

ANALISIS SENTIMEN MASYARAKAT TERHADAP KEBIJAKAN PENERAPAN PPKM DI MEDIA SOSIAL TWITTER DENGAN MENGGUNAKAN METODE XGBOOST

(Analysis Of Community Sentiment On The Policy Of Implementation Of PPKM On Twitter Social Media Using Xgboost Method)

I Putu Angga Purnama Widiarta, Ramaditia Dwiyanaputra, Arik Aranta

^[1]Dept Informatics Engineering, Mataram University

Jl. Majapahit 62, Mataram, Lombok NTB, INDONESIA

Email: anggawidiarta55@gmail.com, [rama, arikaranta]@unram.ac.id

Abstract

Corona Virus Disease (Covid-19) is a virus that causes respiratory infections in humans. Indonesia is a country that has been infected with this virus, the implementation of restrictions on community activities (PPKM) is implemented by the government as a policy to reduce the spread of Covid-19. Pros and cons arise due to the impact of the policy. Therefore, assessing how public opinion or sentiment is towards this policy is important to do. This study aims to implement the XGBoost algorithm in the sentiment classification process. Sentiment analysis targets public opinion on Twitter, the dataset used is 1958 positive tweets and 3980 negative tweets. At the preprocessing stage, case-folding, stopwords removal, tokenizing, and stemming are carried out. Giving weights to terms uses the Term Frequency-Relevance Frequency method to turn each term into a number. In the final stage, classification is carried out by implementing the XGBoost method with optimal hyperparameter scores. K-fold cross validation is used to evaluate model performance. Based on the evaluation results, the best performance was obtained by a model with a hyperparameter value with an $n_estimator$ of 1000, a $learning_rate$ of 0.1, a max_depth of 6, a $subsample$ of 1, a $gamma$ of 0 and utilizing the stem-ming process in preprocessing with an accuracy value of 85.27%, precision of 86.07%, and recall of 85.23%.

Keywords: Sentimen Analisis, PPKM, Covid-19, XGBoost, Term Frequency – Relevance Frequency, K-Fold Cross Validation, Preprocessing, Hyperparameter

1. PENDAHULUAN

Corona Virus Disease (Covid-19) merupakan virus yang dapat menyebabkan infeksi saluran pernafasan pada manusia, mulai dari flu biasa hingga penyakit yang lebih serius seperti *Severe Acute Respiratory Syndrome* (SARS) dan *Middle East Respiratory Syndrome* (MERS). Virus ini ditemukan pada akhir tahun 2019, dimana ini bermula di Wuhan, Provinsi Hubei, China [1].

Berdasarkan data yang diperoleh dari laman *website worldometers*, kasus harian *covid-19* dalam skala global mengalami penurunan yang signifikan, tercatat pada tanggal 1 januari 2022 jumlah kasus harian yang tercatat sebanyak 1.858.097 kemudian per tanggal 15 mei 2022 jumlah kasus harian yang tercatat menunjukkan angka sebanyak 657.158, menurunnya kasus harian ini sangat dipengaruhi oleh faktor dari pemerataan vaksinasi yang sudah dilakukan, kemudian untuk skala di Indonesia, kasus harian yang tercatat

pada tanggal 15 januari 2022 menunjukkan jumlah sebanyak 1054 kasus, tidak seperti kasus harian yang terjadi di lingkup global pada umumnya, kasus harian di Indonesia malah mengalami kenaikan dari bulan Januari sampai bulan Februari yang dimana puncaknya terjadi pada tanggal 17 Februari dimana kasus harian yang tercatat berjumlah 63.956, lalu kurva tersebut tiba – tiba mengalami penurunan dari bulan tersebut hingga per tanggal 27 April yang menunjukkan jumlah sebanyak 617.

Indonesia merupakan salah satu negara yang terjangkit oleh virus *covid-19*, oleh karena itu pemerintah Indonesia menerapkan kebijakan pemberlakuan pembatasan kegiatan masyarakat (PPKM) per tanggal 11 Januari 2021 demi menekan angka persebaran dari virus ini. Kebijakan tersebut berdampak signifikan pada berbagai sektor kehidupan masyarakat, salah satunya adalah dari sektor ekonomi. Pembatasan kemampuan untuk melakukan kegiatan skala besar akan mengakibatkan perekonomian

menjadi semakin sulit, dengan beberapa kegiatan ekonomi berhenti.

Pro dan kontra bermunculan di kalangan masyarakat, khususnya pada media sosial. *Twitter* merupakan salah satu *platform* yang sering digunakan selama periode PPKM. Dilansir dari *Global Digital Statistic "Digital, Social & Mobile in 2019"*, pada tahun 2019 pengguna media sosial di Indonesia berjumlah 150 juta pengguna, *twitter* menyumbang angka terbanyak yang mencakup lebih dari 52% [2]. Opini publik bersifat tidak dibatasi dan bebas di *twitter* [3]. Dalam PPKM, opini publik dinyatakan dalam bentuk reaksi positif, negatif, atau netral terhadap pemerintah. Analisis sentimen dilakukan untuk mengolah opini – opini tersebut dan memperoleh inferensi tekstual dari isi benak masyarakat Indonesia.

Analisis sentimen adalah metode untuk memahami, menganalisis, dan memproses *input* tekstual secara otomatis untuk memperoleh informasi sentimen dari suatu opini [4]. *Machine learning* merupakan metode pendekatan yang menonjol untuk memproses sentimen. Penelitian mengenai sentimen analisis dengan menggunakan pendekatan *machine learning* terkait dengan kebijakan PPKM sebelumnya dengan metode *Support Vector Machine (SVM)* sudah dilakukan dengan nilai akurasi yang diperoleh sebesar 64% [5]. Kemudian pada penelitian yang dilakukan oleh Krisdiyanto, T dkk. Proses analisis opini diklasifikasikan menjadi 2 sentimen yaitu positif atau negatif, proses klasifikasi menggunakan metode *Naive Bayes Classifiers*, diperoleh akurasi sebesar 99% yang termasuk kedalam polaritas positif dan 1% pada polaritas negatif [6].

Berdasarkan pada hal - hal yang telah dipaparkan pada paragraf sebelumnya. Pada penelitian ini, akan dilakukan implementasi penggunaan dari metode *XGBoost* sebagai algoritma klasifikasi, dan mengimplementasikan metode *TF-RF (Term Frequency – Relevance Frequency)* sebagai metode untuk menentukan bobot dari suatu term pada teks. *XGBoost* menghemat waktu, mengoptimalkan sumber daya memori, dan dapat diterapkan secara paralel selama proses implementasi untuk mengelola sentimen.

2. TINJAUAN PUSTAKA

Bryan Pratama, dkk. pada tahun 2019 melakukan studi analisis dengan judul "Sentiment Analysis Of The Indonesian Police Mobile Brigade Corps Based On Twitter Posts Using The SVM And NB Methods" pada studi tersebut dilakukan analisa pada *tweet* dengan kata kunci "Brimob" dimana total *tweet* yang digunakan sebanyak 1000 *tweets*. Studi ini

menggunakan *text mining* dengan didukung oleh *support vector machine (SVM)* untuk mengklasifikasikan sentimen publik terhadap brimob di *twitter*. Akurasi yang diperoleh dengan SVM mencapai 86,96% sedangkan dengan *Naive Bayes* diperoleh akurasi sebesar 86,48% [7].

Tahun 2019, Eka dkk. melakukan studi analisis sentimen pada contoh Gojek dan Grab, menggunakan algoritma *Naive Bayes Classifier*, dan menemukan bahwa akurasi, *recall*, dan presisi metode *Naive Bayes Classifier* masing-masing adalah 72,33%, 73,95%, dan 73,24% [8]. Penelitian tersebut kemudian dilanjutkan oleh D. A. Al-Qudah et al. melakukan penelitian analitik sentimen terhadap penyedia layanan *e-payment* menggunakan algoritma yang disebut *XGBoost* dan membandingkan hasilnya dengan *J84*, *Naive Bayes*, dan *KNN*. Akurasi maksimum didapatkan oleh *KNN* dan *XGBoost* yang masing-masing memiliki nilai *recall* 85,2 persen dan 82,8 persen. Sedangkan dengan menggunakan nilai presisi *Naive Bayes* didapatkan akurasi tertinggi sebesar 72% [9].

ajar Fathur Rachman pada tahun 2020, dalam penelitiannya yang berjudul "Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada media sosial *Twitter*" melakukan penelitian sentimen analisis dengan menggunakan algoritma *Latent Dirichlet Allocation (LDA)* untuk mengelompokkan opini masyarakat dengan tujuan mengetahui topik pembicaraan yang sering dibahas masyarakat terkait dengan wacana vaksinasi, hasil analisis menunjukkan bahwa masyarakat lebih banyak memberikan respon positif terhadap wacana tersebut (30%) dibandingkan dengan respon negatifnya (26%) [10].

Angelina Puput Giovani, dkk. Pada tahun 2020, dalam penelitian dengan judul "Analisis Sentimen Aplikasi Ruang Guru di *Twitter* Menggunakan Algoritma Klasifikasi" melakukan komparasi beberapa algoritma yaitu *Naive Bayes*, *Support Vector Machine*, dan *K-Nearest Neighbour* yang menggunakan *feature selection* dengan yang tidak menggunakan *feature selection*, serta juga membandingkan nilai *Area Under Curve* dari metode – metode tersebut untuk mengetahui algoritma mana yang paling optimal, hasil pengujian menunjukkan bahwa algoritma SVM dengan *feature selection* menjadi algoritma terbaik dengan nilai akurasi 78,55% dan *AUC* 0,853 [11].

Sulaiman Ainin, dkk. Pada tahun 2020, dengan penelitian berjudul "Sentiment Analyses Of Multilingual Tweets On Halal Tourism" menuliskan tentang penelitian yang mereka lakukan pada *tweet* – *tweet* dari rentang waktu 2008 - 2018 yang berkaitan

dengan multilingual halal *tourism* dimana konten dan sentimen dari *tweet – tweet* tersebut dianalisa, mereka menggunakan 19 kata kunci untuk mengesktrak data dari *tweet* dimana 5 kata kunci tersebut adalah bahasa Malaysia, dan sisanya bahasa Inggris. Setelah dilakukan analisa diperoleh kesimpulan bahwa *tweet* terkait pariwisata halal pada negara non muslim melebihi jumlah *tweet* pada negara muslim, penelitian ini menunjukkan bahwa pariwisata halal mulai populer di negara seperti Inggris, Kanada, dan Spanyol [12].

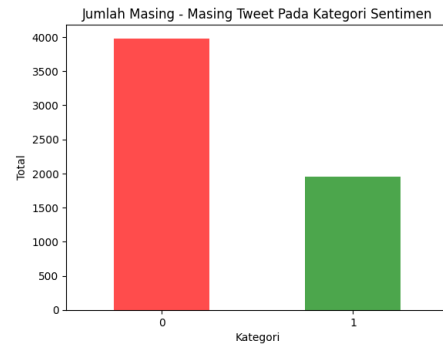
Elena, Podasca pada tahun 2021 dengan penelitian berjudul “Predicting The Movement Direction Of OMXS30 Stock Index Using *XGBoost* and Sentiment Analysis” melakukan prediksi pada indeks pasar saham Swedia menggunakan metode *XGBoost* yang disertakan dengan sentimen analisis dari berita keuangan guna membantu meningkatkan kinerja klasifikasi ketika memprediksi tren harga harian dari indeks pasar saham Swedia yaitu OMXS30. Hasil pada penelitian ini menunjukkan bahwa *XGBoost* memiliki kinerja yang baik dalam mengklasifikasikan tren hari OMXS30 dimana akurasi yang diperoleh mencapai 73% [13].

Pada tahun 2021. Aldiansyah Putra, dkk. dalam penelitiannya berjudul “Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Pada Media Sosial Twitter Menggunakan Algoritma SVM” melakukan penelitian terhadap respons masyarakat di *Twitter* berupa pro dan kontra mereka kepada kebijakan pemerintah dalam pemberlakuan pembatasan kegiatan masyarakat (PPKM), metode yang digunakan pada penelitian tersebut adalah *Support Vector Machine*, dengan memanfaatkan 3000 dataset yang kemudian diperoleh hasil akurasi sebesar 64%. Dari penelitian tersebut, algoritma SVM dapat mengenali *tweet* yang berisikan penolakan PPKM sebagai *tweet* bertendensi negatif dan juga kata – kata yang memiliki hubungan terhadap tendensi negatif tersebut [5].

3. METODE PENELITIAN

3.1. Dataset

Dataset yang digunakan pada penelitian ini memiliki jumlah total sebanyak 5.938 *tweet* dengan jumlah masing-masing yaitu 1958 *tweet* positif dan 3980 *tweet* negatif.



Gambar 1. Jumlah Tweet Pada Setiap Kategori

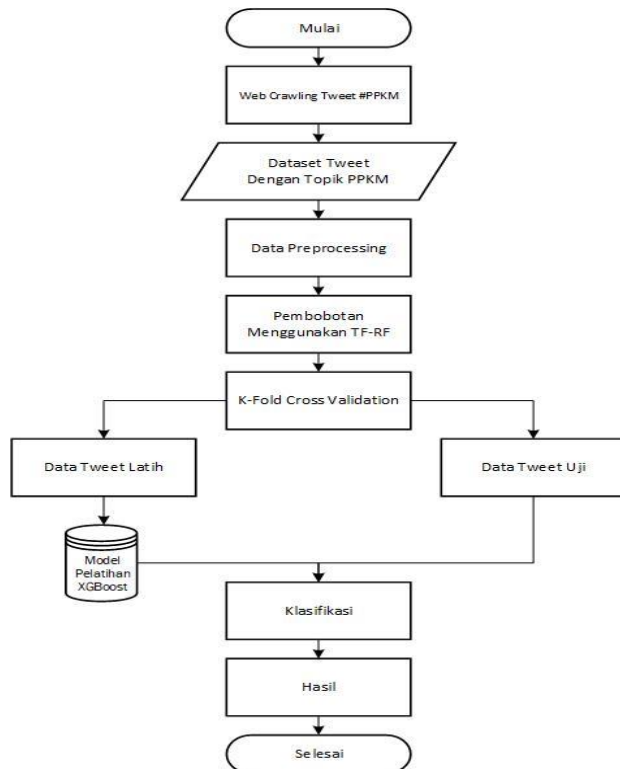
3.2. Alur Klasifikasi Sentimen Tweet

Analisis sentimen masyarakat Indonesia pada media sosial *twitter* terhadap kebijakan pemerintah dalam melaksanakan penerapan PPKM di Indonesia menggunakan *XGBoost* dilakukan dalam beberapa tahapan.

Tahapan dimulai dari *crawling tweets* yang memiliki tagar #ppkm dimana *tweets* tersebut digunakan sebagai *dataset* pada sistem, *tweet – tweet* yang berisikan opini netizen Indonesia terkait dengan penerapan kebijakan PPKM yang dilakukan oleh pemerintah diambil dari *database twitter* dan *tweet – tweet* tersebut memiliki beberapa kriteria yakni *tweet* memuat tagar #ppkm. *Tweet* haruslah menggunakan bahasa Indonesia, dan *tweet – tweet* tersebut adalah *tweets* yang dibuat dalam rentang waktu yang dimulai dari tanggal 1 April 2020 hingga 1 April 2022.

Tweets kemudian dibagi menjadi 2 jenis dataset, yaitu *tweet* yang digunakan sebagai data *training* dan *tweet* yang digunakan sebagai data *testing*, selanjutnya dilakukan *preprocessing* agar *tweets* nantinya menjadi lebih relevan pada saat memasuki proses *training* oleh model, setelah *preprocessing* dilakukan, tahapan selanjutnya yaitu melakukan pembobotan kata (*term*).

Kata yang digunakan adalah *unigram* dari hasil *preprocessing*. Dalam penelitian ini metode *feature weighting* yang digunakan adalah metode TF-RF, baru setelah itu dilakukan *training* model, dan dilanjutkan dengan mengevaluasi model tersebut dengan *testing* data untuk memperoleh tingkat akurasi dari model yang telah dilatih.



Gambar 2. Diagram Proses Klasifikasi

3.2.1. Web Crawling

Tweet dikumpulkan melalui jejaring media sosial *Twitter*. *Tweet* yang dikumpulkan adalah *tweet* yang menggunakan *hashtag* “#ppkm”. “#ppkm” kemudian dimasukkan pada *query* pencarian *tweets* yang digunakan oleh *library snsrape* untuk melakukan *crawling* data pada *tweet – tweet* berbahasa Indonesia yang memuat tagar “ppkm” di dalamnya.

3.2.2. Input Dataset Tweet Training dan Testing

Tweet yang digolongkan sebagai data *training* digunakan untuk membuat model klasifikasi sedangkan *tweet* yang digunakan sebagai data *testing* digunakan untuk menguji model yang telah dibuat.

- Tweet testing* sama seperti *tweets training*, yang membedakan adalah *tweet testing* tidak diketahui jenis kategori sentimennya yang kemudian dimasukkan ke dalam sistem untuk diproses pada *training tweets*.
- Tweet training* adalah *tweets* yang telah dilakukan *preprocessing* pada seluruh *tweets*-nya, lalu dilanjutkan dengan melakukan *term weighting* dengan TF-RF, yang kemudian di-*training* menggunakan algoritma XGBoost.

3.2.3. Text Preprocessing Tweets Dataset

Text preprocessing yang dilakukan pada penelitian ini dibagi menjadi 4 tahap yaitu *casefolding*, *tokenization*, *stopword removal*, dan *stemming*.

- Casefolding* merupakan tahap dimana *dataset* yang ada akan disamaratakan penggunaan huruf kapitalnya. Seluruh *tweet* akan diubah hurufnya menjadi huruf kecil, ini bertujuan agar *tweets* menjadi konsisten pada penggunaan hurufnya, dan mencegah sistem mengalami kebingungan dikarenakan kata yang sama apabila penulisan hurufnya berbeda.
- Tokenization* merupakan tahap untuk mentransformasikan *tweets* menjadi kumpulan kata yang disebut *terms*. Pada *tokenization* juga dilakukan *cleaning* berupa penghilangan tanda baca, simbol, dan karakter, ini dilakukan untuk menghilangkan *noise* pada data awal. Sebelum proses *tokenization*, terlebih dahulu dilakukan proses *casefolding*. Karakter selain huruf dihilangkan dan dianggap *delimiter*.
- Stopword removal* yaitu tahapan untuk menghapus kata – kata yang diyakini memiliki sedikit makna, bahkan hampir tidak bermakna, kata – kata tersebut seperti ‘yang’, ‘yaitu’, ‘di’, ‘tempat’, ‘terus’, ‘walau’, dan masih banyak lainnya.
- Stemming* merupakan tahapan untuk memperoleh kata dasar pada suatu kalimat dengan cara melakukan pemotongan pada imbuhan (*affix*) kata, baik itu *prefix* maupun *suffix* [14]. Proses *stemming* dilakukan dengan menggunakan algoritma Nazief dan Adriani.

3.2.4. Feature Weighting

Penelitian ini menggunakan metode *Term Frequency - Relevance Frequency* (TF-RF) dalam melakukan pembobotan *term*, dimana prosesnya akan dibagi menjadi tahapan-tahapan berikut ini :

- Term Frequency* (TF) adalah faktor yang menentukan bobot dalam sebuah teks yang tergantung pada seberapa sering teks (*term*) tersebut muncul. Bobot *term* akan meningkat seiring dengan banyaknya kemunculan *term* tersebut pada dokumen. $f(d, t)$ merupakan frekuensi kemunculan *term t* pada dokumen d .

$$TF(d, t) = f(d, t) \quad (1)$$

- Relevance Frequency* (RF) adalah metode yang diusulkan oleh Man Lan, pada metode ini

dinyatakan bahwa frekuensi terhadap kemunculan *term* di kategori yang berkaitan dilihat sebagai pertimbangan relevansi dokumen [15]. Persamaan pada RF dijabarkan pada Persamaan (2). b merupakan jumlah dokumen yang mengandung *term* t , sedangkan c adalah jumlah dokumen yang tidak mengandung *term* t .

$$tf_{tdrf} = tf_{td} * \log\left(2 + \frac{b}{\max(1,c)}\right) \quad (2)$$

- c. *Term Frequency - Relevance Frequency* (TF-RF), yaitu tahapan dimana hasil yang diperoleh pada tahap *term frequency* dikalikan dengan hasil yang diperoleh pada tahap *relevance frequency*.

3.2.5. Pelatihan dan Proses Klasifikasi

Maksud dari proses pelatihan ini adalah untuk mengembangkan sebuah model pelatihan yang akan digunakan dalam proses klasifikasi. Algoritma yang pelatihan yang digunakan dalam penelitian ini adalah *eXtreme Gradient Boosting* (XGBoost) *classifier*. Algoritma tersebut adalah algoritma berbasis *tree* yang termasuk ke dalam golongan algoritma *tree* yang sama dengan *decision tree* dan *random forest* [16].

Dengan bantuan prinsip *ensemble*, XGBoost mengubah sejumlah set pembelajar yang lemah (pohon) menjadi model yang kuat sehingga dapat membuat prediksi yang akurat [17]. Metode menambahkan model baru ke pendekatan *ensemble* disebut *boosting*, hal ini dilakukan untuk mengoreksi kesalahan dari model sebelumnya. Model akan ditambahkan satu per satu sampai sampai tidak ada lagi peningkatan yang mungkin dilakukan. Teknik *ensemble* menggunakan model pohon klasifikasi dan regresi yang disebut *tree ensemble models*. Strategi yang dikenal dengan teknik *ensemble* menggabungkan prediksi dari berbagai *tree* menjadi satu [18]. Ini berusaha untuk secara berurutan memodelkan setiap *predictor* menggunakan kesalahan residual dari model sebelumnya.

$$h_0(x) = \text{mean}(Y) \quad (3)$$

$$\hat{Y} = Y - h_0(x) \quad (4)$$

Persamaan (3) dan (4) digunakan untuk menentukan nilai prediksi awal dan kesalahan residual dari model asli. Model pertama dibuat menggunakan Persamaan (3), model berikutnya dibuat menggunakan Persamaan (4). Y adalah nilai residual *error* model awal dan $h_0(x)$ adalah nilai prediksi awal dari model pertama. Model kedua kemudian dibuat menggunakan residual *error* dari model pertama untuk menentukan nilai prediksinya. Kesalahan residual model pertama

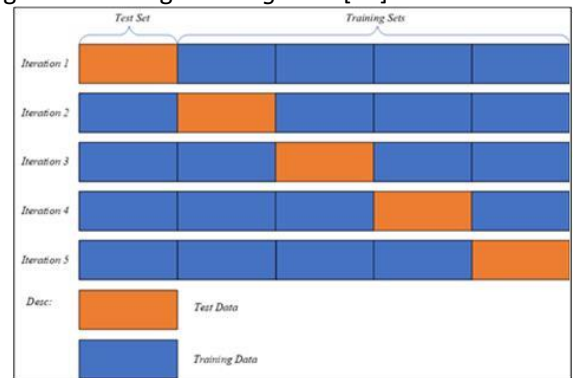
dan kedua kemudian digunakan untuk membuat model ketiga untuk menentukan nilai prediksinya. Sebanyak $n_{estimator}$ ditetapkan, maka proses ini akan terus berulang [18].

$$L^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (5)$$

XGBoost menghasilkan satu set *decision tree* yang mana setiap model pohon bergantung pada pohon sebelumnya. Nilai prediksi awal untuk model pertama di XGBoost akan lemah, tetapi karena lebih banyak model dibangun, bobot diperbarui untuk menghasilkan prediksi yang lebih kuat. Untuk meminimalkan fungsi tujuan, nilai proyeksi dari masing – masing model akan dijumlahkan kemudian dimasukkan ke dalam Persamaan (5) [18]. n adalah jumlah model yang digunakan, l adalah fungsi untuk mengukur selisih antara target y_i dan \hat{y}_i yang diprediksi. $f_t(x_i)$ adalah model baru yang dibangun. Sedangkan Ω adalah fungsi untuk membuat model terhindar dari *overfitting*. Persamaan (5) digunakan untuk mencari nilai keseluruhan.

3.2.6. Pengujian

Pengujian dilakukan dengan menggunakan teknik *k-fold cross validation*, ini merupakan teknik untuk mengevaluasi hasil klasifikasi. *Dataset* dibagi menjadi beberapa *fold*. Pada setiap iterasi yang dilakukan, salah satu *fold* digunakan sebagai *testing data*, dan sisa *fold* digunakan sebagai *training data*. Proses ini dilakukan sebanyak nilai K yang ditetapkan hingga seluruh *fold* digunakan sebagai *testing data* [19].



Gambar 3. Ilustrasi Cross Validation 5 Fold

Apabila dijabarkan, maka prosesnya adalah sebagai berikut ini [19]:

- Jumlah *instance* dibagi sebanyak K *fold*
- Pada iterasi ke-1 adalah saat bagian ke-1 dijadikan sebagai data uji dan empat bagian sisanya dijadikan sebagai data latih. Kemudian dilakukan penghitungan akurasi atau kesamaan atau kedekatan pada hasil

pengukuran dengan menggunakan angka atau data yang sebenarnya berdasarkan porsi dari data tersebut.

- c. Pada iterasi ke-2 yaitu saat bagian ke-2 dijadikan sebagai data uji dan bagian lain sisanya dijadikan sebagai data latih. Kemudian, dilakukan penghitungan akurasi berdasarkan porsi data tersebut.
- d. Lalu seterusnya penghitungan dilakukan hingga mencapai iterasi atau *fold* ke-K. Rata – rata akurasi yang diperoleh dari K buah akurasi akan dijadikan sebagai akurasi final.

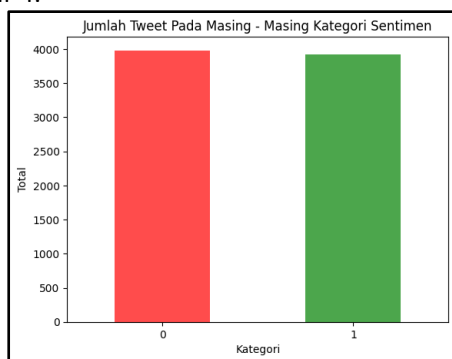
4. HASIL DAN PEMBAHASAN

Pada bagian ini dilaporkan hasil beserta pembahasan penelitian untuk mengukur performa model *XGBoost* dalam melakukan klasifikasi sentimen *tweet* PPKM dengan menggunakan fitur pembobotan TF-RF. Metode – metode tersebut diimplementasikan dengan bahasa pemrograman *python* dan bantuan *library Scikit-learn* serta beberapa *library* lainnya.

Skenario yang diujikan yaitu untuk melihat bagaimana akurasi model dipengaruhi oleh variasi dari *score hyperparameter XGBoost*, pengaruh augmentasi data pada akurasi model *XGBoost*, pengaruh *stemming* pada akurasi model *XGBoost*, pengaruh nilai K pada metode *k-fold cross validation* pada akurasi model *XGBoost*.

4.1. Augmentasi Data

Dikarenakan jumlah *dataset* yang *imbalanced* apabila mengacu pada laman [20], maka disini penulis melakukan augmentasi pada *dataset* sehingga jumlah *dataset* menjadi seperti pada yang ditampilkan oleh Gambar 4.



Gambar 4. Jumlah data setelah diaugmentasi

4.2. Preprocessing

Preprocessing yang telah dilakukan pada *dataset* terdiri menjadi beberapa bagian yang akan dipaparkan sebagai berikut :

4.2.1. Case folding

	tweet	sentiment
0	Karena stimulasi untuk #anakberkebutuhankhusus tidak optimal jika dilak...	0
1	Ye ampun mbh anda seorang Wapres loh.&Gelar anda seorang kiyai @Kiy...	0
2	Mulaik kumat lagi. 😊 #ppkm https://t.co/MNDSPHJU69	0
3	Welcome PPKM bau bau naik level nih, gimana nih mang\r\n#ppkm	0
4	Sesungguhnya Allah tidak menciptakan kita, hambaNya ini, sebagai makhlu...	0

↓

	tweet	sentiment
0	karena stimulasi untuk #anakberkebutuhankhusus tidak optimal jika dilak...	0
1	ye ampun mbh anda seorang wapres loh.&gelar anda seorang kiyai @kiy...	0
2	mulaik kumat lagi. 😊 #ppkm https://t.co/mndspjhu69	0
3	welcome ppkm bau bau naik level nih, gimana nih mang\r\n#ppkm	0
4	sesungguhnya allah tidak menciptakan kita, hambanya ini, sebagai makhlu...	0

Gambar 5. Tahap Case Folding

Pada Gambar 5 disajikan perubahan *tweet* setelah dilakukan *case folding*. Seluruh teks pada *tweets* diubah menjadi huruf kecil demi menghindari *tweets* dengan makna ganda karena ada perbedaan kapital.

4.2.2. Tokenizing

	tweet	sentiment
0	Karena stimulasi untuk #anakberkebutuhankhusus tidak optimal jika dilak...	0
1	Ye ampun mbh anda seorang Wapres loh.&Gelar anda seorang kiyai @Kiy...	0
2	Mulaik kumat lagi. 😊 #ppkm https://t.co/MNDSPHJU69	0
3	Welcome PPKM bau bau naik level nih, gimana nih mang\r\n#ppkm	0
4	Sesungguhnya Allah tidak menciptakan kita, hambaNya ini, sebagai makhlu...	0

↓

	tweet	sentiment
0	karena stimulasi untuk #anakberkebutuhankhusus tidak optimal jika dilak...	0
1	ye ampun mbh anda seorang wapres loh.&gelar anda seorang kiyai @kiy...	0
2	mulaik kumat lagi. 😊 #ppkm https://t.co/mndspjhu69	0
3	welcome ppkm bau bau naik level nih, gimana nih mang\r\n#ppkm	0
4	sesungguhnya allah tidak menciptakan kita, hambanya ini, sebagai makhlu...	0

Gambar 6. Tahap Tokenizing

Pada Gambar 6 disajikan perubahan *tweet* setelah dilakukan *tokenizing*. Kata atau *term* pada *tweet* dipisahkan sehingga menjadi satuan kata atau dalam artian lain disebut dengan *token*.

4.2.3. Stopwords Filtering

	tweet	sentiment
0	[karena, stimulasi, untuk, tidak, optimal, jika, dilakukan, secara, onl...	0
1	[ye, ampun, mbh, anda, seorang, wapres, lohampgelar, anda, seorang, kiy...	0
2	[mulaik, kumat, lagi]	0
3	[welcome, ppkm, bau, bau, naik, level, nih, gimana, nih, mang]	0
4	[sesungguhnya, allah, tidak, menciptakan, kita, hambanya, ini, sebagai,...	0

	tweet	sentiment
0	[stimulasi, optimal, dilakukan, online, memenuhi, kebutuhan, tersebut, ...	0
1	[ye, ampun, mbh, seorang, wapres, lohampgelar, seorang, kiyai, marufami...	0
2	[mulaik, kumat]	0
3	[welcome, ppkm, bau, bau, naik, level, nih, gimana, nih, mang]	0
4	[sesungguhnya, allah, menciptakan, hambanya, makhluk, lemah, mudah, ber...	0

Gambar 7. Tahap Stopwords Filtering

Pada Gambar 7 disajikan perubahan *tweet* setelah dilakukan *stopwords filtering*. *Term* yang tidak relevan dan tidak memiliki makna dibuang, dan *term* yang penting atau relevan disimpan.

4.2.4. Stemming

	tweet	sentiment
0	[stimulasi, optimal, dilakukan, online, memenuhi, kebutuhan, tersebut, ...	0
1	[ye, ampun, mbh, seorang, wapres, lohampgelar, seorang, kiyai, marufami...	0
2	[mulaik, kumat]	0
3	[welcome, ppkm, bau, bau, naik, level, nih, gimana, nih, mang]	0
4	[sesungguhnya, allah, menciptakan, hambanya, makhluk, lemah, mudah, ber...	0

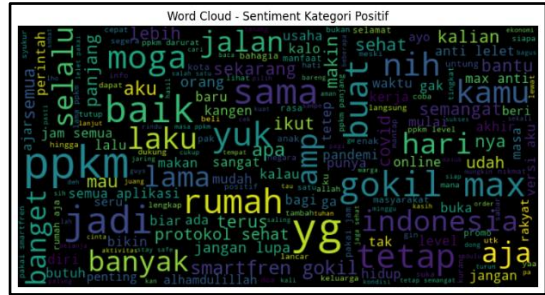
	tweet	sentiment
0	[stimulasi, optimal, laku, online, penuh, butuh, sebut, tetap, operasi,...	0
1	[ye, ampun, mbh, orang, wapres, lohampgelar, orang, kiyai, marufamin, k...	0
2	[mulaik, kumat]	0
3	[welcome, ppkm, bau, bau, naik, level, nih, gimana, nih, mang]	0
4	[sungguh, allah, cipta, hamba, makhluk, lemah, mudah, putus, asa, manus...	0

Gambar 8. Tahap Stemming

Pada Gambar 8 disajikan perubahan *tweet* setelah dilakukan *stemming*. *Term* atau kata dihilangkan imbuhan untuk diubah menjadi kata dasarnya saja. Metode yang digunakan pada proses ini menerapkan algoritma nazief dan adriani [14].

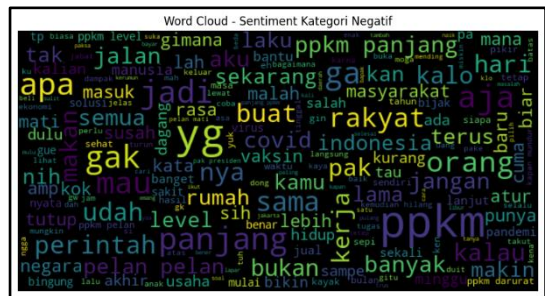
4.3. Word Cloud

Pada subbab ini, ditampilkan *word cloud* dari kedua jenis kategori. *Word cloud* yang ditampilkan merupakan *word cloud* dari masing – masing kategori sentimen dengan *tweet* yang telah melewati *preprocessing* (termasuk *stemming*). Visualisasi *word cloud* bertujuan untuk menampilkan informasi berupa kata – kata yang paling sering muncul dalam *tweets* yang ada.



Gambar 9. Word Cloud Sentimen Kategori Positif

Pada Gambar 9 ditampilkan visualisasi *word cloud* dari *tweets – tweets* dengan kategori positif yang berada pada *dataset* yang telah dilakukan augmentasi.



Gambar 10. Word Cloud Sentimen Kategori Negatif

Pada Gambar 10 ditampilkan visualisasi *word cloud* dari *tweets – tweets* dengan kategori negatif yang berada pada *dataset* yang telah dilakukan augmentasi.

4.4. Term Weighting

Setelah *preprocessing* dilakukan, *term* diolah kembali untuk diubah ke dalam bentuk vektor numerik. Frekuensi kemunculan *term* dicari dengan menggunakan Persamaan (1), kemudian dicari juga relevansi frekuensi *term* dengan menggunakan Persamaan (2). Hasil yang telah diperoleh dari kedua tahapan tersebut lalu dikalikan untuk memperoleh bobot dari TF-RF. Contoh hasil kalkulasi dapat dilihat pada Gambar 11. Vektor yang telah diperoleh kemudian disimpan sebagai model *training*. Model *training* yang telah disimpan dalam proses sebelumnya akan digunakan untuk melakukan klasifikasi terhadap data *testing*.

	hadir	boss	mega	wifi	jackpot	...	phisical	kunyah	perubahan	perubahan	acau	sentiment
0	1.05115	0.727	1.11394	1.25527	1.07918	...	0.0	0.0	0.0	0.0	0.0	1
1	0.00000	0.000	0.00000	0.00000	0.00000	...	0.0	0.0	0.0	0.0	0.0	1
2	0.00000	0.000	0.00000	0.00000	0.00000	...	0.0	0.0	0.0	0.0	0.0	1
3	0.00000	0.000	0.00000	0.00000	0.00000	...	0.0	0.0	0.0	0.0	0.0	1
4	0.00000	0.000	0.00000	0.00000	0.00000	...	0.0	0.0	0.0	0.0	0.0	1

Gambar 11. Word Cloud Sentimen Kategori Negatif

4.5. Pengujian

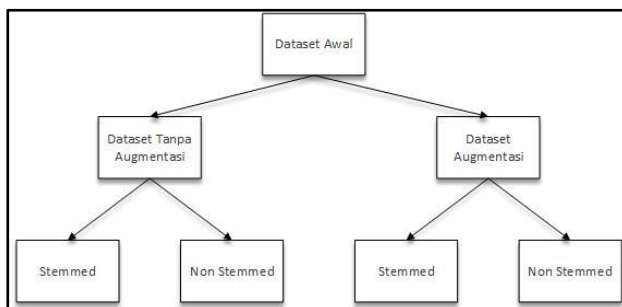
Model *XGBoost* yang digunakan dalam pengujian merupakan model yang telah diperoleh kombinasi

score *hyperparameter* optimalnya terlebih dahulu menggunakan metode *gridsearchCV* dengan nilai akurasi terbaik yang digunakan sebagai acuan dalam memperoleh kombinasi *hyperparameter* optimal. Pada Tabel I ditampilkan nilai *hyperparameter* yang akan diterapkan pada model dalam melakukan pengujian – pengujian selanjutnya.

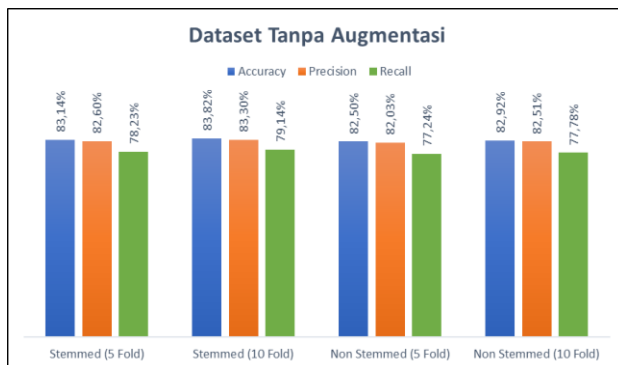
TABEL I. HYPERPARAMETER OPTIMAL

Hyperparameter	Score
n_estimator	1000
gamma	0
subsample	1
learning_rate	0.1
max_depth	6

Pengujian selanjutnya dilakukan untuk mengetahui bagaimana perbandingan hasil akurasi dari proses klasifikasi yang dilakukan antara *dataset* yang telah diaugmentasi dan *dataset* yang tidak diaugmentasi, sekaligus juga untuk melihat perolehan hasil akurasi pada kedua jenis *dataset* yang telah disebutkan sebelumnya ketika *dataset* tersebut melewati proses *stemming* dengan *dataset* yang digunakan tanpa melewati proses *stemming*. Pada Gambar 12 disajikan kombinasi skenario pengujian yang dilakukan.



Gambar 12. Skenario Pengujian

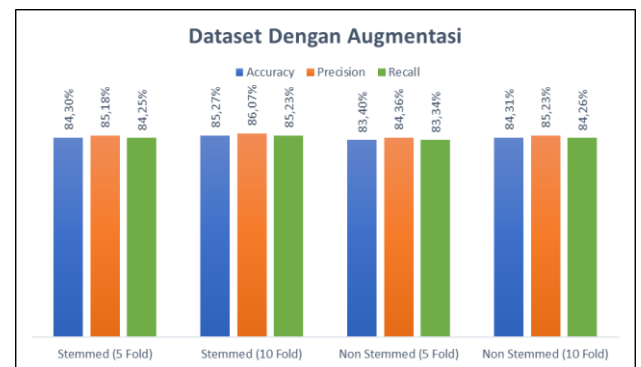


Gambar 13. Grafik Hasil Pengujian Pada Dataset Tanpa Augmentasi

Melalui diagram yang ditampilkan pada Gambar 13, diketahui bahwa rata – rata nilai akurasi, presisi, dan *recall* yang diperoleh pada setiap skema yang dilakukan memiliki rata – rata selisih nilai yang kecil antara satu skema dengan skema atau skenario pengujian yang lainnya. Nilai akurasi, presisi, dan *recall* yang tertinggi berada pada skenario *dataset* yang telah melewati proses *stemming* dengan nilai *fold* yang digunakan yaitu sama dengan 10 *fold*, sementara untuk nilai akurasi, presisi, dan *recall* yang paling rendah berada pada skenario dengan menggunakan *dataset* yang tidak melewati proses *stemming* dengan nilai *fold* sama dengan 5.

Melihat pada Gambar 13, tahapan *stemming* memberikan nilai akurasi, presisi, dan *recall* yang lebih tinggi dari skenario yang tidak melewati proses ini. Bahkan nilai akurasi, presisi, dan *recall* pada skenario tanpa *stemming* dengan nilai *fold* yang sama dengan 10 masih kalah apabila dibandingkan dengan skenario yang melewati tahapan *stemming* tetapi dengan nilai *fold* yang lebih rendah yaitu 5 *fold* saja.

Dilihat secara keseluruhan pada 4 skema yang ditunjukkan pada Gambar 15, diperoleh nilai standar deviasi pada akurasi, presisi, dan *recall* secara berurutan yaitu 0.48% akurasi, 0.45% presisi, dan 0.70% *recall*. Mengacu pada nilai standar deviasi tersebut dapat dinyatakan bahwa perubahan nilai akurasi, presisi, dan *recall* pada masing – masing skenario tidak begitu tinggi dampaknya.



Gambar 14. Grafik Hasil Pengujian Pada Dataset Augmentasi

Gambar 14 menunjukkan hasil pengujian bahwa rata – rata nilai akurasi, presisi, dan *recall* yang diperoleh pada masing – masing skema yang dilakukan memiliki rata – rata selisih nilai yang kecil antara satu skema dengan skema atau skenario pengujian yang lainnya. Apabila dibandingkan dengan grafik pada Gambar 13, nilai selisih yang ada pada Gambar 14 masih lebih besar. Pada *dataset* dengan augmentasi, nilai akurasi, presisi, dan *recall* yang

tertinggi berada pada skenario *dataset* yang telah melewati proses *stemming* dengan nilai *fold* yang digunakan yaitu sama dengan 10 *fold* dimana skema ini juga sama dengan apa yang telah dipaparkan pada *dataset* tanpa augmentasi.

Sementara untuk nilai akurasi, presisi, dan *recall* yang paling rendah berada pada skenario dengan menggunakan *dataset* yang tidak melewati proses *stemming* dengan nilai *fold* sama dengan 5. Sama seperti pada skema yang dilakukan di *dataset* tanpa augmentasi, apabila mengacu pada Gambar 14. Tahapan *stemming* memberikan nilai akurasi, presisi, dan *recall* yang lebih tinggi dari skenario yang tidak melewati proses ini, bahkan nilai akurasi, presisi, dan *recall* pada skenario tanpa *stemming* dengan nilai *fold* yang sama dengan 10 masih kalah apabila dibandingkan dengan skenario yang melewati tahapan *stemming* tetapi dengan nilai *fold* yang lebih rendah yaitu 5 *fold* saja.

Dilihat secara keseluruhan pada 4 skema yang ditunjukkan pada Gambar 14, diperoleh nilai standar deviasi pada akurasi, presisi, dan *recall* secara berurutan yaitu 0.66% akurasi, 0.60% presisi, dan 0.67% *recall*. Mengacu pada nilai standar deviasi tersebut dapat dinyatakan bahwa perubahan nilai akurasi, presisi, dan *recall* pada masing – masing skenario tidak begitu tinggi juga dampaknya.

Berdasarkan dari penjelasan Gambar 13 dan 14, hasil pengujian terbaik dihasilkan pada skenario dengan *dataset* yang telah diaugmentasi, serta melewati proses *stemming* dengan nilai *k-fold* yang digunakan sama dengan 10 dan diperoleh nilai rata – rata yaitu 85.27% akurasi, 86.07% presisi, dan 85.23% *recall*. Peningkatan yang signifikan juga dapat dilihat pada rata – rata nilai *recall* yang berada pada skenario *dataset* tanpa augmentasi dengan *dataset* yang diaugmentasi. Nilai standar deviasi pada akurasi dan presisi di *dataset* dengan augmentasi juga lebih tinggi apabila dibandingkan dengan *dataset* tanpa augmentasi.

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Pada penelitian yang telah dilakukan, diperoleh hasil penelitian yang dapat ditarik kesimpulannya bahwa implementasi metode *XGBoost* dalam melakukan klasifikasi sentimen pada *tweets* dengan menggunakan *dataset* yang telah diaugmentasi dan telah melewati proses *stemming*, menghasilkan nilai akurasi sebesar 85.27%, presisi sebesar 86.07%, dan *recall* sebesar 85.23% setelah diuji dengan

menggunakan teknik *cross validation* dengan nilai *k-fold* yaitu sama dengan 10.

Kemudian pada *dataset* yang mengalami augmentasi cenderung menghasilkan nilai akurasi, presisi, dan *recall* yang lebih tinggi dibandingkan dengan *dataset* yang tidak mengalami augmentasi, hal ini dikarenakan augmentasi *dataset* merujuk pada teknik penambahan variasi atau modifikasi pada *dataset* sehingga meningkatkan jumlah sampel yang tersedia yang menyebabkan model memiliki lebih banyak variasi dan informasi yang dapat dipelajari.

Sementara pada skenario pengujian yang menggunakan tahapan *stemming* cenderung menghasilkan nilai *accuracy*, *precision*, dan *recall* yang lebih tinggi dibandingkan dengan pengujian tanpa tahapan *stemming*. Hal ini terjadi karena imbuhan berupa *prefiks* maupun *sufiks* pada kata yang memiliki kata dasar yang sama malah menambah variasi fitur saat melakukan pembobotan, sehingga memberikan nilai bobot yang berbeda – beda.

5.2. Saran

Untuk penelitian selanjutnya, perlu digunakan lebih banyak *dataset* berupa *tweets* sehingga dapat memperoleh performa model yang lebih baik lagi dikarenakan banyaknya *dataset* maka dapat meningkatkan representasi data yang digunakan pada pelatihan model. Selanjutnya, pada metode pembobotan fitur yang digunakan, dapat dicoba mengimplementasikan metode pembobotan selain TF-RF. Kemudian untuk metode augmentasi yang diterapkan juga dapat dicoba dengan menggunakan metode selain dari melakukan substitusi pada kata untuk melihat bagaimana perbandingan performa pada masing – masing metode yang ada.

DAFTAR PUSTAKA

- [1] W. Zendrato, "Gerakan mencegah daripada mengobati terhadap pandemi covid-19," *Jurnal Education and development*, vol. 8, no. 2, hlm. 242–242, 2020, [Daring]. Tersedia pada: <https://www.sehatq.com/artikel/bahaya-virus->
- [2] R. Yanuarti, "Jurnal Sistem dan Teknologi Informasi Analisis Media Sosial Twitter Terhadap Topik Vaksinasi Covid-19," vol. 6, no. 2, 2021, [Daring]. Tersedia pada: <http://jurnal.unmuhjember.ac.id/index.php/JUSTINDO>
- [3] N. D. Asih dan M. Rosit, "Opini Publik di Media Sosial: Analisis Isi Opini Kandidat Ahok-Djarot dan Anies-Sandi di Twitter," *CoverAge: Journal of Strategic Communication*, vol. 8, no. 2, hlm. 45–56, Mar 2018.
- [4] L. Ardiani, H. Sujaini, dan T. Tursina, "Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap

- Pembangunan di Kota Pontianak,” *Jurnal Sistem dan Teknologi Informasi (Justin)*, vol. 8, no. 2, hlm. 183, Apr 2020, doi: 10.26418/justin.v8i2.36776.
- [5] A. Putra, D. Haeirudin, H. Khairunnisa, dan R. Latifah, “Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Pada Media Sosial Twitter Menggunakan Algoritma Svm,” *Prosiding Semnastek*, 2021.
- [6] T. Krisdiyanto, E. Maricha, dan O. Nurharyanto, “Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naïve Bayes Clasifiers,” *Jurnal CoreIT*, vol. 7, no. 1, 2021.
- [7] B. Pratama dkk., “Sentiment Analysis of the Indonesian Police Mobile Brigade Corps Based on Twitter Posts Using the SVM and NB Methods,” dalam *Journal of Physics: Conference Series*, Institute of Physics Publishing, Mei 2019. doi: 10.1088/1742-6596/1201/1/012038.
- [8] E. Y. and W. A. D. and S. A. Sari, “Sentiment analysis of customer satisfaction on transportation network company using naive Bayes classifier,” *019 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, hlm. 1–6, 2019.
- [9] D. A. Al-Qudah, A. M. Al-Zoubi, P. A. Castillo-Valdivieso, dan H. Faris, “Sentiment analysis for e-payment service providers using evolutionary extreme gradient boosting,” *IEEE Access*, vol. 8, hlm. 189930–189944, 2020, doi: 10.1109/ACCESS.2020.3032216.
- [10] F. F. Rachman dan S. Pramana, “Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter,” *Health Information Management Journal ISSN*, vol. 8, no. 2, hlm. 2655–9129, 2020.
- [11] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, dan W. Gata, “Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi,” *Jurnal Teknoinfo*, vol. 14, no. 2, hlm. 115–123, Jul 2020, doi: 10.33365/JTI.V14I2.679.
- [12] S. Ainin, A. Feizollah, N. B. Anuar, dan N. A. Abdullah, “Sentiment analyses of multilingual tweets on halal tourism,” *Tour Manag Perspect*, vol. 34, hlm. 100658, Apr 2020, doi: 10.1016/J.TMP.2020.100658.
- [13] E. Podasca, “Predicting the Movement Direction of OMXS30 Stock Index Using XGBoost and Sentiment Analysis,” 2021, [Daring]. Tersedia pada: www.bth.se
- [14] D. Wahyudi, T. Susyanto, dan D. Nugroho, “Implementasi dan analisis algoritma stemming nazief & adriani dan porter pada dokumen berbahasa indonesia,” *Jurnal Ilmiah SINUS*, vol. 15, no. 2, 2017.
- [15] A. N. Assidyk, Setiawan, E. Budi, dan K. Isman, “Analisis Perbandingan Pembobotan TF-IDF dan TF-RF pada Trending Topic di Twitter dengan Menggunakan Klasifikasi K-Nearest Neighbor,” *eProceedings of Engineering*, vol. 7, no. 2, 2020.
- [16] T. Chen dan C. Guestrin, “XGBoost: A scalable tree boosting system,” dalam *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Agu 2016, hlm. 785–794. doi: 10.1145/2939672.2939785.
- [17] W. F. Mustika, H. Murfi, dan Y. Widyaningsih, “Analysis Accuracy of XGBoost Model for Multiclass Classification - A Case Study of Applicant Level Risk Prediction for Life Insurance,” dalam *2019 5th International Conference on Science in Information Technology (IC-SITech)*, 2019, hlm. 71–77. doi: 10.1109/IC-SITech46713.2019.8987474.
- [18] M. Riza Kurniawanda, F. Adline, dan T. Tobing, “Analysis Sentiment Cyberbullying in Instagram Comments with XGBoost Method,” *International Journal of New Media Technology*, vol. 9, no. 1, hlm. 28, 2022.
- [19] R. Dwiyanaputra, G. Satya Nugraha, F. Bimantoro, dan A. Aranta, “Deteksi SMS Spam Berbahasa Indonesia menggunakan TF-IDF dan Stochastic Gradient Descent Classifier,” *Jurnal Teknologi Informasi, Komputer, dan Aplikasinya (JTika)*, vol. 3, no. 2, hlm. 200–207, 2021, [Daring]. Tersedia pada: <http://jtika.if.unram.ac.id/index.php/JTIKA/>
- [20] “Imbalanced Data | Machine Learning | Google for Developers.” <https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data> (diakses 18 Juni 2023).