Flowchart Penelitian

a. Studi Literatur

Dalam tahap ini dilakukan pengumpulan dan studi dari penelitian terdahulu yang berkaitan dengan *Text Mining*.

b. Business Understanding

Dalam tahap ini dilakukan analisis fakta yang berkembang dimasyarakat terkait dengan E-materai.

c. Crawling Data

Tahapan ini adalah proses pengumpulan data dimana data yang diambil adalah data komentar e-materai di media sosial X atau twitter.

d. Data Processing

Dalam tahap ini dilakukan *preprocess text* yang berupa *transformation*, *tokenization*, *normalization*, dan *filtering*.

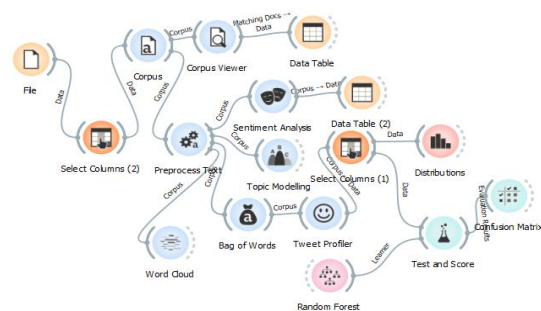
e. Text Mining

Sentimen yang telah didapatkan akan di uji dengan menggunakan algoritma *Random Forest* dan *Bag of Words*.

4. Hasil dan Pembahasan

4.1. Skenario Penelitian

Dalam penelitian ini digunakan aplikasi orange data mining versi 3.37. dimana desain penelitian dari orange data mining sebagai berikut:



Gambar 2 Desain orange datamining

4.1.1. Crawling Data

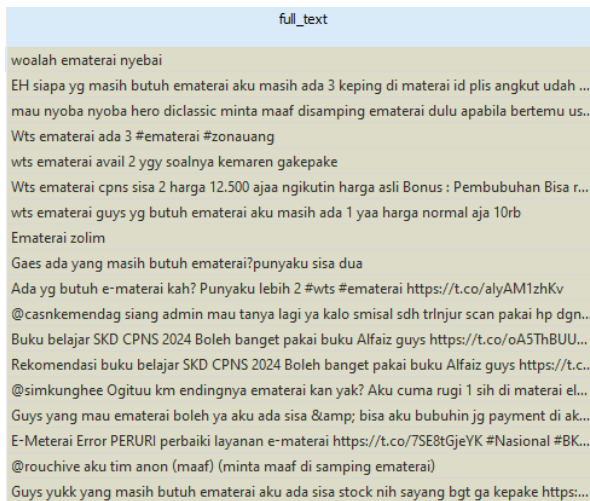
Proses pengumpulan data pada media sosial X atau twitter dilakukan dengan metode *web scraping* melalui *google collabs* dengan menggunakan sintaks berikut ini:

```

filename = 'ematerai.xlsx'
search_keyword = 'E-materai'
since:2024-05-09 until:2024-09-09
lang:id'

limit = 600
  
```

sintaks tersebut menunjukkan bahwa data yang akan diambil memiliki keyword “ematerai” dengan range pengambilan data dimulai dari 05 Mei 2024 sampai dengan 09 September 2024 dan hanya tweet yang berbahasa Indonesia saja. lalu limit data yang diambil hanya sebanyak 600 data tweet.

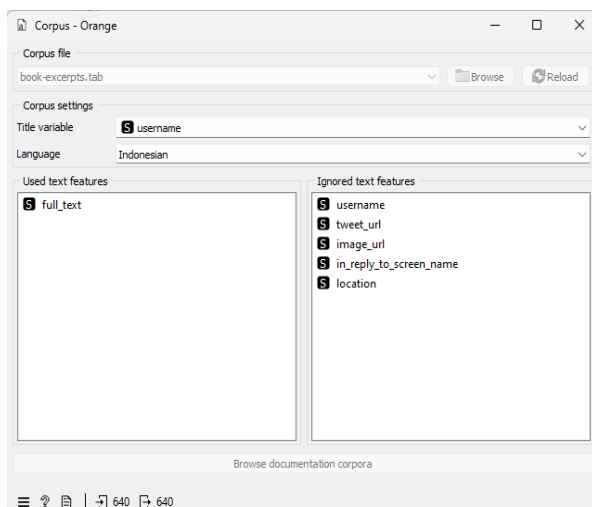


Gambar 3 Hasil Data Tweet

Setelah proses *web scraping* dijalankan didapatkan 640 data tweet tentang ulasan E-materai. Lalu data yang didapatkan akan dimasukan ke orange datamining.

4.1.2. Corpus Viewer

Digunakan *Corpus Viewer* untuk melihat file Text dan untuk memilih fitur yang akan diseleksi, dan hanya memilih *full_text* saja yang berupa data kumpulan tweet tentang E-materai, seperti pada gambar 3 dibawah ini:



Gambar 4 Corpus Viewer

4.1.3. Preprocessing Text

Sebelum data diproses data akan disiapkan terlebih dahulu dengan cara menghapus data yang tidak diperlukan, mengubah data dalam sistem dan beberapa proses lainnya melalui tahapan berikut:

a. Cleansing

Menghilangkan tanda baca dan karakter yang tidak mempengaruhi dalam penelitian seperti karakter tanda baca, situs url dll.

b. Transform Case

Mengubah huruf pada text menjadi huruf kecil agar sesuai dengan bentuk huruf.

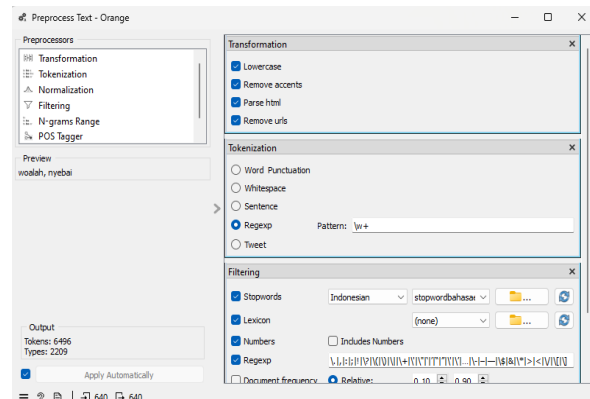
c. Tokenization

Membagi *text* kalimat atau paragraf menjadi bagian-bagian terpisah berdasarkan spasi dan tanda baca.

d. Stopwords

Digunakan untuk menghilangkan sejumlah kata kerja, kata sifat dan kata pembahasan lainnya.

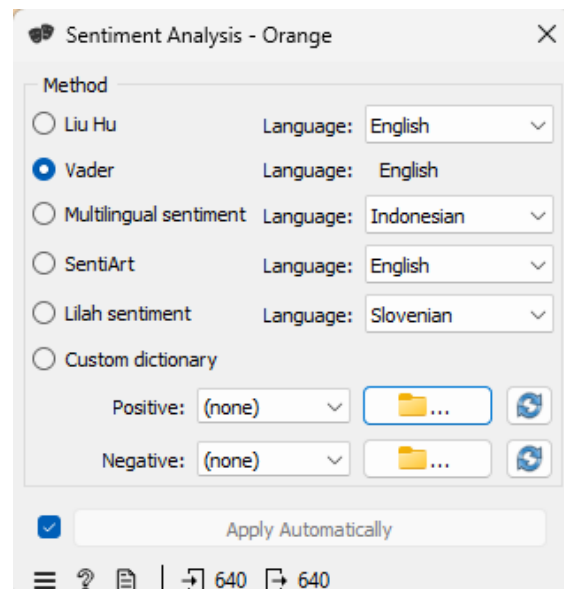
Berikut setingan yang dilakukan di orange datamining:



Gambar 5 Preprocess Text

4.1.4. Analisis Sentimen

Dalam analisis sentimen digunakan metode *vader* dengan tujuan untuk memperoleh hasil sentimen dalam bentuk atribut negatif, netral dan positif, dengan setingan orange data mining sebagai berikut:



Gambar 6 Sentiment Analysis

Dari hasil sentimen analisis didapatkan 55 tweet negatif, 550 tweet netral dan 35 tweet positif, atau 9% tweet negatif, 86% tweet netral dan 5% tweet positif

4.1.5. Tabel data

Tabel data merupakan respon dari luaran proses sebelumnya yaitu *sentiment analysis* seperti pada gambar dibawah ini:

id	compound	full_text	username	positive	negative	neutral
1	0	woalah ematerai nyabai	harjichous	0	0	1
2	0	Eh siapa yg masih butuh ematerai aku...	mayasthoughts	0	0	1
3	0.5574	mau nyoba nyoba hera diclasic mint...	katakurta	0.184	0	0.816
4	0	Wts ematerai ada 3 #ematerai #zonau...	anlghahisa	0	0	1
5	0	wts ematerai awal 2 ygy soalnya kema...	wumlan	0	0	1
6	0.5423	Wts ematerai cpns sisa 2 harga 12.500 ...	Simzen14	0.171	0	0.829
7	0	wts ematerai guys yg butuh ematerai...	selenephille	0	0	1
8	0	Ematerai zolim	madthepigi	0	0	1
9	0	Gaes ada yang masih butuh ematerai...	fleusdedhiver	0	0	1
10	0	Ada yg butuh e-materai kah? Punya...	dikarakina	0	0	1
11	0	@casnikemendag siang admin mau ta...	tutheunknwmn	0	0	1
12	0	Buku belajar SKD CPNS 2024 Boleh ba...	Brovskay	0	0	1
13	0	Rekomendasi buku belajar SKD CPNS...	Brovskay	0	0	1
14	0	@simlunghee Ggtau km endingnya...	atraxia	0	0	1
15	-0.298	Guys yang mau ematerai boleh ya aku...	stipidpeople17	0	0.084	0.916
16	-0.4018	E-Materai Error PERURI perbaiki layan...	KoranMandala...	0	0.197	0.803
17	0	Broschive aku tim anon (maaf) (mint...	heyho5hi	0	0	1
18	0	Guys yukk yang masih butuh ematerai...	makan_indmie	0	0	1
19	0.0772	WOOEE EMATERAI KU LEBIH 2 PLS HA...	stregifelvi	0.206	0	0.794
20	0	Yang butuh e-materai aku bisa bantu...	Sweetzzr	0	0	1
21	0	Wts aku ada sisa ematerai dari https://...	rabbit_dutokki	0	0	1

Dari hasil sentimen analisis pada tabel data didapatkan 55 tweet negatif, 550 tweet netral dan 35 tweet positif, atau 9% tweet negatif, 86% tweet netral dan 5% tweet positif.

4.1.6. Topic Modeling

Topic modeling menjadi salah satu teknik analisis yang menggunakan pendekatan *unsupervised learning*. Hasil *topic modeling* menampilkan 5 topik pada gambar berikut:

Topic	Topic keywords
1	peruri, tempel, pembubuhan, bel, workfess, kemarin, cpns2024, pas, aja, yah
2	cpns, udah, zonauang, beli, bubuhi, workfess, aja, bener, guys, harga
3	beli, web, bknngoid, pas, udh, tempel, pembubuhan, gue, udah, workfess
4	aja, beli, harga, sisa, zonauang, pake, butuh, elektronik, cpns, udah
5	aja, workfess, butuh, bubuhi, beli, udah, cpns, sisa, pake, udh

Gambar 7 Topic Modelling

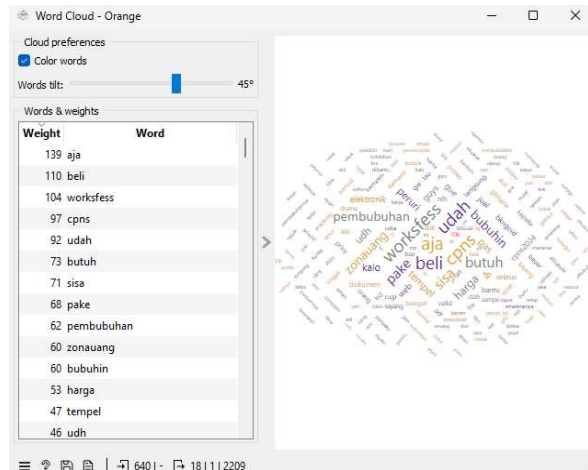
Dalam topic modeling terlihat bahwa 5 topik utama dalam sentiment E-materai adalah sebagai berikut:

1. peruri, tempel, pembubuhan, beli, workfess, kemarin, cpns2024, pas, saja, ya
2. cpns, sudah, zonauang, beli, bubuhi, workfess, saja, bener, guys, harga
3. beli, web, bknngoid, pas, udh, tempel, pembubuhan, gue, uda, workfess
4. saja, beli, harga, sisa, zonauang, pakai, butuh, elektronik, cpns, sudah
5. saja, workfess, butuh, bubuhi, beli, sudah, cpns, sisa, pakai, udh

4.1.7. Word Cloud

Word Cloud digunakan untuk memberikan tampilan visual dari hasil untuk memperkuat hasil *Topic modelling*

Seperti pada gambar berikut:



Gambar 8 Word Cloud

Terlihat dari hasil *word cloud* kata yang paling banyak ditunjukkan pada sentimen E-materai adalah saja, beli, workfess dan cpns.

4.1.8. Bag of Words

Dalam *bag of Words* digunakan setingan IDF (*Inverse Data Frequency*) untuk perhitungan bobot kata unik dalam korpus seluruh dokumen dengan setingan sebagai berikut:

Gambar 9 Bag of Words

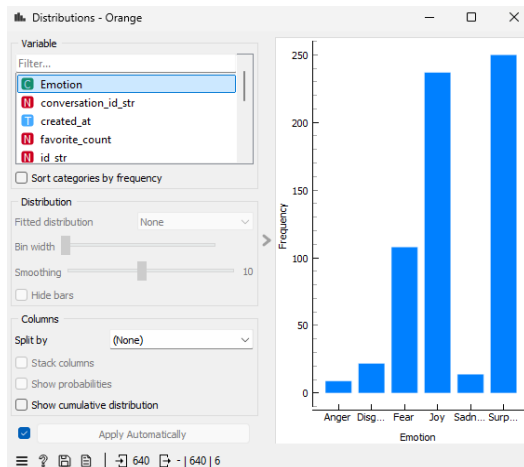
4.1.9. Tweet Profiler

Tweet Profiler mengambil data opini pada setiap dokumen dan akan menghitung probabilitas atau nilai sentiment. Widget ini mendukung 3 kategori yaitu *Ekmasn*, *Plutchiks* dan *Mood State Profiles* (POMS). Dalam penelitian ini digunakan klasifikasi *Ekman* dengan opsi *Multi class* sebagai berikut:

Gambar 10 Tweet profiler

4.1.10. Distribution

Distribution digunakan untuk menunjukkan hasil atribut dari proses sebelumnya yaitu *Tweet Profiler* seperti gambar dibawah

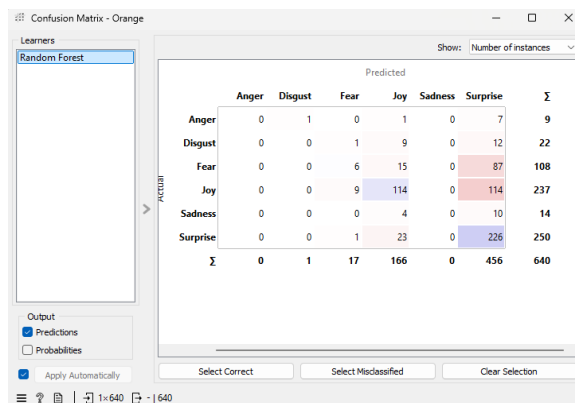


Gambar 11 Distribution

Hasil menunjukkan 6 emosi dalam sentiment E-materai pada media social X atau twitter. Dimana bentuk emosi yang dominan adalah *surprise* dengan 39% (250 tweet), *Joy* dengan 37% (237 tweet), *Fear* dengan 17% (108 tweet), *disgust* dengan 3% (22 tweet), *sadness* dengan 2% (14 tweet) dan *anger* dengan 1% (9 tweet).

4.1.11. Confusion Matrix

Confusion Matrix digunakan untuk memperlihatkan hasil dari prediksi algoritma *Random Forest* sebagai berikut:



Gambar 12 Confusion Matrix

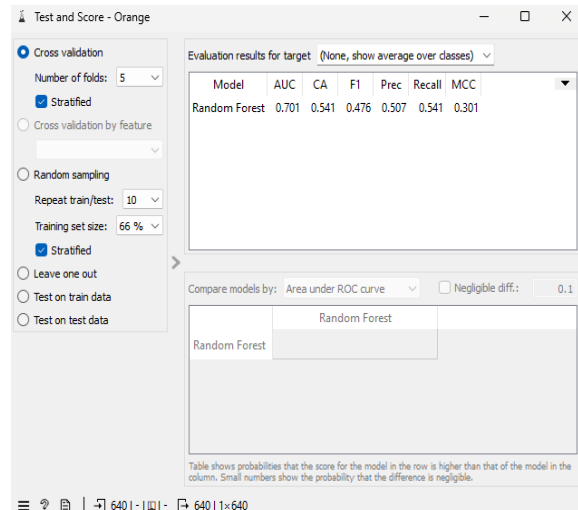
Hasil dari *Confusion Matrix* menunjukkan pada algoritma *Random Forest* sebagai berikut:

Tabel 1 Confusion Matrix Random Forest

Emosi	Actual		Predicted Random Forest	
Anger	9	1%	0	0%
Disgust	22	3%	1	0%
Fear	108	17%	17	3%
Joy	237	37%	166	26%
Sadness	14	2%	0	0%
Surprise	250	39%	456	71%
Total	640	100%	640	100%

4.1.12. Test and Score Random Forest

Pada proses *Test and Score* dilakukan pengujian pada algoritma *Random Forest* pada analisis Sentimen E-materai pada media sosial X atau twitter didapatkan nilai sebagai berikut:



Gambar 13 Test and Score

Terlihat bahwa nilai akurasi pada algoritma *random Forest* sebesar 70.1%, *precision* sebesar 50.7%, *recall* 54.1%, dan *F1-Score* 47.6%.

4.2. Pembahasan

Dari hasil *Crawling data* dari media sosial X atau Twitter yang dilakukan pada periode 05 Mei 2024 sampai dengan 09 September 2024 didapatkan 640 data tweet tentang E-materai. Data yang telah didapat akan melalui tahapan *preproseccing text* dimana data komentar atau opini akan dibersihkan terlebih dahulu, lalu data tweet komentar E-materai didapatkan hasil analysis sentiment yang menunjukkan bahwa mayoritas komentar tentang emateria bersifat netral dengan, 86% tweet netral lalu hanya 9% tweet bersifat negatif dan 5% tweet bersifat positif. Lalu pada *Topic Modeling* dan juga *Word Cloud* menunjukkan kata yang paling banyak keluar saat tweet E-materai adalah beli, workfess dan cpns. Selain itu hasil dari metode algoritma *Random Forest* menunjukkan nilai *Accuracy* sebesar 70.1%, *precision* sebesar 50.7%, *recall* 54.1%, dan *F1-Score* 47.6%.

5. Kesimpulan dan Saran

5.1. Kesimpulan

Berdasarkan hasil dari analisa dapat ditarik kesimpulan bahwa algoritma *Random Forest* metode *Bag of Words* memberikan akurasi sebesar 70.1 % pada analisis sentimen media sosial X atau twitter pada pembahasan e-materai. Dari hasil tweet yang didapat menunjukkan 86% opini twitter yang bernilai sentimen netral, 9% opini twitter bernilai negatif dan 5% opini twitter yang bernilai positif. Selain itu kata terbanyak pada tweet e-materai adalah tentang beli, worfess dan cpns.

5.2. Saran

Dalam menganalisis sentimen data yang didapatkan hanya 640 data saja di sebabkan adanya pembatasan pengambilan data pada media sosial X atau twitter yang telah dibatasi. Diharapkan jumlah data bisa ditambah. Selain itu juga dapat menggunakan algoritma klasifikasi lainnya selain *Random Forest*.

SUMBER RUJUKAN

Referensi

- [1] S. N. Fauza Tuanaya, "Fungsi Bea Meterai Dalam Surat Perjanjian," *Notarius*, vol. 13, no. 2, pp. 879–889, 2020, doi: 10.14710/nts.v13i2.31290.
- [2] V. P. Wulandari, "Kedudukan Hukum Meterai Dalam Perjanjian Perdata Di Kota Palangka Raya," *Moral. J. Ilmu Huk.*, vol. 5, no. 1, pp. 50–72, 2019.
- [3] M. Z. Z. H. Safaat and M. M. Gulton, "Kekuatan Hukum Materai Elektronik Dalam Perjanjian Jual Beli Online," *J. Lex Suprema*, vol. 5, no. 1, pp. 101–117, 2023, [Online]. Available: <https://aptika.kominfo.go.id/2020/01/revolusi-industri-4-0/>.
- [4] I. A. Pribadi, "Peruri terus lakukan pemulihan agar layanan e-materai kembali normal," *Antara*, Sep. 05, 2024. [Online]. Available: <https://www.antaranews.com/berita/4310067/peruri-terus-lakukan-pemulihan-agar-layanan-e-materai-kembali-normal>
- [5] A. Fatkhudin, F. A. Artanto, N. A. Safla, and D. Wibowo, "Decision Tree Berbasis SMOTE dalam Analisis Sentimen Penggunaan Artificial Intelligence untuk Skripsi," *Remik Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 8, no. April, pp. 494–505, 2024, doi: 10.33395/remik.v8i2.13531.
- [6] F. A. Artanto, "Analisis Sentimen Opini Publik terhadap Fenomena Bunuh Diri Mahasiswa di Twitter Menggunakan Algoritma Naive Bayes," *Satesi J. Sains Teknol. dan Sist. Inf.*, vol. 4, no. 1, pp. 70–76, 2024, doi: 10.54259/satesi.v4i1.2908.
- [7] Norlaila, W. W. Winarno, and E. T. Luthfi, "Analisis sentimen masyarakat tentang tambang di indonesia pada twitter menggunakan data mining," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 9, no. 3, pp. 1091–1099, 2024, doi: <https://doi.org/10.29100/jupi.v9i3.5402>.
- [8] I. Afdhal, R. Kurniawan, I. Iskandar, R. Salambue, E. Budianita, and F. Syafria, "Penerapan Algoritma Random Forest Untuk Analisis Sentimen Komentar Di YouTube Tentang Islamofobia," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 1, pp. 122–130, 2022, [Online]. Available: <http://ojs.serambimekkah.ac.id/jnkti/article/view/4004/pdf>
- [9] R. Pohan *et al.*, "Implementasi Algoritma Support Vector Machine dan Model Bag-of-Words dalam Analisis Sentimen mengenai PILKADA 2020 pada Pengguna Twitter," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 10, pp. 4924–4931, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [10] A. A. Rahman, "Implementasi Pembobotan BoW dan TF-IDF Pada Algoritma Random Forest untuk Analisis Sentimen," Universitas Pendidikan Indonesia, 2023. [Online]. Available: <http://repository.upi.edu/id/eprint/102235>
- [11] F. A. Larasati, D. E. Ratnawati, and B. T. Hanggara, "Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 9, pp. 4305–4313, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [12] M. R. Adrian, M. P. Putra, M. H. Rafialdy, and N. A. Rakhmawati, "Perbandingan Metode Klasifikasi Random Forest dan SVM Pada Analisis Sentimen PSBB," *J. Inform. Upgris*, vol. 7, no. 1, pp. 36–40, 2021, doi: 10.26877/jiu.v7i1.7099.
- [13] F. A. Artanto, "Support Vector Machine Berbasis Particle Swarm Optimization Pada Analisis Sentimen Anggota KPPS," *J. FASILKOM (teknologi Inf. dan Ilmu KOMputer)*, vol. 14, no. 1, pp. 75–79, 2024, doi: <https://doi.org/10.37859/jf.v14i1.6795>.
- [14] A. Wahyuningtyas, I. S. Sitanggang, and H. Khotimah, "Deteksi Spam pada Twitter Menggunakan Algoritme Naive Bayes," *J. Ilmu Komput. dan Agri-Informatika*, vol. 7, no. 1, pp. 31–40, 2020, doi: 10.29244/jika.7.1.31-40.
- [15] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naive Bayes Classifier Berbasis Konversi Ikon Emosi," *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.
- [16] A. Fatkhudin, A. Khambali, F. A. Artanto, and N. A. P. Zade, "Implementasi Algoritma Clustering K-Means Dalam Pengelompokan Mahasiswa Studi Kasus (Prodi Manajemen Informatika)," *J. Minfo Polgan*, vol. 12, no. 2, pp. 777–783, 2023, doi: 10.33395/jmp.v12i2.12494.
- [17] I. Rosyadi, F. A. Artanto, S. E. Rahmawati, H. Tri, and B. Joyo, "Decision Tree Dalam Analisis Keputusan Pembelian Program Pada Perkumpulan Penggiat Programmer Indonesia," *J. Fasilkom*, vol. XII, no. III, pp. 141–144, 2022.
- [18] M. F. I. Haq, I. Rosyadi, M. Nasir, and A. Khambali, "Sentiment Analisis Ulasan Aplikasi Livin Pada Google Play Store," *J. Surya Inform.*, vol. 14, no. 1, pp. 24–29, 2024.
- [19] H. H. Kusumawardani, I. Rosyadi, F. A. Artanto, F. I. Arzha, and N. A. Rachmayani, "Analisis Decision Tree dalam Pengaruh Digital Marketing terhadap Penerimaan Siswa Baru," *Remik*, vol. 6, no. April, pp. 225–231, 2022.
- [20] A. Miftahusalam, A. F. Nuraini, A. A. Khoirunisa, and H. Pratiwi, "Perbandingan Algoritma Random Forest, Naive Bayes, dan Support Vector Machine Pada Analisis Sentimen Twitter Mengenai Opini Masyarakat Terhadap Penghapusan Tenaga Honorer," *Semin. Nas. Off. Stat.*, no. 1, pp. 563–572, 2022, doi: 10.34123/semnasoffstat.v2022i1.1410.
- [21] E. Daniati and H. Utama, "Analisis Sentimen Dengan Pendekatan Ensemble Learning Dan Word Embedding Pada Twitter," *J. Inf. Syst. Manag.*, vol. 4, no. 2, pp. 125–131, 2023, doi: 10.24076/joism.2023v4i2.973.
- [22] D. Krisnandi, R. N. Ambarwati, A. Y. Asih, A. Ardiansyah, and H. F. Pardede, "Analisis Komentar Cyberbullying Terhadap Kata Yang Mengandung Toksisitas Dan Agresi Menggunakan Bag of Words dan TF-IDF Dengan Klasifikasi SVM," *J. Linguist. Komputasional*, vol. 6, no. 2, pp. 36–41, 2023, doi: 10.26418/jlk.v6i2.85.