

## Analisis Sentimen Review Film pada Twitter menggunakan Metode Klasifikasi Hybrid SVM, Naïve Bayes, dan Decision Tree

Mochammad Alfi Rizky Reynaldhi<sup>1</sup>, Yuliant Sibaroni<sup>2</sup>

<sup>1,2,3</sup> Universitas Telkom, Bandung

<sup>1</sup>itsmocha@students.telkomuniversity.ac.id, <sup>2</sup>yuliant@telkomuniversity.ac.id

---

### Abstrak

Sebuah film dapat dikatakan memiliki keberhasilan atau kegagalan bergantung pada ulasan yang didapat. Di masa seperti ini, terdapat peningkatan jumlah penonton pada film yang ditayangkan secara daring dan secara tidak langsung meningkat juga ulasan yang didapat pada sebuah film. Oleh karena itu penelitian ini bertujuan untuk membangun model analisis sentimen yang baik untuk melakukan klasifikasi ulasan pada film untuk memberikan label positif dan negatif. Metode klasifikasi yang digunakan ialah Hybrid yang menggabungkan algoritma SVM karena dapat dimanfaatkan untuk *high dimensional space*, Naïve Bayes yang mampu mengolah data dalam jumlah besar dengan akurasi yang tinggi, dan Decision Tree karena mampu mempertimbangkan seluruh kemungkinan hingga berujung sebuah kesimpulan. Ketiga metode tersebut dikombinasikan menggunakan metode *ensemble voting*. Berdasarkan hasil evaluasi didapatkan bahwa akurasi yang diperoleh 0.8192 dan F1-score 0.78. Hasil penelitian ini menunjukkan bahwa algoritma klasifikasi *Hybrid* dapat membantu dalam mengklasifikasikan ulasan film.

**Kata kunci :** Hybrid, SVM, Naïve Bayes, Decision Tree, Analisis Sentimen, Film

---

### Abstract

A movie can be said to have success or failure depending on the reviews it gets. Nowadays, the amount of movies streaming viewers increases and indirectly increases the reviews of the movies. Therefore, this study aims to build a good sentiment analysis model for classifying movie reviews to give positive and negative labels. The classification method used is Hybrid which combines the SVM algorithm because it can be used for high dimensional space, Naïve Bayes which is able to process large amounts of data with high accuracy, and Decision Tree because it is able to consider all possibilities to lead to a conclusion. The three methods are combined using the ensemble voting method. Based on the evaluation results, it was found that the accuracy obtained was 0.8192 and the F1-score was 0.78. The results of this study indicate that the Hybrid classification algorithm can assist in classifying film reviews.

**Keywords:** Hybrid, SVM, Naïve Bayes, Decision Tree, Sentiment Analysis, Movie

---

## 1. Pendahuluan

### Latar Belakang

Keberhasilan atau kegagalan sebuah film bergantung pada ulasan yang didapatkan, dimana terdapat ribuan ulasan yang mempengaruhi *rating* dari sebuah film. Pada tahun 2016 misalnya, menurut Motion Picture of America Dodd yang mengatakan bahwa rata-rata enam ratus film diproduksi setiap tahunnya dan akan terus bertambah [1]. Bisa dibayangkan berapa film yang telah ada saat ini dengan jumlah ulasan pada setiap filmnya, ditambah lagi kemudahan menonton film dengan menggunakan layanan *streaming* yang tersedia. Selain itu, film bukan hanya merupakan sumber hiburan atau *entertainment*, tetapi juga merupakan sumber utama dari pemasaran global. Ketertarikan masyarakat pada suatu film dapat diperoleh dari pendapat maupun opini yang dibagikan pada salah satu media sosial yang populer yaitu Twitter [2].

Twitter merupakan media yang mewadahi ribuan sampai jutaan opini yang disebut dengan *tweet* dan dapat dimanfaatkan untuk mengumpulkan opini masyarakat terhadap suatu film. Untuk mengetahui pengaruh dari opini tersebut, maka dilakukannya analisis sentimen pada opini yang telah didapatkan. Analisis sentimen memiliki fokus utama yaitu untuk memprediksi polaritas suatu opini yang diungkapkan dalam suatu bahasa yang kemudian akan diklasifikasikan ke dalam kategori positif atau negatif. Dengan demikian sebuah film dapat dikatakan berhasil atau gagal dari opini yang telah diklasifikasikan tersebut[3,4].

Pada penelitian sebelumnya, peneliti ingin membuat beberapa model klasifikasi untuk memprediksi keberhasilan film *box office* dimana hasil akurasi yang didapatkan setiap model kemudian dibandingkan antara satu dengan yang lainnya, diantara lain adalah Naïve Bayes 59.07%, SVM 60.39%, KNN 60.76%, dan Random Forest 22.97% [5]. Oleh sebab itu, pada penelitian ini berfokus untuk membangun model klasifikasi yang optimal dengan menggabungkan algoritma SVM, Naïve Bayes, Decision Tree menjadi Hybrid. Metode Hybrid

ini juga merupakan *gap* dari penelitian sebelumnya. Selain itu, penelitian ini juga berfokus untuk menghasilkan model yang dapat mengklasifikasikan opini atau *tweet* dengan optimal.

### Topik dan Batasannya

Berdasarkan latar belakang yang telah disampaikan sebelumnya, maka rumusan masalah dari penelitian ini adalah mengukur performansi dan akurasi dari penggabungan metode SVM, Naïve Bayes, Decision Tree secara Hybrid dalam pengklasifikasian *tweet review* film. Selain itu, untuk mengetahui bagaimana cara menghasilkan model klasifikasi yang optimal. Sedangkan untuk batasan masalah pada penelitian ini menggunakan sebanyak 1023 data *tweet* pengguna Twitter mengenai *review* film Nevertheless, Riverdale, Vincenzo, dan Never Have I Ever yang diambil pada bulan Juli 2021.

### Tujuan

Tujuan dari penelitian tugas akhir ini adalah untuk mengetahui sentimen mengenai film yang sedang populer pada tahun 2021 dan untuk membangun model analisis sentimen dan klasifikasi *review* film menggunakan metode Hybrid yang terdiri dari algoritma klasifikasi SVM, Naïve Bayes, dan Decision Tree.

### Organisasi Tulisan

Setelah ini akan dijelaskan mengenai studi terkait penelitian serupa beserta dengan hasil yang diperoleh. Lalu, pada bagian tiga akan dijelaskan system yang dibangun peneliti dan juga penjelasan mengenai gambaran beserta teori-teori terkait penelitian ini. Pada bagian empat yaitu evaluasi, akan dijelaskan mengenai hasil pengujian yang didapatkan yaitu perhitungannya akurasi, *precision*, *recall*, *F1-Score*, dan analisis dari hasil pengujian. Bagian akhir adalah kesimpulan dari hasil penelitian dan juga saran atau masukan untuk penelitian selanjutnya.

## 2. Studi Terkait

Telah dilakukan penelitian serupa mengenai analisis sentimen sebelumnya, dan beberapa diantaranya menggunakan dataset yang sudah ada salah satunya adalah dataset IMDB yang berisi *review* film. Selain menggunakan dataset yang sudah ada, beberapa penelitian sebelumnya juga menggunakan dataset yang diambil melalui media sosial salah satunya Twitter yang juga digunakan oleh peneliti untuk digunakan sebagai dataset. Bukan hanya dataset, terdapat penelitian sebelumnya yang menggunakan metode *hybrid* untuk menggabungkan beberapa metode untuk menghasilkan performa yang baik untuk klasifikasi.

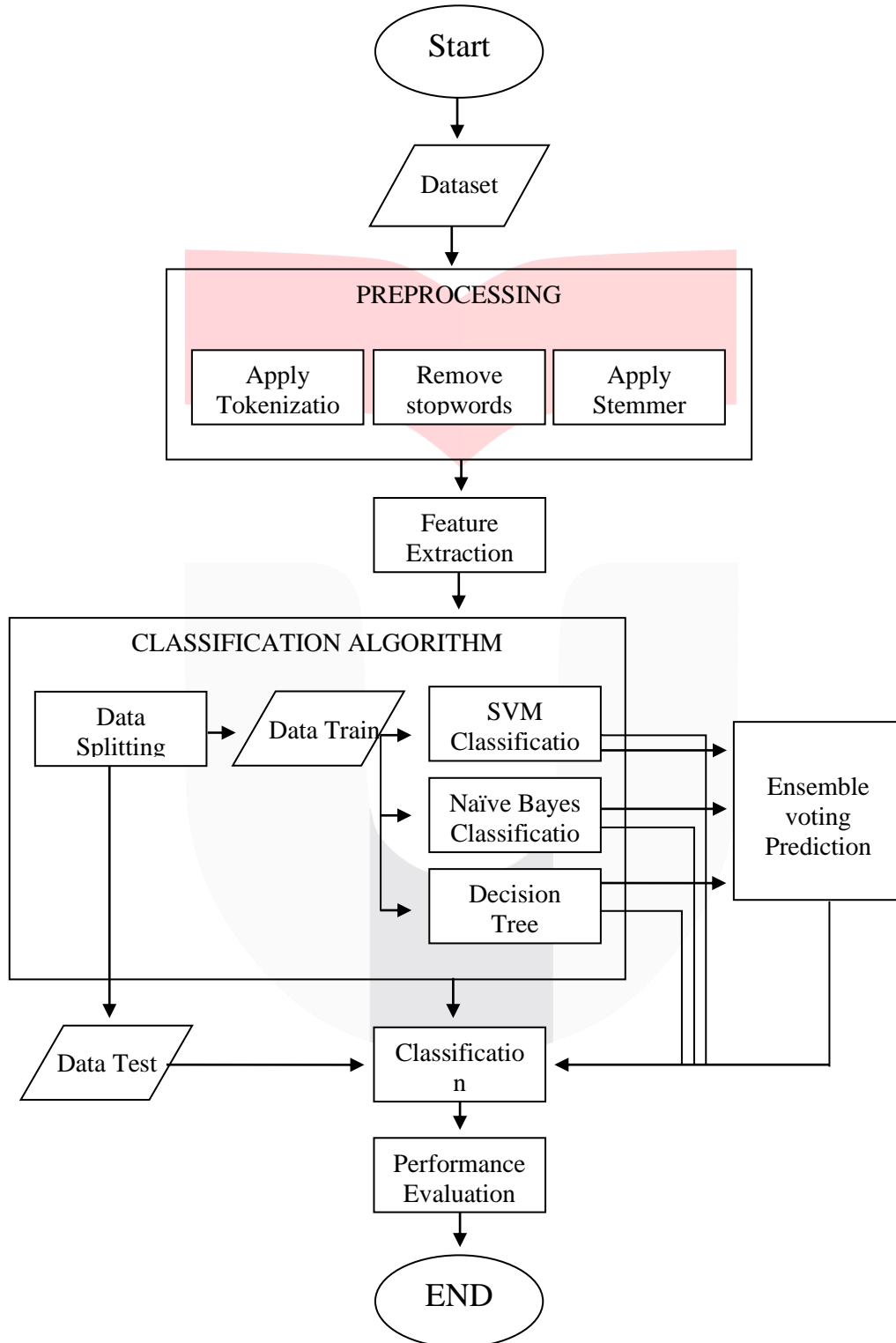
Dimulai pada penelitian [1] yang menggunakan dataset dari IMDB yang berisi mengenai *review* film, dan penelitian [9] yang menggunakan dataset dari Steam yang berisi mengenai *review* game. Kedua penelitian [1, 9] menggunakan metode yang bekerja secara individu. Penggunaan beberapa metode bertujuan untuk melakukan perbandingan performansi, khususnya untuk mengetahui metode yang memiliki akurasi terbaik. Dimana pada penelitian [9] metode Decision Tree memiliki akurasi yang baik yaitu 75% dibanding metode Naïve Bayes. Lalu pada penelitian [1] menggunakan empat metode dengan hasil akurasinya yaitu SVM 60.39%, KNN 60.76%, Naïve Bayes 59.07%, dan Random Forest 22.97%. Dengan hasil yang didapatkan, peneliti [1] menyarankan untuk mencoba mengkombinasikan beberapa metode untuk menghasilkan performansi yang baik.

Metode penggabungan yang bisa disebut Hybrid tersebut telah dilakukan dilakukan pada penelitian [7] yang menggabungkan Gradient Boosting dan Extra Trees classifier menggunakan Ensemble Voting, dimana hasilnya mendapatkan akurasi 86.5% dibanding hasil akurasi yang didapatkan metode klasifikasi Naïve Bayes 76.74% dan Decision Tree 69.77%. Lalu pada penelitian [8] juga menggunakan Ensemble Voting dengan mengambil hasil probabilitas maksimum, dengan hasil akhir yang menunjukkan bahwa kombinasi dari beberapa metode memiliki akurasi yang baik dibanding metode individu seperti IBK, J48, NBTree, One R, Zero R, Naïve Bayes, dan Decision Stump. Selain itu, metode Ensemble Voting juga memiliki performa yang lebih baik dibandingkan beberapa metode klasifikasi homogen seperti Bagging dan versi *boosting* dari metode Decision Stump. Selain itu, metode ensemble juga digunakan pada penelitian [23] dengan melakukan analisa pada tiga kernel yang ada pada SVM. Dimana masing-masing kernel yaitu SVM – linier dan RBF mampu memprediksi kelas yang salah dengan baik dan SVM-Adaboost mampu memberikan performansi yang baik pada data yang *unbalance*.

Tidak hanya metode hybrid untuk melakukan klasifikasi, penggunaan pemrosesan data tekstual untuk menentukan sebuah teks kedalam suatu sentimen positif atau negatif sebagai salah satunya adalah Textblob yang merupakan *library* yang populer dalam melakukan sentimen pada sebuah kata juga digunakan pada penelitian analisis sentiment seperti pada penelitian [24,25]. Dimana pada penelitian [24] membandingkan Textblob dengan AFINN yang juga merupakan salah satu teknik yang digunakan pada analisis sentimen. Hasilnya seluruh metode yang diuji pada penelitian [24] yang menggunakan Textblob memiliki performansi terbaik, salah satu hasilnya pada metode Logistic Regression yang menggunakan Textblob memiliki akurasi 0.95 sedangkan dengan AFINN mendapatkan akurasi 0.88.

### 3. Sistem yang Dibangun

Pada gambar 1 menjelaskan bagaimana sistem klasifikasi dibuat pada penelitian ini. Memanfaatkan beberapa contoh referensi dari penelitian sebelumnya sebagai gambaran umum untuk membuat sistem ini.



Gambar 1. Sistem Klasikasi Analisis Sentimen

### 3.1 Pengumpulan Data

Dataset pada penelitian ini merupakan hasil *crawling tweet* yang berisi opini mengenai film yang sedang populer, diantara lain adalah Riverdale, Never Have I Ever, Nevertheless, dan Vincenzo. Pengambilan data sebanyak 1023 *tweet* pada periode bulan Juli 2021, yang masing-masing berisi mengenai opini tentang keempat film tersebut.

### 3.2 Preprocessing

Terdapat tiga tahapan dalam melakukan proses *preprocessing* yaitu *Tokenization*, *Stopword Removal*, dan *Stemming*. Perlu diketahui jika pada tahap ini memanfaatkan *Natural Language Toolkit* (NLTK) yang mengidentifikasi kata berbahasa Inggris. Penjelasan tahapan-tahapan tersebut adalah sebagai berikut:

- *Tokenization*. Tokenisasi merupakan suatu hal yang penting dalam analisis *lexical*, dimana pada tahap ini akan memisahkan suatu input ke dalam sebuah frasa, kata, simbol, atau elemen lain yang disebut token [10]. Dalam proses ini, seluruh input akan diubah menjadi huruf kecil. Lalu, angka, karakter *emoticon*, dan tanda baca akan dihapus.
- *Stopword Removal*. Tahap ini berfungsi untuk menghapus suatu kata yang tidak signifikan. Kata yang sering muncul pada sebuah dokumen akan dipertimbangkan sebagai kata yang tidak signifikan karena tidak adanya statistik yang relevan. Maka kata tersebut harus dihapus [11].
- *Stemming*. Pada tahapan ini, sebuah kata yang telah dilakukan tokenisasi akan diubah menjadi kata dasar agar tidak ada redundansi. Mengasumsikan seluruh kata yang telah melewati tahap *stemming* akan berhubungan secara semantik [12].

### 3.3 Analisis Sentimen

Tahap ini menggunakan *library* pada bahasa pemrograman python yaitu *textblob* yang merupakan sebuah *library* untuk melakukan *natural language processing* pada dataset. Melakukan perhitungan pada polaritas dan subjektivitas pada suatu teks sehingga menghasilkan pelabelan analisis sentimen yang dikategorikan sebagai berikut:

- “Positif”
- “Negatif”

Sebelum melakukan pelabelan pada *tweet*, terlebih dahulu dilakukan kalkulasi atau perhitungan dari polaritas masing-masing sentimen pada *tweet*. Setiap *tweet* tersebut memiliki bobot yang terdapat pada setiap kata yang telah ditentukan oleh *library textblob* yang pada penelitian ini menggunakan *corpus* bahasa Inggris. Jika sentimen negatif lebih besar daripada sentimen positif, maka *tweet* akan mendapatkan label positif dan begitu juga sebaliknya. Contoh pelabelan sentimen dapat dilihat pada tabel 1 dimana *compound* merupakan penentu polaritas dari *tweet*. Jika *compound* > 0, maka bisa polaritas akan menjadi 1 dan menyimpulkan bahwa sentimen yang ada pada *tweet* tersebut adalah positif.

**Tabel 1. Contoh Pelabelan Sentimen**

Sentiment	Tweet	Polarity	Subjectivity	Negative	Positive	Compound
<b>Positif</b>	sunday is a perfect day where i can start watching vincenzo	1.0	1.0	<b>0.0</b>	<b>0.316</b>	0.5719
<b>Negatif</b>	i will never forgive netflix for cancelling this show but continuing freaking riverdale	0.0	0.0	<b>0.42</b>	<b>0.0</b>	-0.6712

Hasil dari pelabelan terhadap dataset yang berisi *tweet* opini mengenai film pada penelitian ini dapat dilihat pada table 2 dibawah.

**Tabel II. Contoh Pelabelan**

Sebelum <i>Preprocessing</i>	Setelah <i>Preprocessing</i>	Kategori
Y'all, what's with this shit of 'Nevertheless' ?? Did she really end up with that guy?δÿα"	shit nevertheless really end guy planning watch	Negatif
@tinnkky Billions is excellent. When youâ€™re in the mood for a comedy..never have I ever on Netflix is also very good	billions excellent mood comedy never ever netflix also good	Positif

### 3.4 Feature Extraction

Tahap ini biasanya disebut *feature extraction* atau vektorisasi dimana tahapan ini cukup penting untuk melakukan analisis pada teks. Kata-kata akan diubah ke dalam bentuk *integer* atau *float*, tergantung menggunakan algoritma apa yang diimplementasikan pada *machine learning*. Dalam penelitian ini, dua teknik yang paling sering dan populer untuk digunakan adalah Count Vectorizer dan TF-IDF Vectorizer [13].

Pada penelitian ini, peneliti menggunakan Count Vectorizer yang salah satu teknik yang sederhana dan memiliki hasil yang relatif memuaskan. Dimana menurut penelitian [15], cara kerja dari teknik ini adalah dengan menghitung berapa kali sebuah kata muncul dalam dokumen dan menggunakan nilai satuan sebagai bobotnya atau diubah dalam tokenisasi. Selain itu alasan lain peneliti menggunakan teknik ini karena mampu mengembalikan kata-kata dalam bentuk bilangan bulat dan juga menjadi alasan tidak menggunakan teknik TF-IDF Vectorizer karena mengembalikan bilangan float.

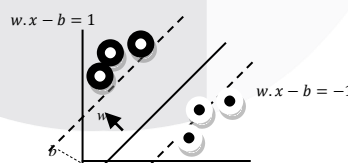
### 3.5 Model Algoritma

Dalam membuat sistem klasifikasi ini, data yang dimiliki harus dibagi menjadi dua yaitu data latih dan data test. Hal tersebut dapat dimanfaatkan untuk melatih sistem klasifikasi dengan menggunakan metode yang dipilih untuk membangun metode hybrid ini adalah yaitu SVM, Naïve Bayes, dan Decision Tree. Sistem klasifikasi tersebut akan dilatih sehingga dapat menghasilkan *hyperplane* dan dapat melabeli kelas "Negatif" dan "Positif" menggunakan bahasa pemrograman Python.

- Klasifikasi Support Vector Machine (SVM)

Metode ini digunakan karena berhasil diimplementasikan pada beberapa penelitian [1,6,7] dalam pengklasifikasian sentimen. SVM menggunakan persamaan klasifikasi berikut untuk pengklasifikasian.

$$f(\vec{x}) = \text{sign}(\vec{w}^t \cdot \vec{x} + b) \quad (1)$$



**Gambar 2. Hyperplane pada SVM**

$w$  mewakili vektor bobot yang menunjukkan pentingnya fitur-fitur tersebut.  $x$  merupakan input yang diberikan ke pengklasifikasi dan juga bias [16]. Algoritma melakukan identifikasi dan mempertimbangkan fitur mana yang lebih penting dan mencoba untuk menyesuaikan bobot. Ketika sebuah input masuk, maka algoritma akan mengonversi input tersebut ke dalam ruang fitur pelatihan dan kemudian menghitung perkalian antara vektor bobot dan input. Pada analisis sentimen jika sebuah persamaan adalah positif, maka algoritma akan mengeluarkan positif dan sebaliknya jika negatif. SVM akan menghasilkan *hyperplane* yang optimal saat memisahkan kedua kategori.

- Klasifikasi Naïve Bayes

Naïve Bayes merupakan algoritma pengklasifikasian model linier yang dikategorikan sederhana namun sangat efisien dalam memberikan pelabelan. Metode ini mengasumsikan pengaruh nilai atribut pada kelas

tertentu untuk tidak bergantung pada atribut lainnya. Asumsi pada algoritma ini dinamai class-conditional independent yang menjadikan metode klasifikasi ini dianggap naif. Teorema dasar metode Bayes ditulis sebagai berikut [12]:

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)} \quad (2)$$

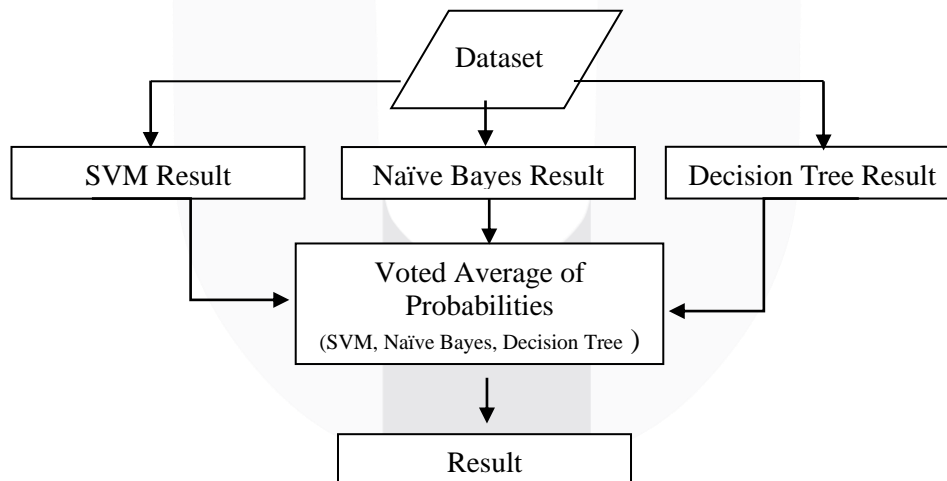
Dimana  $P(B|A)$  merupakan probabilitas dari dokumen(teks) yang dimiliki oleh kelas B [21].

- Klasifikasi Decision Tree

Merupakan metode berbasis pohon yang digunakan untuk membantu klasifikasi. Metode ini membutuhkan bagging dan boosting untuk meningkatkan akurasi karena cenderung over fit [14]. Struktur pada algoritma ini seperti sebuah pohon, dimana terdapat *root node*, *child node*, dan *leaf* sebagai tempat untuk menampilkan label suatu data.

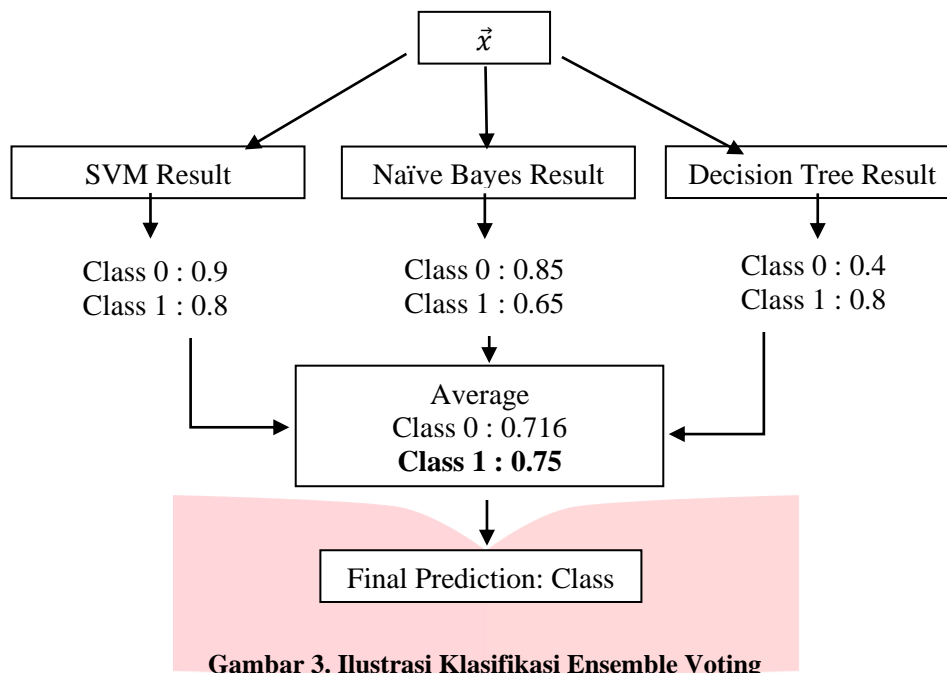
- Klasifikasi Hybrid – Ensemble Voting

Hybrid merupakan sebuah metode yang digunakan untuk mengombinasikan berbagai macam metodologi seperti mengombinasikan beberapa sistem klasifikasi [19]. Metode ensemble adalah teknik pembelajaran yang menggunakan beberapa model dasar untuk menghasilkan satu model prediksi yang optimal. Poin utamanya adalah menemukan satu model yang paling baik memprediksi output. Ide utama di balik penggunaan konsensus beberapa algoritma *machine learning* adalah untuk mengatasi keterbatasan beberapa algoritma, sebagai kekurangan dari satu algoritma mungkin menjadi kekuatan pada algoritma lain. Misalnya, Naive Bayes mengasumsikan bahwa kehadiran satu fitur tidak terkait dengan fitur lain [20]. Selain itu keuntungan dari penggunaan metode Hybrid adalah mendapatkan performansi akurasi tinggi, kompleksitas komputasi yang baik, dan lebih fleksibel serta *robust* dalam mengolah data dengan dimensi yang tinggi. Walaupun begitu, terdapat juga kelemahan dari metode ini adalah hasil yang didapatkan bergantung pada *feature selection* yang digunakan serta *classifier* yang dipilih [22].



**Gambar 2. Diagram Proses Ensemble Voting**

Seperti yang dapat dilihat pada gambar 2 mengenai diagram proses Ensemble Voting. Metode pengklasifikasian yang digunakan akan menghasilkan output yang akan digabungkan aturan rata-rata dari *probability voting*. Selanjutnya, kelas yang direpresentasikan memiliki rata-rata nilai yang tinggi akan dipilih untuk hasil akhir [8].



Kemudian pada gambar 3 merupakan ilustrasi bagaimana ensemble voting bekerja. Pada ilustrasi ini, misal diberikan input  $x$ , maka peluang  $x$  menurut model pertama yaitu SVM yang akan masuk ke kelas 0 adalah 0,9 dan masuk kelas 1 adalah 0,8. Selanjutnya peluang menurut model kedua yaitu Naïve Bayes yang akan masuk ke kelas 0 adalah 0,85 dan masuk kelas 1 adalah 0,65. Terakhir pada model ketiga yaitu Decision Tree, peluang  $x$  yang masuk pada kelas 0 adalah 0,4 dan pada kelas 1 adalah 0,8. Selanjutnya model ensemble melakukan perhitungan rata-rata pada peluang  $x$  yang masuk ke kelas 0 = 0,716 yang merupakan nilai rata-rata dari peluang  $x$  yang dimiliki oleh ketiga model yang masuk pada kelas 0. Lalu rata-rata peluang  $x$  untuk kelas 1 = 0,75 yaitu nilai rata-rata dari peluang  $x$  yang dimiliki oleh ketiga model yang ada pada kelas 1. Jadi berdasarkan model ensemble voting yang digunakan,  $x$  akan diklasifikasikan dalam kelas 1 karena memiliki nilai rata-rata peluang yang lebih tinggi daripada nilai rata-rata peluang kelas 0.

### 3.6 Sistem Evaluasi

Sistem ini digunakan peneliti untuk mengukur tingkat ketepatan data dengan memanfaatkan *precision* (*PREC*) dan *recall* (*REC*) untuk menghitung skor yang didapat. Dimana *Confusion Matrix* digunakan untuk mendapatkan hasil data *actual* yang kemudian dilakukan perhitungan prediksi klasifikasi. Dapat dilihat pada tabel 2 dibawah.

**Tabel II. Contoh Confusion Matrix**

	Prediksi Positif	Prediksi Negatif
Aktual Positif	TP	FN
Aktual Negatif	FP	TN

*Precision* – Proporsi data yang diprediksi benar dengan persamaan:

$$\text{Precision} = \frac{TP}{FP+TP}$$

*Recall* – Proporsi relevan yang diidentifikasi benar dengan persamaan:

$$\text{Recall} = \frac{TP}{FN+TP}$$

*Accuracy* – Proporsi jumlah total data yang diprediksi benar dengan persamaan:

$$\text{Accuracy} = \frac{TP+TN}{TN+FN+FP+TP}$$

*F1 - Score* – Mengkombinasikan nilai *precision* (*PREC*) dan *recall* (*REC*) untuk mendapatkan nilai *mean* [17], dengan persamaan:

$$F1\text{-Score} = \frac{2PREC \times REC}{PREC + REC}$$

#### 4. Evaluasi

##### 4.1 Hasil Pengujian

Pada tabel 3 merupakan nilai dari model yang dibuat dari SVM, Naïve Bayes, Decision Tree, dan Hybrid. Model tersebut menggunakan Count Vectorizer untuk pembobotan.

**Tabel III. Hasil Percobaan Classification Report 80:20**

Classifier	Feature	Accuracy	Precision	Recall	F1-Score
SVM	Count Vectorizer	0.8192	0.885	0.765	0.78
Naïve Bayes		0.8182	0.885	0.765	0.78
Decision Tree		0.7977*	0.88*	0.74*	0.75*
Hybrid		0.8182	0.885	0.76	0.775

Pada hasil yang ada pada tabel 11 dapat diperhatikan bahwa “\*” memperlihatkan metode hybrid memiliki hasil yang lebih baik dibandingkan dengan Decision Tree dan memiliki performa sedikit lebih rendah daripada Naïve Bayes. Kemudian pada tabel 4 merupakan hasil performansi untuk data yang dilakukan *split* 70:30. Dapat dilihat bahwa tidak terlalu ada perbedaan yang signifikan, namun metode Hybrid lebih unggul daripada Decision Tree.

**Tabel IV. Hasil Percobaan Classification Report 70:30**

Classifier	Feature	Accuracy	Precision	Recall	F1-Score
SVM	Count Vectorizer	0.8202	0.89	0.76	0.78
Naïve Bayes		0.8189	0.89	0.76	0.78
Decision Tree		0.7977*	0.88	0.74*	0.76*
Hybrid		0.8189	0.88	0.76	0.78

Kemudian dilakukan eksplorasi lebih pada klasifikasi karena peneliti ingin mengetahui hasil yang didapatkan oleh metode klasifikasi hybrid jika data dilakukan *oversampling*, *undersampling*, dan juga split data dengan perbandingan yang berbeda. *Original* merupakan kondisi data yang awal yang belum dilakukan *resampling*. Dimana dapat dilihat pada tabel 4 merupakan hasil eksplorasi peneliti menggunakan beberapa kondisi data dan split data 80:20.

**Tabel V. Hasil Percobaan Classification Report Hybrid**

Data Condition	Accuracy	Precision	Recall	F1
Original	0,8182	0,88	0,76	0,77
Oversampling	0,7723	0,84	0,77	0,75
Undersampling	0,7737	0,84	0,77	0,76

Eksplorasi terakhir adalah hasil *report* dengan menggunakan split data 70:30 yang akan ditunjukkan pada tabel 5 dibawah ini.

**Tabel VI. Hasil Percobaan Classification Report Hybrid**

Data Condition	Accuracy	Precision	Recall	F1
Original	0,8189	0,88	0,76	0,78
Oversampling	0,7677	0,84	0,76	0,75
Undersampling	0,7703	0,84	0,76	0,75

Dari kedua tabel diatas menunjukkan bahwa pemberlakuan *resampling* pada data menyebabkan hasil khususnya pada akurasi dan F1 menurun. Hal ini bisa terjadi karena banyaknya data *dummy* hasil *resample* pada kondisi data Oversampling dan juga pada kondisi data Undersampling, dimana hasil keduanya dipengaruhi oleh perbandingan data yang digunakan.

##### 4.2 Analisis Hasil Pengujian

Pada tabel hasil percobaan 4 dan 5 menggunakan fitur Count Vectorizer, pengondisian data, dan split data



juga dapat dibandingkan jika akurasi dan secara evaluasi bahwa metode Hybrid memiliki tingkat akurasi yang lebih tinggi pada *split data* 70:30. Hal ini terjadi karena metode hybrid merupakan yang merupakan

## 5. Kesimpulan

Dalam penelitian ini, peneliti telah mendapatkan hasil eksperimen klasifikasi opini pengguna Twitter tentang *review* film menggunakan metode Hybrid. Dimana metode ini mampu memiliki performansi yang lebih baik dibandingkan sala satu *classifier* yang bekerja secara individu. Selain membandingkan dengan metode lain, pada penelitian ini juga membandingkan performansi metode Hybrid dengan *split data* yang berbeda yang mendapatkan hasil akhir evaluasi metode menggunakan fitur *Count Vectorization* tersebut didapatkan pada metode Hybrid hasil penggabungan ketiga metode dengan *ensemble voting* mendapatkan skor presisi 0.88, recall 0.76, F1-Score 0.78, dan akurasi 0.8189. Hal ini menunjukkan bahwa kombinasi metode Hybrid yang terdiri dari SVM, Naïve Bayes, dan Decision Tree memiliki memiliki pengolahan dalam pengklasifikasian baik pada skenario kedua yaitu untuk *split data training* dan *test* 70:30 dibanding 80:20. Pada hasil penelitian ini menunjukkan bahwa metode Hybrid mampu melakukan klasifikasi sentimen analisis pada *review* film ditunjukkan dari hasil evaluasi yang baik, ditunjukkan juga pada peningkatan akurasi dan F1-score yang didapatkan. Meskipun demikian, perbedaan strategi atau skenario dapat mempengaruhi hasil yang didapatkan memanfaatkan metode Hybrid.

## REFERENSI

- [1] Kapardhi K. G. 2018. Sentiment Analysis of Tweets to Classify The Box Office Success of Movies. National College of Ireland.
- [2] W. R. Bristi, Z. Zaman and N. Sultana. 2019. Predicting IMDb Rating of Movies by Machine Learning Techniques. 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-5.
- [3] S. A. El Rahman, F. A. AlOtaibi, and W. A. Alshehri. 2019. Sentiment Analysis of Twitter Data. International Conference on Computer and Information Sciences (ICCIS), Sakaka, Saudi Arabia, pp. 1-4.
- [4] F. H. Khan, S. Bashir and U. Qamar. TOM: Twitter opinion mining framework using hybrid classification scheme. Computer Engineering Department, College of Electrical and Mechanical Engineering, National University of Sciences and Technology (NUST), Islamabad, Pakistan.
- [6] M. Wongkar and A. Angdresey. 2019. Sentiment Analysis Using Naïve Bayes Algorithm of The Data Crawler: Twitter. Fourth International Conference on Informatics and Computing (ICIC), Semarang, Indonesia, pp. 1-5.
- [7] Shawni Dutta, Samir Bandyopadhyay. 2020. Forecasting of Campus Placement for Student Using Ensemble Voting Classifier. Asian Journal of Research in Computer Science (AJRCOS), 5(4), pp. 1 – 12.
- [8] Isha Gandhi, Mrinal Pandey. 2015. Hybrid Ensemble of Classifier using Voting. International Conference on Green Computing and Internet of Things (ICGCIoT).
- [9] Z. Zhuo. 2018. Sentiment Analysis of Steam Review Datasets using Naïve Bayes and Decision Tree Classifier.
- [10] T. Verma and D. Gaur. 2014. Tokenization and Filtering Process in RapidMiner. Int. J. Appl. Inf. Syst., vol. 7, no. 2, pp. 16–18.
- [11] A. Hotho, A. Nürnberger, and G. Paaß. 2005. A Brief Survey of Text Mining. LDV Forum - Gld. J. Comput. Linguist. Lang. Technol., vol. 20, pp. 19–62.
- [12] S. Vijayarani, M. J. Ilamathi, and M. Nithya. 2015. Preprocessing Techniques for Text Mining - An Overview. Int. J. Comput. Sci. Commun. Networks, vol. 5, no. 1, pp. 7–16.
- [13] A. Kanitkar. 2018. Bollywood Movie Success Prediction using Machine Learning Algorithms. 3<sup>rd</sup> International Conferences on Circuits, Control, Communication and Computing (I4C), Bangalore, India, pp. 1-4
- [14] Avinash, M., Sivasankar, E. 2019. A study of feature extraction techniques for sentiment analysis. Emerging Technologies in Data Mining and Information Security, vol. 814, pp. 475–486. Springer, Singapore.
- [15] S. Sarlis, I. Maglogiannis. 2020. On The Reusability of Sentiment Analysis Datasets Applications with Dissimilar Contexts. IFIP Advances in Information and Communication Technology, vol 583.
- [16] Q. I. Mahmud, A. Mohaimen, M. S. Islam, and Marium-E-Jannat. 2017. A Support Vector Machine Mixed with Statistical Reasoning Approach to Predict Movie Success by Analyzing Public Sentiments. 20<sup>th</sup> International Conference of Computer and Information Technology (ICCIIT), Dhaka, pp. 1-6.
- [17] Charu Nanda, Mohit Dua, Garima Nanda. 2018. Sentiment Analysis of Movie Reviews in Hindi Language using Machine Learning. International Conference on Communication and Signal Processing, India.
- [18] D. Lavanya and D. Rani. 2012. Ensemble Decision Tree Classifier for Breast Cancer Data. International Journal of Information Technology Convergence and Services (IJITCS), vol. 2, pp. 17-24.
- [19] D. Pawel, S. Zbigniew. 2013. Hybrid Methods in Data Classification and Reduction. Institute of Computer Science, University of Rzeszow, Rzeszow, Poland.
- [20] H. A. Nayyar and A. El-Hag. 2020. Two-Layer Ensemble-Based Soft Voting Classifier for Transformer Oil Interfacial Tension Prediction. Department of Electrical and Computing Engineering, University of Waterloo, Waterloo, Canada.
- [21] S. V. Wawre, S. N. Deshmukh. 2015. Sentiment Classification using Machine Learning Techniques. International Journal of Science and Research (IJSR).
- [22] Rabia Aziz Musheer, C. K. Verma, Namita Srivastava. 2017. Dimension reduction methods for microarray data: a review. Department of Mathematics & Computer Application, Maulana Azad National Institute of Technology Bhopal-462003 (M.P.), India.
- [23] Chintan Dedhia, Jyoti Ramteke. 2017. Ensemble model for Twitter Sentiment Analysis. International Conference on Inventive Systems and Control (ICISC-2017).
- [24] Rashid Khan, Furqan Rustam, Khadija Kanwal, Arif Mehmood, Hyu Sang Choi. 2021. US Based Covid-19 Tweets Sentiment Analysis Using Textblob and Supervised Machine Learning Algorithms. International Intelligence (ICAI), Islamabad, Pakistan.
- [25] Reza Hermansyah, Riyanarto Sarno. 2020. Sentiment Analysis about Product and Service Evaluation of PT Telekomunikasi Indonesia Tbk from Tweets Using Textblobs, Naïve Bayes and K-NN Method. International Seminar on Application for Technology of Information and Communication (iSemantic).

