

# Analisis Sentimen Mahasiswa di Surabaya Terhadap Pelayanan Vaksinasi COVID-19 Menggunakan Beberapa Classifier

Meliana Kusuma Pangkasidhi, Henry Novianus Palit, Andre Gunawan

Program Studi Informatika Fakultas Teknologi Industri Universitas Kristen Petra

Jl. Siwalankerto 121 – 131 Surabaya 60236

Telp. (031) – 2983455, Fax. (031) – 8417658

E-Mail: melianapang029@gmail.com, hnpalit@petra.ac.id, andre.gunawan@petra.ac.id

## ABSTRAK

Indonesia merupakan salah satu negara yang sedang berjuang menangani pandemi virus COVID-19 pada saat ini dengan melaksanakan vaksinasi COVID-19. Pemerintah sekarang sedang berjuang mengajak masyarakat untuk melakukan vaksinasi dengan cara memaksimalkan pelayanan vaksinasi COVID-19. Realitanya, pelayanan vaksinasi masih memiliki kendala pada beberapa aspek. Untuk melihat berbagai pandangan mengenai pelayanan vaksinasi yang telah dilaksanakan, maka dari itu dilakukan penelitian di bidang analisis sentimen untuk menganalisa opini masyarakat. Dalam penelitian ini akan menggunakan Naïve Bayes, Support Vector Machine (SVM), Random Forest, dan Light Gradient Boosting Machine (LGBM) untuk melakukan klasifikasi teks dan dibandingkan performanya menggunakan *evaluation metrics*. Terdapat dua jenis dataset yang digunakan, yaitu dataset kuesioner dan dataset media sosial. Model kuesioner akan diuji menggunakan dataset media sosial, sedangkan model media sosial menggunakan dataset media sosial yang melalui *split dataset*. Hasil pengujian mengatakan bahwa model yang dilatih dengan dataset media sosial menghasilkan performa lebih baik daripada model kuesioner. Model terbaik untuk klasifikasi aspek dan sentimen adalah Random Forest.

**Kata Kunci:** analisis sentimen, klasifikasi teks, vaksinasi COVID-19, naïve bayes, SVM, random forest, LGBM

## ABSTRACT

*Indonesia is one of the countries that are currently struggling to deal with the COVID-19 virus pandemic by providing vaccination. The government is currently trying to persuade the public to do vaccination by maximizing COVID-19 vaccination services. In reality, vaccination services still have problems with some aspects. To see various insights on vaccination services that have been implemented, therefore a research was conducted in the field of sentiment analysis to analyze public opinion. In this research, classifiers that will be used are Naïve Bayes, Support Vector Machine (SVM), Random Forest, and Light Gradient Boosting Machine (LGBM) to perform text classification and their performances will be compared with evaluation metrics. There are two types of datasets used, namely questionnaire dataset and social media dataset. The questionnaire model will be tested using a social media dataset, while the social media model will use social media dataset that will be split. The testing results show that the model trained with the social media dataset produces better performance than the questionnaire model. Of these four classifiers, the best model for aspect and sentiment classification is Random Forest.*

**Keywords:** sentiment analysis, text classification, vaccination, COVID-19, naïve bayes, SVM, random forest, LGBM

## 1. PENDAHULUAN

Indonesia merupakan salah satu negara yang sedang berjuang menangani dan memberantas pandemi virus COVID-19 pada saat ini dengan melaksanakan vaksinasi COVID-19. Pemerintah sekarang sedang berjuang mengajak masyarakat untuk melakukan vaksinasi dengan cara memaksimalkan pelayanan vaksinasi COVID-19. Realitanya, pelayanan vaksinasi masih memiliki beberapa kendala yang dirasakan oleh masyarakat sebagai penerima vaksinasi. Beberapa adalah terbatasnya fasilitas ruang tunggu dan jumlah kursi untuk calon penerima vaksin atau pasca menerima vaksin [10], sertifikat vaksin yang belum diterima [14], dan data sertifikat vaksin yang salah [5]. Untuk melihat berbagai *insight* mengenai pelayanan vaksinasi, maka penelitian analisis sentimen ini dilakukan. Penelitian ini akan berfokus pada tanggapan penerima vaksin terhadap pelayanan vaksinasi sebagai salah satu bentuk usaha baik pemerintah dalam menangani pandemi ini. Penelitian ini berfokus pada satu kota dahulu, yaitu Surabaya, dengan melihat tanggapan para mahasiswa yang melakukan vaksinasi di kota Surabaya. Sedangkan mahasiswa dipilih karena dianggap mampu bertanggung jawab atas komentar yang dilontarkan. Penelitian ini akan melakukan perbandingan dan analisis sentimen terhadap pelayanan vaksinasi COVID-19 dengan menggunakan Naïve Bayes (NB), Support Vector Machine (SVM), Light Gradient Boosting Machine (LGBM), dan Random Forest (RF). Keempat metode ini akan dibandingkan dan dipilih berdasarkan evaluasi metrik yang paling baik. Metode NB banyak digunakan beberapa penelitian pada bidang sentiment analysis dan NB ini menjadi pemenangnya [4]. Kemudian, SVM ini juga digunakan karena beberapa penelitian mengatakan bahwa SVM ini memiliki akurasi tertinggi untuk klasifikasi [11]. Selain itu, ada LGBM dan RF. LGBM adalah pemenang untuk *multiclass classification* sentimen dalam hal akurasi dan F-score. Sedangkan, RF berada di posisi kedua [1].

## 2. PENELITIAN SEBELUMNYA

Penelitian Sentiment Analysis sebelumnya, terdapat penelitian yang mengangkat masalah penanganan COVID-19 yang menimbulkan pro dan kontra di masyarakat. Metode yang digunakan adalah NB. Penelitian ini bertujuan untuk memahami tren opini publik tentang COVID-19 di Indonesia dari perspektif umum maupun ekonomi. Namun, hasil dari penelitian ini menunjukkan ketidakpuasan terhadap kinerja pemerintah, bukan usaha baik pemerintah [12]. Selain itu, terdapat penelitian yang menggunakan SVM dan RF untuk rekomendasi produk menggunakan analisis sentimen dan menghasilkan kesimpulan bahwa SVM berhasil untuk prediksi dan akurasi klasifikasinya, sedangkan RF efektif untuk akurasi dan robustness [8]. Penelitian lainnya melakukan perbandingan beberapa metode untuk menganalisis sentimen topik umum, menunjukkan bahwa LGBM adalah metode yang paling unggul, kemudian disusul oleh

RF. Metode yang digunakan adalah Logistic Regression, SVM, RF, Multinomial Naïve Bayes, Extreme Gradient Boosting, LGBM dan Deep Learning [1]. Oleh karena itu pada penelitian ini akan dilakukan dengan membandingkan beberapa metode yang dianggap bagus berdasarkan penelitian-penelitian terdahulu, yaitu NB, SVM, RF, dan LGBM.

### 3. DATASET

Terdapat dua jenis dataset yang digunakan, yaitu dataset kuesioner dan dataset media sosial. Pertanyaan yang ditanyakan pada pengumpulan adalah aspek seputar pelayanan vaksinasi, seperti kemudahan mendapatkan informasi vaksinasi, tempat pelaksanaan, ketepatan waktu vaksinasi dimulai dan waktu tunggu vaksinasi, ketertiban dalam penerapan protokol kesehatan, dan kemudahan mendapat dan ketepatan informasi dalam sertifikat vaksinasi. Ada dua jenis model yang diuji. Mode pertama adalah model yang dilatih dengan keseluruhan dataset kuesioner dan kemudian di *testing* menggunakan dataset media sosial. Hal ini untuk melihat apakah model dataset kuesioner dapat melakukan klasifikasi dengan baik. Model kedua adalah model yang di *train* dan *test* menggunakan dataset media sosial dengan melakukan pembagian dataset sebesar 80:20. Pengumpulan dataset kuesioner berhasil mengumpulkan sebanyak 305 responden. Namun ternyata ada 1 responden yang belum melakukan vaksinasi dosis 1 dan dosis 2. Sedangkan dataset media sosial berhasil mengumpulkan 112 responden. Setelah semua data disortir, total jumlah dataset kuesioner adalah 2985 data yang siap digunakan. Didalamnya terdiri dari 2494 data positif dan 491 data negatif. Terdapat 597 data untuk aspek Informasi, Waktu, Tempat, Protokol, dan Sertifikat. Sementara itu, dataset media sosial memiliki 558 data. Didalamnya terdiri dari 423 data positif dan 135 data negatif. Terdapat 122 data aspek Informasi, 104 data aspek Tempat, 125 data aspek Waktu, 107 data aspek Protokol, dan 100 data aspek Sertifikat.

### 4. METODE

#### 4.1 Naïve Bayes

Naïve Bayes (NB) adalah pengklasifikasi probabilistik berdasarkan teorema Bayes, dengan mempertimbangkan asumsi independen yang kuat [6]. Naïve Bayes terkenal karena sederhana dan kinerja yang baik dalam klasifikasi teks [15]. Jenis Naïve Bayes yang cocok untuk klasifikasi teks adalah Multinomial Naïve Bayes (MNB). MNB ini mengasumsikan distribusi multinomial sehingga dapat digunakan untuk fitur diskrit seperti jumlah kata dalam klasifikasi teks [1]. MNB membutuhkan input jumlah fitur integer, seperti *word count* [15]. Setiap teks berisi istilah yang diberikan probabilitas berdasarkan jumlah kemunculannya dalam teks tertentu. Kemudian Naïve Bayes dapat mempelajari pola dataset yang telah dilabeli dan membandingkan konten di semua dokumen dengan menyusun daftar kata serta kemunculannya. Daftar kemunculan kata tersebut dapat digunakan untuk mengklasifikasikan teks baru ke dalam kategori yang tepat, menurut probabilitas tertingginya [17].

#### 4.2 Support Vector Machine

Support Vector Machine (SVM) adalah salah satu metode terbaik yang sering digunakan karena tingkat akurasi nya yang baik. Metode SVM memiliki beberapa kelebihan salah satunya adalah dapat diterapkan pada data yang berdimensi tinggi. Disisi lain kekurangan yang dimiliki SVM adalah sulit digunakan untuk data yang jumlah besar. SVM bekerja dengan cara mencari sebuah *hyperplane* atau garis pembatas pemisah antar kelas yang mempunyai margin atau jarak antar *hyperplane* dengan data paling terdekat pada setiap kelas yang paling besar [16].

#### 4.3 Light Gradient Boosting Machine

Light Gradient Boosting Machine (LGBM) adalah algoritma versi Gradient Boosting lain yang menggunakan *leaf-wise algorithm* untuk menumbuhkan Decision Tree secara vertikal [1]. *Leaf-wise algorithm* ini bekerja dengan menambah cabang hingga tidak bisa dipecah lagi atau mencapai kedalaman maksimumnya. Gradient-based One-Side Sampling (GOSS) dan Exclusive Feature Bundling (EFB) digunakan untuk menangani data dan fitur dengan jumlah besar. GOSS digunakan oleh LGBM untuk mengurangi jumlah *instance* data dan menjaga akurasi untuk Decision Tree saat *training*. Hal ini dilakukan dengan menyimpan *instance* dengan gradient besar dan membuang *instance* dengan gradient kecil secara acak. Kemudian, telah dibuktikan bahwa perlakuan ini dapat menghasilkan estimasi gain yang lebih akurat daripada sampling acak seragam, dengan target *sampling rate* yang sama, terutama ketika nilai information gain memiliki rentang yang besar [7]. Sedangkan, LGBM menggunakan EFB untuk mengatasi *sparsity* dalam dataset. EFB menggabungkan fitur-fitur dan menghasilkan fitur yang lebih sedikit, namun tetap mempertahankan fitur yang penting [1].

#### 4.4 Random Forest

Random Forest (RF) adalah metode pembelajaran untuk klasifikasi, maupun regresi. RF membangun sejumlah Decision Tree saat training berlangsung. Saat mendapatkan *case* baru, setiap Decision Tree yang dibuat akan dijalankan dan menghasilkan *class / label / output*. Final output yang dipilih berdasarkan voting terbanyak atau *class* dengan jumlah output terbanyak. Pada proses *training*, RF akan mengambil data training secara acak [3].

#### 4.5 Bag Of Words

*Bag Of Words* (BOW) adalah sebuah teknik yang terkenal dalam pemrosesan teks. BOW menghasilkan *list* kata, yang sering disebut sebagai *vocabulary*, dari sebuah dataset. Setiap kalimat direpresentasikan sebagai vektor dengan setiap kata diwakili dengan nilai numerik. Misalnya, sebuah kata akan diberi nilai 1 jika kata yang muncul itu terdapat di *vocabulary* atau diberi nilai 0 jika tidak ada di *vocabulary*. Walaupun sederhana, namun BOW ini sering digunakan karena BOW masih sangat ampuh dalam memprediksi [1].

#### 4.6 Fasttext

Fasttext merupakan proyek *word embedding open-source* yang dilakukan oleh Facebook Research. Fasttext sendiri merupakan hasil studi lanjutan dari Word2Vec. Word2Vec memiliki keterbatasan dimana Word2vec ini melakukan vektorisasi kata berdasarkan *vocabulary* yang telah dipelajari saja sehingga Word2vec tidak bisa memprediksi vektor dari kata yang jarang ditemukan dan menghasilkan *Out-of-Vocabulary* (OOV). Dari sini dikembangkanlah teknik vektorisasi berbasis *character-level* yang dapat mengatasi permasalahan OOV, yaitu Fasttext. Fasttext lebih mempelajari dan mempertimbangkan struktur kata dan menggali informasi semantik kata tersebut daripada mempelajari representasi kata [2]. Oleh karena itu Fasttext memiliki kemampuan membuat vektor kata yang tidak pernah ditemukan, walaupun kata tersebut mengandung *typo*[18].

#### 4.7 Preprocessing Data

Setiap baris data yang akan masuk kedalam model akan melalui tahap-tahap persiapan, seperti *case folding*, normalisasi, penghapusan *stopword*, penghapusan duplikasi kata, dan *stemming*. Pertama adalah tahap *case folding*, *case folding* dilakukan dengan mengubah semua huruf dalam dokumen menjadi huruf kecil atau *lowercase*, menghapus angka, *emoticon*, simbol, tanda baca, dan

*whitespace*. Kemudian pada tahap normalisasi kata akan dilakukan pengubahan kata yang disingkat menjadi kata seharusnya. Aktivitas normalisasi kata ini akan dilakukan dengan cara membuat *list* yang terdiri dari kata singkatan disertai dengan kata seharusnya. Kemudian setiap kata akan dibandingkan dengan *list* yang ada, apabila ditemukan kata singkatan maka kata tersebut akan diganti dengan kata seharusnya sesuai *list*. Selanjutnya adalah proses penyaringan untuk menghilangkan kata yang tidak relevan pada kalimat teks dengan cara membandingkan *stopword list*. Selanjutnya adalah penghapusan kata duplikasi. Tahap ini akan melihat setiap kata dari sebuah kalimat, apabila kata selanjutnya adalah kata yang sebelumnya telah ada, maka kata tersebut akan dihapus. Terakhir adalah *stemming* yang dilakukan untuk menemukan kata dasar dari sebuah kata dengan menghilangkan semua imbuhan.

## 5. PENGUJIAN

### 5.1 Pengujian Parameter Tuning

Pengujian ini akan dilakukan pada SVM, RF, dan LGBM menggunakan GridSearchCV untuk mencari parameter terbaik. Sedangkan NB tidak dilakukan *parameter tuning*. Parameter yang diuji pada SVM adalah C, kernel, dan gamma. Untuk RF adalah *max\_depth*, *criterion*, *max\_features*, *min\_samples\_split*, dan *n\_estimators*. Untuk LGBM adalah *boosting\_type*, *learning\_rate*, *max\_depth*, *num\_leaves*, dan *n\_estimators*.

#### 5.1.1 Model Kuesioner

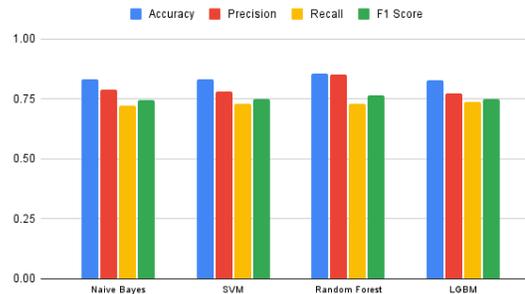
Pengujian model kuesioner ini menghasilkan kesimpulan bahwa kombinasi parameter SVM yang paling baik untuk model aspek adalah kernel RBF, nilai C sebesar 2, dan nilai gamma sebesar 0.1, serta menghasilkan nilai *accuracy*, *recall*, *precision*, dan *f1-score* yang sama, yaitu 81%. Terdapat kenaikan sebesar 0.2% untuk *accuracy*, *recall*, *precision*, dan *f1-score* pada model aspek. Untuk RF menghasilkan performa paling baik pada hasil GridSearchCV dengan nilai *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 82.4%, 83.3%, 82.4% dan 82.5%. Kemudian LGBM juga menghasilkan performa paling baik dengan kenaikan 0.5% untuk *accuracy*, 0.9% untuk *precision*, 0.1% untuk *recall*, dan 0.3% untuk *f1-score* dari model yang menggunakan *default parameters*. Untuk NB dengan parameter *defaultnya* menghasilkan nilai *accuracy* 81.8%, *precision* 81.8%, *recall* 82%, dan *f1-score* 82%. Keseluruhan hasil pengujian dapat dilihat di Gambar 1. Secara keseluruhan, RF mengungguli tiga *classifier* lainnya.



Gambar 1. Pengujian Parameter pada Model Kuesioner untuk Aspek

Sedangkan untuk model sentimen, dapat dilihat pada Gambar 2 bahwa SVM mengalami peningkatan performa diiringi dengan kenaikan 1.98% untuk *accuracy*, penurunan 0.7% untuk *precision*, kenaikan 7.8% untuk *recall*, dan kenaikan 7.2% untuk *f1-score*. Pengujian parameter RF untuk model sentimen juga menunjukkan hasil yang lebih baik dengan parameter *max\_depth* sebesar 60, *min\_samples\_split* sebesar 10, dan *n\_estimators* sebesar 200

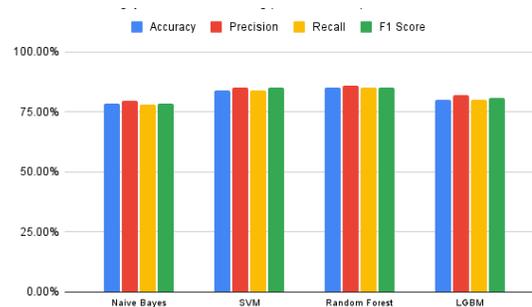
dengan nilai *accuracy* 85.4%, *precision* 85.1%, *recall* 73%, dan *f1-score* 76.5%. Sedangkan LGBM menghasilkan performa terbaiknya dengan nilai *accuracy* 82.9%, *precision* 77.3%, *recall* 73.6%, dan *f1-score* 75.1%. Untuk Naive Bayes dengan parameter *defaultnya* menghasilkan nilai *accuracy* 83.3%, *precision* 78.7%, *recall* 72.3%, dan *f1-score* 74.6%. Perbandingan hasil pengujian dapat dilihat pada Gambar 2.



Gambar 2. Pengujian Parameter pada Model Kuesioner untuk Sentimen

#### 5.1.2 Model Media Sosial

Pada pengujian model media sosial akan menggunakan GridSearchCV dengan rasio *split dataset* sebesar 80:20. Berdasarkan hasil eksperimen untuk model aspek dataset media sosial pada Gambar 3, dapat dilihat bahwa percobaan SVM berperforma baik dengan nilai *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 84%, 85%, 84%, dan 84%. Begitu pula yang terjadi pada RF, hasil uji coba yang paling baik menghasilkan nilai *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 85%, 86%, 85%, dan 85%. Percobaan LGBM juga menunjukkan nilai evaluasi model sebesar 80%, 82%, 80%, dan 81% untuk masing-masing nilai *accuracy*, *precision*, *recall*, dan *f1-score*. Untuk NB dengan parameter *defaultnya* menghasilkan nilai *accuracy* 78.5%, *precision* 79.8%, *recall* 78.1%, dan *f1-score* 78.3%. Hasil uji coba parameter pada model media sosial untuk aspek dapat dilihat di Gambar 3.



Gambar 3. Pengujian Parameter pada Model Media Sosial untuk Aspek

Untuk model sentimen, parameter terbaik yang dihasilkan GridSearchCV untuk SVM adalah kernel RBF, nilai C sebesar 10, dan gamma sebesar 0.1. Dengan parameter yang disebutkan, SVM dapat menghasilkan nilai *accuracy* sebesar 85%, *precision* sebesar 79%, *recall* sebesar 79%, dan *f1-score* sebesar 79%. Sedangkan RF menunjukkan nilai evaluasi model terbaiknya sebesar 87% untuk *accuracy*, 85% untuk *precision*, 76% untuk *recall*, dan 79% untuk *f1-score*. Parameter yang digunakan adalah *max\_depth* sebesar 40. LGBM pun menunjukkan respon positif dengan kenaikan performa dengan nilai 87%, 84%, 77%, 80% untuk *accuracy*, *precision*, *recall*, dan *f1-score*. Parameter yang digunakan LGBM adalah

boosting type 'dart', learning\_rate sebesar 0.8, max\_depth sebesar 10, dan num\_leaves sebesar 8. Untuk NB dengan parameter defaultnya menghasilkan nilai accuracy 84.8%, precision 80.25%, recall 76.1%, dan f1-score 77.7%. Keseluruhan hasil pengujian dapat dilihat pada Gambar 4.



Gambar 4. Pengujian Parameter pada Model Media Sosial untuk Sentimen

## 5.2 Pengujian Undersampling

Melihat hasil pengujian sebelumnya, pada pengujian model sentimen terdapat perbedaan nilai precision dan recall yang cukup jauh. Hal ini dikarenakan imbalance dataset. Untuk mengatasi masalah ini, maka pengujian Undersampling ini dilakukan. Namun karena dataset media sosial terlalu sedikit, maka pengujian ini tidak akan dilakukan pada model media sosial. Karena akan mempengaruhi classifier dimana terlalu banyak informasi untuk data positif yang terbuang untuk menyeimbangkan jumlah data positif dan negatif. Pengujian teknik Undersampling ini akan dilakukan pada model dataset kuesioner saja dengan menggunakan metode Random Undersampling [9][13]. Pengujian ini menggunakan 50:50 sebagai rasio distribusi kelas setelah sampling berdasarkan rekomendasi penelitian sebelumnya [9]. Hasil penerapan RUS ini dibandingkan dengan model tanpa penerapan undersampling (None). Hasil pengujian undersampling dapat dilihat pada Tabel 1.

Tabel 1. Pengujian Teknik Undersampling

Sampling	Accuracy	Precision	Recall	F1-Score
<b>NB</b>				
None	83.3%	78.7%	72.3%	74.6%
RUS	76.3%	70.4%	75.3%	71.6%
<b>SVM</b>				
None	83%	78%	73%	75%
RUS	83.5%	77.7%	76.2%	76.9%
<b>RF</b>				
None	85.4%	85.1%	73%	76.5%
RUS	84.5%	78.9%	83.2%	80.5%
<b>LGBM</b>				
None	83%	78%	72%	74%
RUS	79.5%	73.3%	77.9%	74.8%

Setelah diterapkannya Undersampling pada model media sosial untuk sentimen, dapat dilihat bahwa dari empat classifier yang digunakan, Naive Bayes terlihat tidak merespon pengaruh

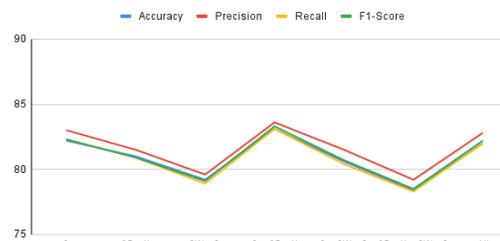
Undersampling dengan baik. Hanya nilai recall yang meningkat sebesar 3% namun terjadi penurunan untuk accuracy, precision, dan f1-score. Sedangkan SVM ini merespon baik dengan meningkatnya nilai accuracy, recall, dan f1-score. Walaupun nilai precision menurun, tapi bisa dilihat bahwa penurunan precision pada SVM tidak begitu signifikan, hanya sekitar 0.3%. Hal ini membuat juga nilai evaluasi model pada SVM lebih merata dan maksimal. Walaupun RF dan LGBM dan merespon penerapan undersampling ini dengan menurunnya nilai accuracy dan precision, namun nilai recall dan f1-score juga meningkat setelah Undersampling. Hasil akhirnya menunjukkan nilai evaluasi model yang lebih seimbang. Dapat ditarik kesimpulan bahwa RF dan LGBM masih merespon positif penerapan Undersampling.

## 5.3 Pengujian Preprocessing Data

Pada pengujian ini akan dilakukan eksperimen berupa penggantian tahap-tahap preprocessing data dan melihat tahap preprocessing data mana saja yang membantu model bekerja secara maksimal.

### 5.3.1 Model Kuesioner

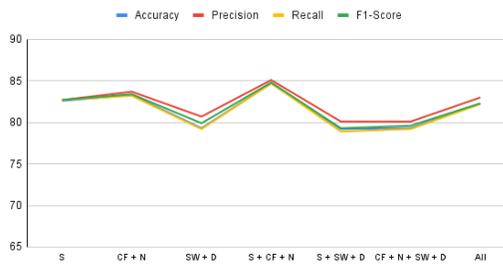
Hasil pengujian ini untuk model aspek dapat dilihat bahwa NB, SVM, RF, dan LGBM menunjukkan evaluasi model terbaik dengan stemming, case folding, dan normalisasi. Namun penggunaan kombinasi case folding + normalisasi, penghapusan stopwords + penghapusan duplikasi kata, stemming + penghapusan stopwords + penghapusan duplikasi kata, dan case folding + normalisasi + hapus stopwords + hapus duplikasi kata menurunkan performa semua classifier jika dibandingkan semua preprocessing data. Sedangkan dengan stemming saja, nilai evaluasi model tiap classifier bisa menyamai semua preprocessing data. Ini membuktikan bahwa tahapan stemming penting. Case folding dan normalisasi membantu menghapus simbol-simbol dan juga mengembalikan kata singkatan ke kata aslinya, sehingga perpaduan stemming + case folding + normalisasi lah yang terbaik. Sedangkan penghapusan stopwords dan penghapusan duplikasi tidak menyumbangkan hasil yang signifikan. Hal disebabkan karena ada kata penting yang mungkin dianggap sebagai stopwords dan terhapus. Berikut adalah keterangan singkatan pada Gambar 5, Gambar 6, Gambar 7, dan Gambar 8, yaitu S = Stemming, CF = Case Folding, N = Normalisasi, SW = Hapus Stopword, D = Hapus Duplikasi



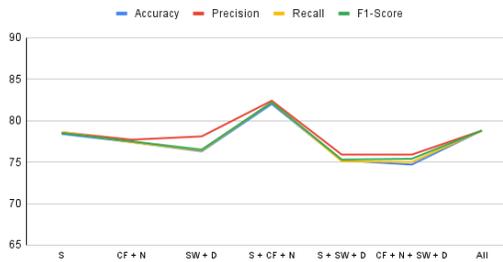
Gambar 5. Pengujian Preprocessing Data pada Model Naive Bayes untuk Aspek Kuesioner



Gambar 6. Pengujian Preprocessing Data pada Model SVM untuk Aspek Kuesioner

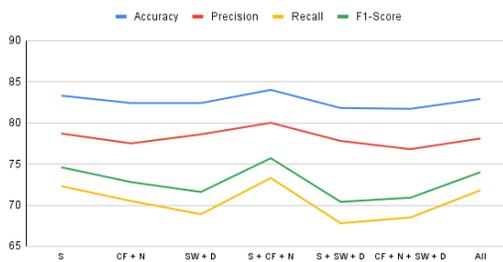


Gambar 7. Pengujian *Preprocessing Data* pada Model Random Forest untuk Aspek Kuesioner

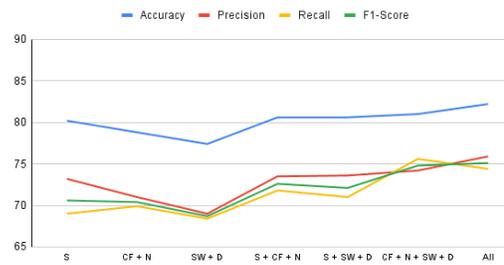


Gambar 8. Pengujian *Preprocessing Data* pada Model LGBM untuk Aspek Kuesioner

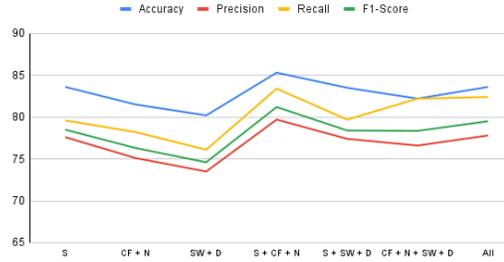
Hasil uji coba untuk model sentimen menunjukkan bahwa kombinasi *preprocessing data* yang baik adalah *stemming*, *case folding*, dan normalisasi, dan semua tahap *preprocessing data*. Pada *classifier* NB, *stemming* sangat berperan penting dalam meningkatkan performanya, karena penggunaan *stemming* saja bisa menghasilkan nilai evaluasi model yang sedikit lebih baik daripada semua *preprocessing data*. Juga kombinasi yang tidak ada *stemming* justru menurunkan performa NB. Namun kombinasi *stemming* + penghapusan *stopword* + penghapusan duplikasi kata menurunkan performa NB walaupun terdapat *stemming*. Hal ini dapat karena kemungkinan ada kata penting terdeteksi sebagai *stopword* dan terhapus. *Stemming* juga penting bagi RF dan SVM. Namun kombinasi *stemming* juga perlu diperhatikan. Kombinasi terbaik adalah *case folding* + normalisasi jika melihat Gambar 9. Apabila membandingkan *stemming* + penghapusan *stopword* + penghapusan duplikasi dan yang tanpa *stemming* pada SVM, maka kombinasi dengan *stemming* masih lebih baik. LGBM juga menghasilkan performa baik pada penggunaan *stemming* saja. Selain itu, percobaan kombinasi semua *preprocessing data* kecuali *stemming* juga tidak memberikan nilai performa yang lebih baik. Hasil pengujian dapat dilihat di Gambar 9, Gambar 10, Gambar 11, dan Gambar 12.



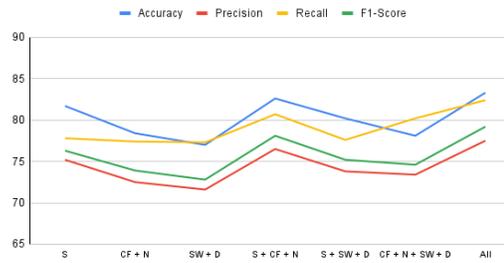
Gambar 9. Pengujian *Preprocessing Data* pada Model Naïve Bayes untuk Sentimen Kuesioner



Gambar 10. Pengujian *Preprocessing Data* pada Model SVM untuk Sentimen Kuesioner



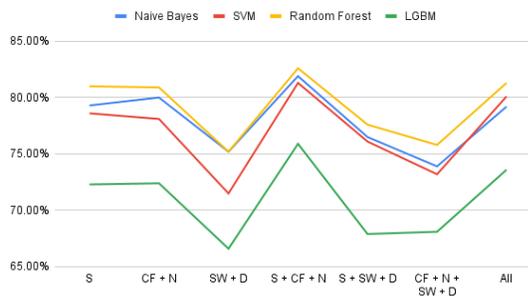
Gambar 11. Pengujian *Preprocessing Data* pada Model Random Forest untuk Sentimen Kuesioner



Gambar 12. Pengujian *Preprocessing Data* pada Model LGBM untuk Sentimen Kuesioner

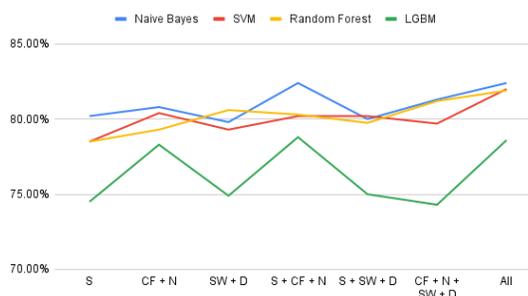
### 5.3.2 Model Media Sosial

Pengujian *preprocessing data* ini menggunakan K-Fold Cross Validation dengan jumlah *fold* yang digunakan adalah 10. Gambar 13 dan Gambar 14 mencantumkan rata-rata *accuracy* dari semua *fold*. Untuk model aspek, dapat dilihat pada Gambar 13 dimana *classifier* menghasilkan nilai rata-rata akurasi tertinggi dengan *stemming*, *case folding*, dan normalisasi. Terlihat juga bahwa penghapusan *stopword* dan penghapusan duplikasi kata cenderung menghasilkan nilai evaluasi model yang kurang bagus. Hal ini disebabkan oleh tidak adanya proses normalisasi kata, sehingga kata yang seharusnya melalui *stemming* tidak terdeteksi. Selain itu, tidak ada tahap *case folding* yang menghapus simbol dan penghapusan *stopword* yang kurang tepat. Kata yang terdeteksi sebagai *stopword* mungkin adalah kata penting. Pada NB, RF, dan LGBM, memperlihatkan kembali seberapa penting tahap *stemming* dilakukan. Pada percobaan *stemming* saja, ketiga *classifier* ini dapat menghasilkan rata-rata akurasi yang menyaingi rata-rata akurasi model seluruh tahap *preprocessing*. Maka, hal ini juga menjelaskan mengapa *stemming* + *case folding* + normalisasi paling baik. Berikut ini keterangan singkatan pada Gambar 13 dan Gambar 14, S = *Stemming*, CF = *Case Folding*, N= Normalisasi, SW = Hapus *Stopword*, D= Hapus Duplikasi



**Gambar 13. Rata-rata Akurasi Pengujian Preprocessing Data pada Model Media Sosial untuk Aspek**

Pada hasil eksperimen pada model sentimen untuk dataset media sosial ini, kombinasi *preprocessing data* yang dapat membantu keempat *classifier* adalah seluruh tahap *preprocessing data*. Uniknya, penggunaan penghapusan *stopword* + penghapusan duplikasi kata yang cukup dihindari pada beberapa eksperimen sebelumnya, namun pada percobaan model sentimen ini RF menghasilkan performa yang cukup baik. Walaupun bukan yang terbaik, namun penggunaan penghapusan *stopword* + penghapusan duplikasi kata lebih unggul daripada beberapa kombinasi *preprocessing data* lainnya. Keempat *classifier* menunjukkan performa yang kurang stabil pada percobaan *preprocessing data*. Hal ini dapat disebabkan karena persebaran data data positif dan negatif tidak merata. Hasil percobaan *preprocessing data* pada model media sosial untuk sentimen dapat dilihat pada Gambar 14.



**Gambar 14. Rata-rata Akurasi Pengujian Preprocessing Data pada Model Media Sosial untuk Sentimen**

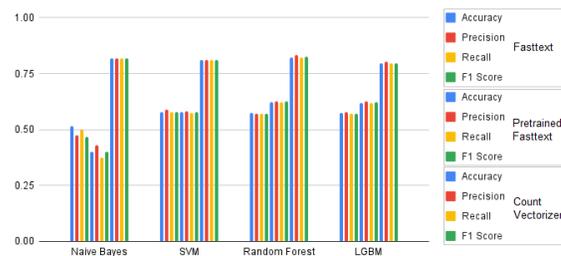
## 5.4 Pengujian Fasttext

Pengujian ini membandingkan kinerja BOW dan FastText. Terdapat 2 jenis model FastText yang akan digunakan, yaitu model FastText akan dilatih dataset sendiri dan model *pre-trained* FastText Bahasa Indonesia yang dilatih lagi dengan dataset sendiri.

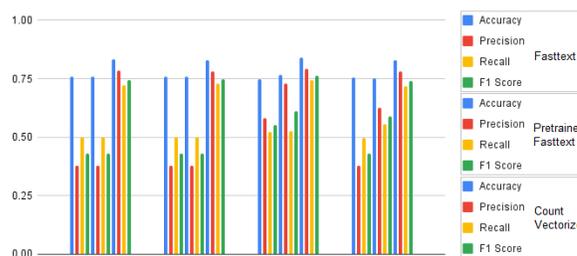
### 5.4.1 Model Kuesioner

Apabila melihat Gambar 15 dan Gambar 16 dapat dilihat bahwa memang BOW masih jauh lebih unggul daripada menggunakan Fasttext ataupun *Pretrained* Fasttext. Namun penggunaan model *pretrained* Fasttext berhasil menghasilkan nilai evaluasi model yang lebih baik daripada model Fasttext yang dilatih sendiri. Namun penggunaan model Fasttext ini terlihat hanya bekerja dengan baik pada *classifier* RF dan LGBM. RF dan LGBM menghasilkan nilai evaluasi model yang jauh lebih baik apabila menggunakan *pretrained* model pada model aspek maupun sentimen. Namun pada *classifier* NB dan SVM justru terjadi penurunan performa saat menggunakan *pretrained model* pada model aspek. Pada model sentimen, performa NB dan SVM cenderung tidak naik ataupun tidak turun. Pada model Fasttext yang dilatih sendiri, RF menduduki peringkat teratas untuk model aspek

dan sentimen dengan nilai *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 57.3%, 57.1%, 57.2%, 57.19% dan 74.7%, 58.3%, 52.3%, 55.1%. Penggunaan *pretrained* Fasttext dan RF pada model aspek menghasilkan nilai *accuracy*, *precision*, *recall*, dan *f1-score* tertinggi, yaitu 62.3%, 62.7%, 62.3%, dan 62.5%. Begitu pula untuk model sentimen, RF menghasilkan nilai *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 76.7%, 73%, 52.8%, dan 61.3%. Dari hasil eksperimen penerapan Fasttext dan *pretrained* Fasttext ini membuktikan bahwa penggunaan *pretrained* Fasttext lebih baik dalam meningkatkan kinerja *classifier*. Perbandingan teknik vektorisasi bisa dilihat pada Gambar 15 dan Gambar 16.



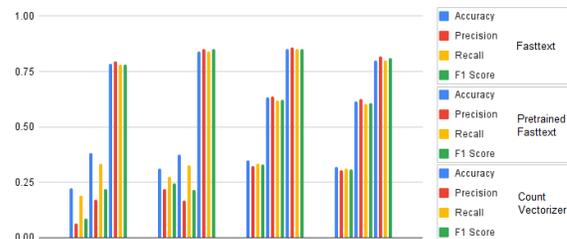
**Gambar 15. Implementasi Fasttext dan Pretrained Fasttext Model Aspek Kuesioner**



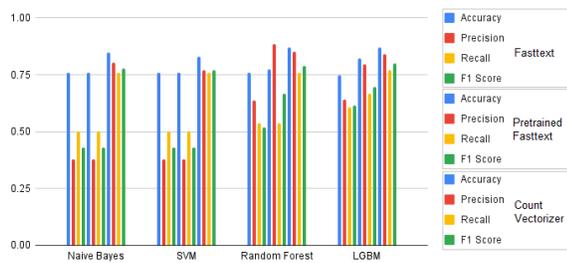
**Gambar 16. Implementasi Fasttext dan Pretrained Fasttext Model Sentimen Kuesioner**

### 5.4.2 Model Media Sosial

Pada hasil pengujian model media sosial ini menginformasikan bahwa *pretrained* Fasttext model mampu meningkatkan performa *classifier* RF dan LGBM pada model aspek dan sentimen. NB dan SVM menghasilkan nilai evaluasi model yang sama untuk model sentimen pada penerapan model Fasttext dan *pretrained* Fasttext. Hanya sedikit kenaikan nilai pada penerapan model Fasttext ke *pretrained* Fasttext untuk model aspek. Pada penerapan model Fasttext sendiri, LGBM lebih unggul untuk klasifikasi sentimen dengan nilai *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 75%, 64%, 60.7%, 61.7%. Untuk model aspek, semua *classifier* gagal memprediksi. Pada penerapan *pretrained* Fasttext, RF unggul pada klasifikasi aspek dengan nilai evaluasi model sebesar 63.3%, 63.9%, 61.9%, 62.1%. Pada model sentimen, LGBM mengungguli dengan nilai evaluasi model sebesar 82.1%, 79.8%, 66.7%, dan 69.5%. Hasil uji coba bisa dilihat pada Gambar 17 dan Gambar 18.



**Gambar 17. Hasil Implementasi Fasttext dan Pretrained Fasttext Model Aspek Media Sosial**



**Gambar 18. Hasil Implementasi Fasttext dan Pretrained Fasttext Model Aspek Media Sosial**

## 6. KESIMPULAN

Secara keseluruhan, model terbaik untuk melakukan klasifikasi aspek dan sentimen pada dua jenis dataset adalah RF. Walaupun begitu, NB dan SVM juga bersaing kuat dengan RF dalam klasifikasi. Namun ternyata LGBM ini kurang cocok untuk melakukan klasifikasi pada dataset kecil. Untuk *preprocessing data* yang terbaik berdasarkan pengujian yang dilakukan adalah *stemming + case folding + normalisasi* dan semua *preprocessing data*. Selain itu, terlihat juga bahwa model kuesioner mampu memprediksi dataset media sosial dengan baik, namun nilai evaluasi modelnya masih kalah dengan model media sosial sendiri. Hasil analisa data opini terhadap pelayanan vaksinasi COVID-19 bahwa para responden memberikan respon positif pada pelayanan vaksinasi yang diberikan. Namun dari semua aspek, aspek protokol kesehatan dan waktu lah yang perlu diperhatikan. Selain itu, terdapat juga kenaikan kualitas pelayanan vaksinasi dengan kenaikan sentimen positif dan penurunan sentimen negatif pada dosis 2 vaksinasi. Pada dosis pertama, terdapat sejumlah 1227 data positif dari seluruh aspek dan 293 data negatif. Sedangkan dosis kedua memiliki 1267 data positif dan 198 data negatif.

## 7. SARAN

Untuk penelitian selanjutnya, bisa melakukan teknik pengumpulan data yang lebih merata untuk setiap kelasnya, serta mengumpulkan data dalam skala yang lebih besar. Kemudian, memperluas variabel kelas aspek ataupun sentimen mengenai pelayanan. Untuk mengatasi *imbalance dataset* bisa mencoba teknik *undersampling* lainnya, dan mencoba Fasttext pada dataset yang lebih besar, serta mencoba *classifier* yang disediakan langsung oleh Fasttext.

## 8. REFERENSI

- [1] Alzamzami, F., Hoda, M., & Saddik, A. E. 2020. Light Gradient Boosting Machine for General Sentiment Classification on Short Texts: A Comparative Evaluation. Retrieved from <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9099543>.
- [2] Athiwaratkun, B., Wilson, A. G., & Anandkumar, A. 2018. Probabilistic FastText Multi-Sense Word Embeddings. Retrieved from <https://arxiv.org/pdf/1806.02901.pdf>.
- [3] Baid, P., Gupta, A., & Chaplot, N. 2017. Sentiment Analysis of Movie Reviews using Machine Learning Techniques. From [https://www.researchgate.net/Sentiment\\_Analysis\\_of\\_Review\\_using\\_Machine\\_Learning\\_Techniques/](https://www.researchgate.net/Sentiment_Analysis_of_Review_using_Machine_Learning_Techniques/)
- [4] Charlyn, et al. 2021. Twitter Sentiment Analysis towards COVID-19 Vaccines in the Philippines Using Naive Bayes. <https://doi.org/10.3390/info12050204>.
- [5] NN. 2021. Cara Memperbaiki Data Sertifikat Vaksin yang Salah, Jangan Panik Dulu!. Retrieved from <https://news.detik.com/berita/d-5721350/cara-memperbaiki-data-sertifikat-vaksin-yang-salah-jangan-panik-dulu>.
- [6] Dey, et al. 2016. Sentiment Analysis of Review Datasets using Naive Bayes' and K-NN Classifier. <https://doi.org/10.5815/ijieeb.2016.04.07>.
- [7] Ke, G., et al. 2017. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. Retrieved from <https://papers.nips.cc/paper/2017/file/6449f44a102-Paper.pdf>
- [8] Khanvilkar & Vora. 2018. *Sentiment Analysis for Product Recommendation Using Random Forest*. Retrieved from <https://www.researchgate.net/profile/Sentiment-Analysis-for-Product-Recommendation-Using-Random-Forest.pdf>
- [9] Mountassir, A., Benbrahim, H. & Berrada, I. 2012. An empirical study to address the problem of Unbalanced Data Sets in sentiment classification. IEEE 2012 IEEE International Conference. doi:10.1109/ICS MC.2012.6378300
- [10] Musyaddad, A. A. 2021. Ombudsman Temukan Sejumlah Masalah Vaksinasi Covid di Surabaya. Retrieved from <https://ombudsman.go.id/pengumuman/r/artikel--ombudsman--temukan-sejumlah-masalah-vaksinasi-covid-di-surabaya>.
- [11] Priyavrat, & Singh, A. J. 2017. Sentiment Analysis: A Comparative Study of Supervised Machine Learning Algorithms Using Rapid miner. Int. J. Res. Appl. Sci. Eng. Technol., vol. 5, pp. 80–89.
- [12] Pristiyono, et al. 2020. *Sentiment analysis of COVID-19 vaccine in Indonesia using Naïve Bayes Algorithm*. IOP Conference Series: Materials Science and Engineering. Retrieved from <https://iopscience.iop.org/article/10.1088/1757-899X/1088/1/012045/meta>
- [13] Prusa, J., Khoshgoftaar, T. M., Dittman, D. J. & Napolitano, A. 2015. Using Random Undersampling to Alleviate Class Imbalance on Tweet Sentiment Data. IEEE 2015 IEEE International Conference. doi:10.1109/IRI.2015.39
- [14] Salman, Ghinan. 2021. Punya Keluhan Soal Sertifikat Vaksin COVID-19, Warga Surabaya Bisa Laporkan ke Layanan Ini. Retrieved from <https://regional.kompas.com/read/2021/08/27/203228078/punya-keluhan-soal-sertifikat-vaksin-covid-19-warga-surabaya-bisa-lapor-ke>.
- [15] Scikit-learn Developers. 2021. MultinomialNB. Retrieved from [https://scikit-learn.org/stable/modules/generated/sklearn.naive\\_bayes.MultinomialNB.html](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html)
- [16] Somantri, O., & Apriliani, D. 2018. Support Vector Machine Berbasis Feature Selection Untuk Sentiment Analysis Kepuasan Pelanggan Terhadap Pelayanan Warung Dan Restoran Kuliner Kota Tegal. <https://doi.org/10.25126/jtiik20185867>
- [17] Ting, S.L, Ip, W. H., & Tsang, A. H. C. 2011. Is Naive Bayes a Good Classifier for Document Classification?. Retrieved from <https://www.researchgate.net/Is-Naive-Bayes-a-Good-Classifier-for-Document-Classification.pdf>
- [18] Young, J. C., & Rusli, A. 2019. Review and Visualization of Facebook's FastText Pretrained Word Vector Model. doi:10.1109/icesi.2019.8863015