

# Studi Analisis Metode Analisis Sentimen pada YouTube

Santi Thomas<sup>1</sup>, Yuliana<sup>2</sup>, Noviyanti. P<sup>3</sup>

<sup>1,2,3</sup> Teknologi Informasi, Institut Shanti Bhuana

<sup>1</sup>santi.thomas@shantibhuana.ac.id, <sup>2</sup>yuliana@shantibhuana.ac.id, <sup>3</sup>noviyanti@shantibhuana.ac.id

**Abstract**—Abstract—It cannot be denied that YouTube is now the most popular video sharing website. Opinions written from viewers could be an important asset for the development of a company or YouTubers. In this report, the technique used to analyze those comments is called sentiment analysis or opinion mining techniques. Although this knowledge or technique is not new, it is still important to continue studying this, given the importance of the viewer comments. This paper is the result of a literature review of several studies that have been conducted using various methods in sentiment analysis. The purpose of this analytical study is to obtain a theoretical basis that can support further research by studying the working methods, advantages, and disadvantages of each method. For the reason of having no knowledge on how the method works will affect the results and will be a waste of time. The results of comparisons from research that have been done, showing that the Naïve bayes algorithm has a higher accuracy, then SVM then DT. But this is a preliminary result as no study has used all methods at once in a case.

**Keywords**— Sentiment Analysis, method, YouTube, Analytical Study

**Intisari**— Tidak dapat dipungkiri bahwa YouTube sekarang merupakan situs web berbagi video yang paling banyak diminati. Opini yang ditulis dari para viewer bisa menjadi aset yang penting untuk perkembangan suatu perusahaan ataupun para YouTuber. Teknik yang digunakan untuk menganalisa komentar ini disebut dengan analisis sentimen atau teknik *opinion mining* dan meskipun ilmu ini bukan merupakan suatu hal yang baru tapi tetap penting untuk terus dipelajari mengingat pentingnya komentar viewer tersebut. Paper ini merupakan hasil dari *literatur review* atas sejumlah penelitian yang telah dilakukan dengan menggunakan berbagai metode dalam analisis sentiment . Tujuan dari studi analisis ini adalah untuk mendapatkan landasan teori yang bisa mendukung untuk penelitian selanjutnya dengan mempelajari cara kerja, kelebihan dan kekurangan dari setiap metode karena ketidak tahuan akan cara kerja metode akan mempengaruhi hasil dan menghabiskan waktu. Hasil perbandingan dari penelitian yang sudah dilakukan, didapatkan hasil bahwa algoritma NB memperoleh tingkat akurasi yang paling tinggi, lalu diikuti oleh SVM kemudian DT. Tapi ini hanya hasil sementara sampai dilakukan penelitian atas 1 video dengan menggunakan semua metode

**Kata Kunci**— Analisis Sentimen, metode, YouTube, Studi Analisis

## I. PENDAHULUAN

Sejak tahun 2005, dimana YouTube pertama kali dikenalkan hingga sampai saat ini, mengalami peningkatan popularitas yang cukup tajam. Menurut data dari *Hootsuite*, sebuah situs layanan manajemen konten yang menyediakan layanan media daring yang terhubung dengan berbagai situs jejaring sosial seperti *Youtube*, *Facebook*, *Instagram*, *Twitter*, *Line*, *Pinterest*, *Whatsapp*, *Fb Messenger*, *Wechat* (*Weixin*), *QQ*, *Qzone*, *Tiktok* (*Douyin*), *Sina Weibo*, *Reddit*, *Douban*, *Linkedin*, *Baidu Tieba*, *Skype*, *Snapchat*, dan *Vibe* , mengatakan

jumlah pengguna YouTube pada tahun 2020 di seluruh dunia mencapai 2 milyar pengguna dimana setiap hari orang menonton 1 milyar jam video dan 500 video diupload setiap menitnya. Di Indonesia sendiri, YouTube menduduki peringkat 1 dari seluruh pengguna media sosial dengan persentase sebesar 88% [1].

Youtube merupakan salah satu dari social Network yang berkategori *Multimedia sharing* [2]. Pada umumnya, file-file Multimedia yang dibagikan berupa file gambar, audio, dan video melalui internet. Fitur-fitur yang disediakan multimediasharing antara lain mengirim dan menanggapi komentar, pemberian *rate* (peringkat), berbagi (*share*) dalam bentuk alamat (*URL*). Proses yang dilakukan meliputi unggah (*upload*) dan unduh (*download*). Diantara social network yang ada saat ini seperti *Flicker* dan *Picasa*, YouTube tetap menjadi *platform sharing* video terbesar di Internet.

Sudah banyak penelitian dilakukan tentang pola interaksi pengguna di YouTube ini. Interaksi ini berupa komentar yang biasanya digunakan untuk menunjukkan perasaannya, memberikan dukungan emosi, kesan, atau nasehat sebagai tanggapan atas video[3] . Perilaku pengguna banyak digunakan untuk berbagai kepentingan seperti tujuan pemasaran dimana perusahaan dapat menggunakan informasi untuk menargetkan pesan dan produk mereka [4], dalam politik untuk mendukung strategi kampanye, untuk mengevaluasi pelayanan dalam pemerintahan [5] dsbnya. Karena itu, YouTube tidak hanya berfungsi sebagai hiburan saja, tetapi juga untuk memasarkan *brand* dari perusahaan, pendidikan, olahraga, kesehatan, dokumentari , dll [6]. Teknik yang dipakai untuk menganalisa komentar adalah analisis sentimen (*sentiment analysis*) yaitu suatu proses menganalisa dimana jutaan orang membagikan pengalaman mereka akan berbagai masalah dalam bentuk pendapat, saran, komentar, masukan, dll . Ada beberapa metode yang dapat digunakan dalam analisis sentiment. Penelitian ini bertujuan untuk mendalami cara kerja dari setiap metode yang ada, sehingga untuk penelitian selanjutnya dapat menggunakan metode yang tepat dengan hasil yang akurat tanpa membuang banyak waktu.

## II. TINJAUAN PUSTAKA

Ada beberapa pendapat para ahli tentang definisi analisis sentimen. Dalam penelitian dengan judul *Arabic sentiment analysis: Lexicon-based and corpus-based*. mengatakan analisis sentimen adalah jenis *natural language* yang digunakan untuk melacak *mood* publik tentang sebuah produk atau topik tertentu [7]. Analisis sentimen kadang disebut juga sebagai *opinion mining*, dimana melibatkan sistem untuk mengumpulkan dan memeriksa pendapat tentang produk yang diposting di blog, komentar, review atau tweet. Menurut Bing Liu [8] analisis sentiment merupakan suatu bidang ilmu pengetahuan yang menganalisa pendapat, perasaan, evaluasi, penilaian, sikap dan emosi terhadap kualitas seperti produk, layanan, organisasi,

individu, masalah, peristiwa, topik dan atributnya. Analisis sentimen umumnya merupakan permasalahan *Supervised Learning*.

Secara sederhana cara kerja analisis sentimen dapat dibagi menjadi 3 langkah yaitu : klasifikasi , evaluasi dan visualisasi data [9].

#### A. Klasifikasi

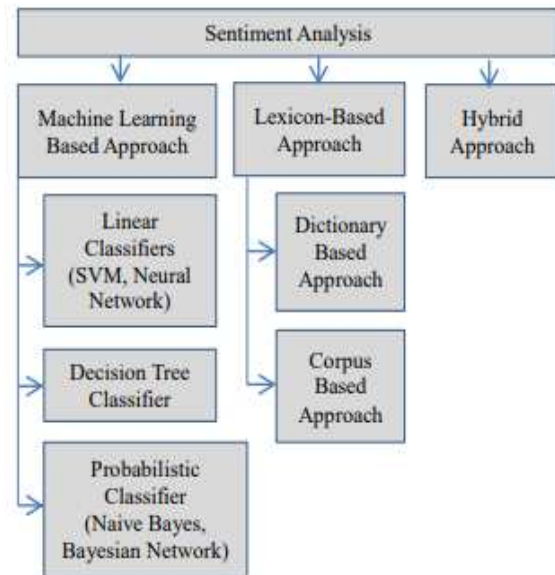
Kemampuan mengklasifikasikan komentar YouTube secara akurat membantu untuk memahami perilaku media sosial yang muncul dan potensi dampaknya. Penelitian oleh Sadia,dkk dengan judul *An Overview of Lexicon-Based Approach For Sentiment Analysis* membedakan antara jenis komentar dan memilih jenis komentar untuk analisis atau menganalisa perbedaan antara jenis komentar pada video yang berbeda [3]. Perbedaan kategori video juga mempengaruhi jenis komentar yang dilontarkan. Banyak video, terutama di kategori "Olahraga" dan "Berita dan Politik" menarik banyak kata-kata kotor dan konten yang menghina, sementara konten lainnya mengarah ke provokasi rasial. Peneliti ini membagi komentar menjadi 10 kategori yaitu : informasi (*information*) , nasehat (*advice*) , kesan (*impression*), pendapat (*opinion*), tanggapan atas komen sebelumnya (*responses to previous comments*), ekspresi perasaan pribadi (*Expression of personal feelings*), percakapan umum (*General conversation*), proses situs (*site processes*), deskripsi konten video (*video content description*) dan kategori *non-response* (*non-response categories*). Sementara Siesdorfer dalam penelitiannya membagi komentar menjadi 4 kategori yaitu : *seft-promotion* yaitu komen tentang konten video, *propaganda* yaitu komen tentang agama dan sosial, *abusive comments* yaitu komen yang sangat rasial dan *miscellaneous comments* yaitu komen yang tidak termasuk dalam kategori apapun karena bersifat netral [10].

Manfaat skema klasifikasi adalah memungkinkannya peneliti dan praktisi untuk mengidentifikasi jenis dari komentar dan memisahkannya untuk dianalisa. Analisis ini dapat bersifat kualitatif untuk menyelidiki bahasa pengguna ataupun kuantitatif untuk analisis sentimen.

Metode Analisis Sentimen dapat dibagi menjadi 3 yaitu :

1. *Machine learning*: fitur-fitur di dalamnya dapat mengenali sentimen (sudut pandang seseorang) dalam sebuah teks. Metode *machine learning* kini semakin bertambah populer karena dapat dinilai representatif. Yang termasuk didalam *Machine Learning* adalah metode SVM, Neural Network, Decision Tree dan Naïve Bayer.
2. *Lexicon-based* yaitu menggunakan berbagai kata yang dinilai dengan skor polaritas untuk mengetahui tanggapan masyarakat/pengguna/konsumen mengenai suatu topik. Pendekatan dengan kamus dan corpus termasuk didalam metode *Lexicon-based* ini.
3. Campuran (Hybrid) yaitu menggabungkan 2 metode yang ada.

Skema ketiga metode ini dapat dilihat dalam gambar 1 dan dibahas lebih lanjut dalam bab 3.



Gambar 1. Metode Analisis Sentimen untuk Klasifikasi Data

#### B. Evaluasi

Skema evaluasi adalah sebagai berikut: tinjauan dataset, *pre-processing* yang meliputi *case folding* (mengubah semua huruf menjadi huruf kecil semua), *tokenizer* (membagi kalimat menjadi kata), penghapusan *stopwords* (menghapus kata-kata yang sering muncul tetapi tidak memiliki makna) dan *lemmatization* (mengubah kata menjadi kata dasar), lalu klasifikasi dan evaluasi.

Untuk mengukur kinerja suatu model pada *machine learning* digunakan *confusion matrix* yang memberikan perbandingan hasil antara model (sistem) dengan hasil klasifikasi. Confusion Matrix dapat digambarkan seperti pada gambar 2:

		Nilai sebenarnya	
		TRUE	FALSE
Nilai prediksi	TRUE	TP (True Positive) <i>Correct result</i>	FP (False Positive) <i>Unexpected result</i>
	FALSE	FN (False Negative) <i>Missing result</i>	TN (True Negative) <i>Correct absence of result</i>

Gambar 2. Confusion Matrix

TP merupakan data positif yang diprediksi benar.

TN merupakan data negatif yang diprediksi benar.

FP merupakan data negatif tapi diprediksi sebagai data positif.

FN merupakan data positif tapi diprediksi sebagai data negatif

Tahap selanjutnya dalam evaluasi adalah menggunakan matrik evaluasi seperti *Precision*, *Recall*, *F-score*, dan *Accuracy*. *Precision* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. *Recall* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif.

Accuracy merupakan rasio prediksi Benar (positif dan negatif) dengan keseluruhan data . F1 Score merupakan perbandingan rata-rata precision dan recall yang dibobotkan [11].

Rumus yang digunakan adalah sebagai berikut :

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$\text{F-score} = \frac{2*(\text{Recall}*Precision)}{\text{Recall}+Precision} \quad (4)$$

C. Visualisasi Data

Langkah terakhir dalam analisis sentimen adalah visualisasi data. Visualisasi data dilakukan menggunakan bagan sesuai kebutuhan perusahaan atau siapa saja yang memanfaatkan data-data ini. Biasanya orang menggunakan teknik yang sudah dikenal, seperti grafik, histogram, atau matriks. Namun dapat juga berupa wordcloud atau peta interaktif karena hasil akhir dari sentiment analysis bisa sangat bervariasi.

III. PEMBAHASAN

Pada bagaian ini akan dibahas tentang metode-metode yang sudah digunakan dalam analisis sentiment pada YouTube.

A. Lexicon-Based Approach

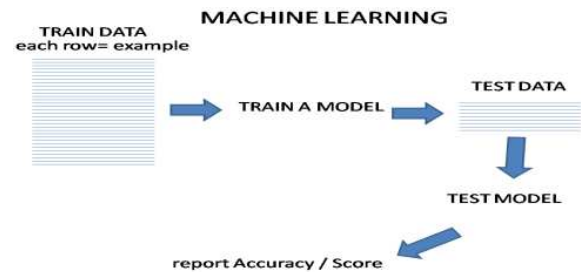
Lexicon-Based pada umumnya menggunakan kamus (Dictionary Based Approach) seperti SentiWordNet (SWN) untuk menggambarkan popularitas (positif, negatif atau netral). SWN dikembangkan untuk bahasa Inggris tetapi sekarang sudah mulai dikembangkan untuk bahasa lain termasuk bahasa Indonesia [6]. SWN merupakan hasil anotasi otomatis dari semua synset pada WordNet. Konsep utama pada WordNet adalah synset, di mana synset merupakan kumpulan dari 1 atau lebih kata yang memiliki makna sama. Selain itu penggunaan metode negation detection juga penting karena kata ‘tidak’ membuat sebuah kata yang tadinya positif menjadi negatif, misalnya bagus menjadi tidak bagus. Social Media Aware Phrase Detection (SMAPD) digunakan untuk mendeteksi kata yang sering dipakai dalam media masa tetapi belum ada di kamus SWN. Metode ini masih dapat dikembangkan dikemudian hari.

Metode ini cenderung lebih mudah diaplikasi daripada Machine learning Based Approach tapi kekurangannya adalah membutuhkan keterlibatan manusia dalam proses analisis [3], kamus yang digunakan masih memiliki keterbatasan [12] dan masih belum bisa menganalisa gaya bahasa sarkasme [13]. Karena itu SWN dan lexicon lainnya cocok digunakan untuk berita, blogs, produk review dan movie review dimana orang hanya menggunakan gaya bahasa tertentu. Berbeda dengan media sosial lainnya (seperti YouTube) dimana bahasa yang digunakan biasanya informal dan bahasa slangs atau gaul .

Corpus-Based diterapkan oleh 2 peneliti untuk komentar berbahasa Arab [7][14]. Kerangka umum untuk Corpus-Based ada dua bagian; yang pertama mewakili pembentukan corpus, model serta model pengujian yang kedua mengumpulkan data teks dari sumber, menerapkan evaluasi dan mengekstrak fitur-fitur seperti yang disebutkan dalam bagian pertama. Karena Dataset dan leksikon Arab yang tersedia untuk umum jarang ditemukan dan terbatas, perlu pembuatan set data dan untuk klasifikasi menggunakan metode SVM memberikan akurasi tertinggi dibandingkan metode lainnya. Dengan meningkatnya size leksikon, keakuratan juga meningkat.

B. Machine Learning Based Approach (ML)

Merupakan inferensi terhadap data dengan pendekatan sistematis. Inti dari ML ini adalah untuk membuat model sistematis yang mereflesikan pola-pola data. Biasanya menggunakan statika atau aljabar linear [15]. Merupakan bagian dari Artificial Intelligence (AI) yang berfokus pada pembuatan mesin yang dapat belajar tanpa diprogram secara eksplisit.



Gambar 3. Algoritma analisis sentiment dengan Machine Learning Based Approach

1. Naïve Bayer (NB)

NB adalah klasifikasi probabilitas linear yang mempunyai asumsi independent yang kuat [16]. NB mengadopsi pendekatan Bag-of-Words (BOW) Persamaan dari probabilitas prior:

$$P(H) = \frac{N_j}{N} \quad (5)$$

Nj : Jumlah data pada suatu class

N : Jumlah total data

Persamaan dari teorema bayes:

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (6)$$

X : Data class belum diketahui

H : Hipotesis data class spesifik

P(H|X) : Probabilitas Hipotesis H terhadap kondisi

X P(H) : Probabilitas Hipotesis H

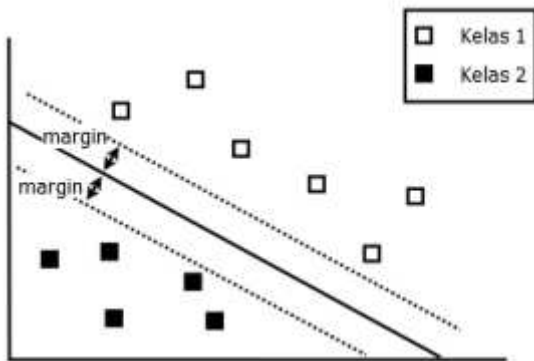
P(X|H) : Probabilitas X terhadap kondisi hipotesis H

P(X) : Probabilitas X

Keuntungan dari metode ini adalah sederhana, cepat dan berakurasi tinggi. Selain itu hanya membutuhkan sedikit data pelatihan untuk memperkirakan parameter (sarana dan varians variabel) yang diperlukan untuk klasifikasi [17].

## 2. Support Vector Machine (SVM)

*Support Vector Machine (SVM)* merupakan salah satu metode dalam *supervised learning* yang biasanya digunakan untuk klasifikasi (seperti *Support Vector Classification*) dan regresi (*Support Vector Regression*)[18]. SVM digunakan untuk mencari *hyperplane* terbaik dengan memaksimalkan jarak antar kelas. *Hyperplane* adalah sebuah fungsi yang dapat digunakan untuk pemisah antar kelas. Dalam 2-D fungsi yang digunakan untuk klasifikasi antar kelas disebut sebagai *line whereas*, fungsi yang digunakan untuk klasifikasi antar kelas dalam 3-D disebut *plane similarly*, sedangkan fungsi yang digunakan untuk klasifikasi di dalam ruang kelas dimensi yang lebih tinggi disebut *hyperplane*.



Gambar 4. *Hyperplane* yang memisahkan 2 kelas [28]

Data yang digunakan dengan  $x_i \in \mathbb{R}^d$  dan label dinotasikan dengan  $Y_i \in \{1,2\}$  untuk  $i = 1,2,\dots,1$  yang mana 1 adalah banyaknya data. Persamaan SVM adalah sebagai berikut:

$$\text{Jika } Y_i = +1, \text{ maka persamaan } x_i * w + b > +1 \quad (7)$$

$$\text{Jika } Y_i = -1, \text{ maka persamaan } x_i * w + b < -1 \quad (8)$$

Sebuah penelitian dilakukan dengan komentar bahasa Indonesia atas video produk *smartphone* [19] menggunakan FVEC-SVM. Karena komentar bahasa Indonesia mengandung banyak sekali jumlah *slang* dan strukturnya tidak dapat digeneralisasikan, maka metode STRUCT berkinerja  $\pm 5\%$  lebih sedikit dari FVEC, dan FVEC dilakukan tanpa penghitungan leksikon karena kekurangan *support library*. Selain itu, pemilihan kernel yang sesuai fungsi di SVM perlu dieksplorasi untuk mengungguli metode opini mining. Dari hasil percobaan menunjukkan bahwa FVEC-SVM menggunakan fungsi kernel linier untuk opini mining memiliki kinerja terbaik dibandingkan dengan fungsi kernel polinomial derajat 2, derajat polinomial 3, dan RBF. FVEC-SVM secara umum menggunakan fungsi kernel linier yang dapat mengklasifikasikan bahasa Indonesia komentar video YouTube tentang ulasan *smartphone* dengan akurasi 62,76%.

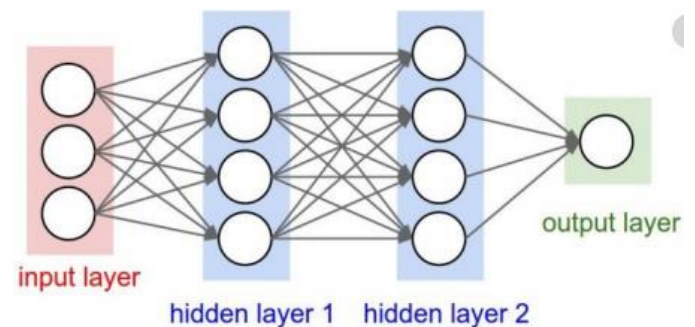
Penelitian serupa [20] dengan pendekatan struktural (STRUCT) dan SVM atas video trailer film Indonesia. Untuk menentukan ranting video. Dalam melakukan analisis sentimen untuk menghasilkan rating sentimen memiliki permasalahan dimana tidak semua komentar relevan kepada topik yang bersangkutan sehingga tidak semua komentar dapat dijadikan parameter dalam menentukan rating sentimen. Karena itu ditambahkan parameter *like* dimana jumlah *like* yang ada pada

masing-masing komentar juga ikut dalam menentukan hasil dari rating sentiment.

Hasil klasifikasi STRUCT-SVM menggunakan linear kernel adalah *precision* 66%, *recall* 50%, *F-measure* 53%, dan akurasi sebesar 70%. Sedangkan dengan *RBF kernel* adalah *precision* 67%, *recall* 53%, dan *F-measure* 56% dan akurasi sebesar 71%. Hal tersebut menunjukkan bahwa STRUCT-SVM dengan RBF kernel lebih baik daripada linear kernel dalam melakukan klasifikasi dan menentukan rating sentiment.

## 3. Neural Network (NN)

Cara kerja NN meniru cara kerja otak manusia yaitu dengan syaraf-syaraf yang mengirimkan informasi. NN merupakan inti dari *Deep Learning*. *Deep Neural Network* merupakan *Neural Network* dengan banyak layer [21].



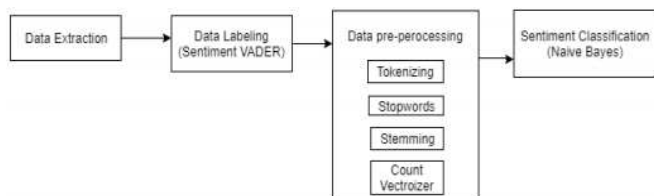
Gambar 5. Deep Neural network

Penelitian [22] dengan komentar berbahasa Portugis untuk menganalisa channel “*Joice Hasselmann*” yang mempunyai jumlah komentar terbanyak. Model *Deep Neural Network* ini menggunakan 6 layer sebagai berikut : *Embedding layer* untuk menghilangkan informasi yang tidak penting, *convolucional layer* ukuran 64 dan 16 untuk mengekstrak fitur dengan level yang berbeda, *dense layer* untuk mempelajari hubungan antara space fitur dengan kategori output, *dropout layer* untuk menghindari *overfitting* dan memastikan hasil yang didapat diantara range  $[-1; +1]$ .

Sebagai perbandingan, dilakukan juga klasifikasi menggunakan pengklasifikasi statistik Bayesian dengan distribusi Gaussian. Hasil yang diperoleh adalah *Deep Neural Network* dengan 6 layer ini mendapat akurasi yang lebih baik daripada model statistic dengan akurasi 60%-84% dan metode ini mempunyai hasil yang bagus untuk memprediksi komentar sentiment YouTube.

## C. Hybrid Approach

Merupakan gabungan 2 metode untuk mendapatkan klasifikasi sentiment yang lebih. Salah satu contoh Hybrid Approach yaitu menggabungkan Sentimen Vader dengan NB [23].



Gambar 6. Alur framework Hybrid Approach

Pendekatan Sentimen VADER diterapkan untuk memberi label pada komentar. Untuk melakukan klasifikasi biner, komentar netral telah dihapus sehingga menjadi klasifikasni binary dan pengklasifikasi NB dengan 70% data training dan 30% data testing dengan akurasi 79,78% dan skor F1 83,72%. Confussion Matrix yang menunjukkan bahwa pengklasifikasi telah bekerja dengan sangat baik dalam memprediksi sentimen komentar. Pendekatan leksikon Sentiment VADER yang digunakan untuk teks media sosial berdampak baik pada pengklasifikasi NB dalam memprediksi sentimen.

Contoh Hybrid approach yang lain seperti menggunakan Machine Learning untuk menganalisa sentimen dari komentar dan NB digunakan untuk menghitung polarity dari komentar. Sedangkan gabungan metode NB dan Lexicon based. NB digunakan untuk memprediksi komentar penerimaan. Studi tentang dataset yang besar dengan menggunakan SentiWordNet dan metadata YouTube menunjukkan ketergantungan yang kuat antara sentimen berbeda yang diungkapkan dalam komentar, peringkat komentar, dan orientasi topik dibahas dalam konten video. 2 contoh ini diungkapkan dalam penelitian yang dilakukan oleh Seindoufer dalam penelitiannya yang berjudul How Usefull Are Your Comment? [10]

#### IV. HASIL PERBANDINGAN PENGGUNAAN METODE DALAM MACHINE LEARNIG

Sudah banyak penelitian yang membandingkan metode yang berbeda untuk satu kasus sentiment analisis. Salah satunya adalah penelitian dengan judul Analysis and Implementation Machine Learning for YouTube Data Classification by Comparing the Performance of Classification Algorithms oleh Amanda dkk [24] dengan menganalisa video berkategori kesenian dan ilmu dengan menggunakan metode Random Forest, SVM dan Naive Bayes dengan hasil Naive Bayes memiliki akuransi terbaik. Hasil penelitian ini dapat dilihat dalam tabel 1.

Tabel 1. Hasil Penelitian dengan Random Foreset, SVM dan Naive Bayes

Model	Precision(%)	Recall (%)	Acuracy
Random Forest	87	82	0.8188
SVM	87	82	0.8228
Naive Bayes	<b>90</b>	<b>88</b>	<b>0.8779</b>

Sementara peneliti dari Thailand menentukan 2 video berbahasa Thailand berkategori musik dan iklan dengan SVM, MNB dan DT [25]. Hasil yang didapat untuk Video berkategori music NB yang paling tinggi lalu SVM kemudian DT, sedangkan video berkategori iklan, SVM yang paling tinggi, lalu NB kemudian DT. Hasil dapat dilihat dalam tabel 2.

Tabel 2. Hasil Penelitian untuk 2 video berbahasa Thailand dengan metode SVM, MNB dan DT

Genre	Classifiers		
	SVM (%)	NB (%)	DT (%)
MV	82.67	84.48	79.97
Advertisement	76.14	75.33	75.43

Penelitian menarik dari Indonesia dengan judul An Evaluation of Preprocessing Steps and Tree-based Ensemble Machine Learning for Analysing Sentiment on Indonesian YouTube Comments oleh Aribowo dkk [26] , menggunakan perbandingan antara NB, SVM, Decision Tree, Random Forest, dan Extra Tree classifier terhadap 2 dataset yang berhubungan dengan Covid 19. Hasil yang didapat bahwa Extra Tree classifier kemudian diikuti oleh Random Forest, SVM dan NB. Hasil penelitian ini dapat dilihat dalam tabel 3 .

Tabel 3. Hasil penelitian 2 dataset berbahasa Indonesia

Machine Learning	Dataset 1		Dataset 2	
	Akuransi	F Measure	Akuransi	F Measure
Naive Bayes	73.94	74.66	81.29	81.63
SVM	83.41	83.42	81.38	82.49
Decision Tree	83.58	83.58	79.29	79.28
Random Forest	86.76	86.76	83.35	83.36
Extra Tree	<b>88.19</b>	<b>88.2</b>	<b>84.45</b>	<b>84.84</b>

Dari penelitian [24][25] dapat dilihat bahwa NB memperoleh tingkat akuransi yang paling tinggi, lalu diikuti oleh SVM kemudian DT. Berbeda dengan apa yang didapat oleh Aribowo dkk [26], Extra Tree dan Random Forest memiliki nilai tertinggi. Tetapi untuk NB, SVM dan DT justru DT emppunyai nilai lebih tinggi untuk dataset 1 dan SVM memperoleh nilai lebih tinggi untuk dataset 2 sedangkan NB memperoleh nilai terendah dari semua algoritma yang dipakai.

#### V. MULTIMODAL SENTIMEN ANALISIS

Yang menjadi trend saat ini adalah penggunaan multimodal data untuk meningkatkan akuransi. Ini berarti selain bentuk text juga digunakan data yang lain sebagai informasi yaitu audio dan visual. Ekpresi wajah bisa menunjukkan emosi manusia. Emosi manusia pada dasarnya dibagi menjadi 6 sentimen sederhana yaitu marah, sedih, terkejut, takut, jijik dan bahagia serta emosi nonbasic seperti kelelahan, sakit, setuju, terlibat, keingintahuan, kesal dan

berpikir [27]. Sentimen emosi seperti menggerakkan alis, bibir, berkerut, dsbnya sedangkan sentimen audio dipengaruhi oleh gaya berbicara, penutur, kalimat dan bahasa. Variasi suara bisa menunjukkan berbagai emosi sehingga ucapan yang sama bisa mempunyai nilai sentimen berbeda.

Penelitian [28] terhadap video YouTube berbahasa Spanyol menggunakan *deep learning* karena model dari *deep learning* memetakan input ke beberapa ruang fitur, maka berbagai bentuk input dari multiple modal ini dapat juga diproyeksikan dalam model ini.

Ketentuan untuk video yang bisa dianalisa adalah orang harus ada di depan kamera dengan wajah yang kelihatan jelas dan tidak boleh ada musik atau animasi. *Software* yang digunakan adalah *Okao Vision*. Lalu para anatator menentukan video ini pernilai positif atau negative. Sementara audio menggunakan *OpenEAR15* secara otomatis menghitung intensitas nada dan suara.

Fusi multimodal dilakukan dengan menggabungkan bahasa, audio dan visual menjadi vektor fitur umum sehingga menghasilkan satu vector untuk setiap video. Untuk klasifikasi berupa klasifikasi binary digunakan SVM dengan kernel linear untuk menemukan *hyperplane* dengan max margin 16. Tools yang digunakan adalah WEKA. Untuk setiap percobaan dilakukan 10 kali validasi silang untuk seluruh dataset.

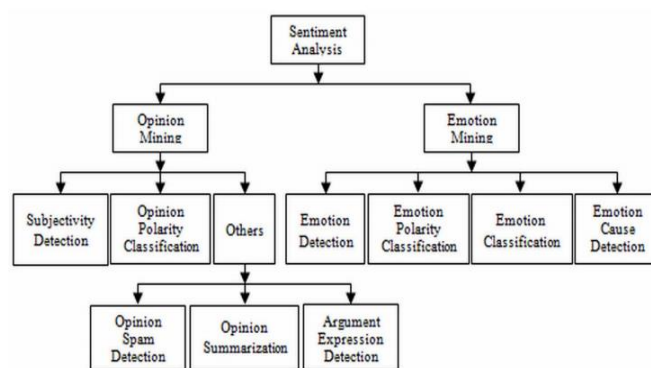
Hasil menunjukkan bahwa penggabungan text, audio dan video menunjukkan hasil akurasi yang lebih tinggi. Sedangkan kalau hanya text, audio atau video saja maka nilai akurasi tertinggi adalah text. Tabel 4 menunjukkan hasil dari penelitian.

Tabel 4. Hasil Penelitian Multimodal

Modality	Akuransi (%)
Text	64.94
Visual	61.04
Audio	46.75
Text-visual	73.68
Text-audio	68.42
Audio-visual	66.23
Text-audio-visual	75.00

Ketika diimplementasikan dengan dataset bahasa Inggris hasil yang didapat juga sama yaitu text dengan nilai tertinggi dan audio yang paling rendah.

Dengan alasan bahwa sentimen itu dipengaruhi oleh emosi maka kategori eksplorasi sentiment dibagi menjadi 2 yaitu *opini mining* yang didistribusikan dengan pendapat dan *emotion mining* yang didistribusikan dengan artikulasi emosi. Gambar 8 menunjukkan taksonomi eksplorasi sentimen.



Gambar 8. Taksonomi eksplorasi sentimen  
(Sumber: Yadollahi, Shahraki & Zaiane, 2017)

Meskipun polaritas opini sering kali berkorelasi dengan emosi tapi tetap sulit untuk menentukan popularitinya. Misalnya, seseorang dapat tersenyum yang sebenarnya mengungkapkan pendapat negatif, sehingga menjadikan analisis sentimen multimodal sebagai penelitian yang kompleks dan menantang.

Penggunaan multimodal ini mempunyai keuntungan untuk mengatasi ambiguitas, memberikan tambahan informasi sentimen dan meningkatkan koneksi ke lingkungan dunia nyata (*grounding*) [28] tetapi juga tingkat kesulitan yang tinggi [27] karena pengaruh budaya, perkembangan bahasa, konteks yang beragam dan perubahan dari single modal ke multimodal serta tools yang perlu dikembangkan lagi terutama untuk audio yang mempunyai akurasi yang rendah.

Analisis sentimen multimodal ini dapat digunakan untuk prediksi pasar, analisa bisnis dan deteksi fokus baru seperti deteksi emosi dari catatan kasus bunuh diri [27].

## VI. KESIMPULAN DAN SARAN

Paper ini berisi informasi mengenai perbandingan metode-metode yang bisa digunakan untuk melakukan proses analisis sentimen. Dari pembahasan diatas dapat dilihat bahwa yang masih menjadi tantangan saat ini untuk analisis sentiment adalah sulit menentukan sentimen terhadap kata ambigu, sarkasme, maupun slang. Meskipun sudah mempunyai sentiment lexicon (bahkan dalam orientasi khusus domain), tidak berarti kata yang ada dalam lexicon sesuai dengan sentiment dalam kalimat yang spesifik. Selain itu belum bisa diaplikasikan untuk video dengan komentar multibahasa karena kata *slang* tidak bisa diterjemahkan begitu saja. Serta pengembangan lebih lanjut untuk multimodal sentiment analisis. Untuk pengklasifikasian teks dengan jumlah data yang tidak terlalu banyak *Naïve Bayes* menjadi pilihan yang cocok sedangkan *Support Vector* cocok dalam mengklasifikasikan teks dengan jumlah dataset yang besar. *Lexicon Based* memang lebih mudah dan sederhana daripada *Machine Learning* tapi kurang lengkapnya *sentiment lexicon* bisa menjadi penghambat dalam menggunakan metode ini. Dari penelitian Amanda dan Sarakit [24][25] dengan menggunakan SVM, NB dan DT dapat dilihat bahwa NB memperoleh tingkat akurasi yang paling tinggi, lalu diikuti oleh SVM kemudian DT. Sebaliknya penelitian oleh Aribowo

dkk [26], Extra Tree dan Random Forest memiliki nilai tertinggi. Tetapi untuk NB, SVM dan DT justru DT mempunyai nilai lebih tinggi untuk dataset 1 dan SVM memperoleh nilai lebih tinggi untuk dataset 2 sedangkan NB memperoleh nilai terendah dari semua algoritma yang dipakai.

Untuk penelitian selanjutnya akan menganalisa beberapa video rohani yang akhir-akhir semakin semarak muncul sebagai akibat dari pandemi Virus Corona dengan menggunakan metode Naïve Bayes karena metode ini cocok digunakan dengan jumlah data yang sedikit dan tidak kompleks. Diharapkan dari Analisa video rohani ini dapat membantu peningkatan kualitas untuk video selanjutnya.

#### Referensi

- [1] We Are Social & Hootsuite, "Indonesia Digital report 2020," *Glob. Digit. Insights*, p. 247, 2020, [Online]. Available: <https://datareportal.com/reports/digital-2020-global-digital-overview>.
- [2] I. P. A. E. Pratama, *Komputer dan Masyarakat*. Bandung: Informatika Bandung, 2014.
- [3] A. Sadia, F. Khan, and F. Bashir, "An overview of lexicon-based approach for sentiment analysis," *Int. Electr. Eng. Conf.*, no. IEEC, pp. 1–6, 2018.
- [4] B. I. Permata, A. E. Prihatini, and Widiartanto, "Pengaruh Kualitas Produk Dan Periklanan Youtube Terhadap Loyalitas Pengguna Brand Wardah Kosmetik Di Kota Semarang," *J. Ilmu Adm. Bisnis*, vol. 7, no. 3, pp. 250–257, 2018.
- [5] N. Riyastika, "Analisis Penggunaan Youtube Pemerintah Provinsi DKI Jakarta Sebagai Sarana Komunikasi Ditinjau Dari Sudut Pandang Political PR," *Lib.Ui.Ac.Id*, no. Januari, p. 21, 2014, [Online]. Available: [www.lib.ui.ac.id](http://www.lib.ui.ac.id).
- [6] S. Christina and D. Ronaldo, "A Survey of Sentiment Analysis," *A Surv. ot Sentim. Anal. Using Sentirwordnet Bhs. Indones.*, vol. 12, no. 2, pp. 201–241, 2018, doi: 10.5715/jnlp.13.3\_201.
- [7] N. A. Abdulla, N. A. Ahmed, M. A. Shehab, and M. Al-Ayyoub, "Arabic sentiment analysis: Lexicon-based and corpus-based," *2013 IEEE Jordan Conf. Appl. Electr. Eng. Comput. Technol. AEECT 2013*, 2013, doi: 10.1109/AEECT.2013.6716448.
- [8] B. Liu, "Sentiment analysis: Mining opinions, sentiments, and emotions," *Sentim. Anal. Min. Opin. Sentim. Emot.*, no. May, pp. 1–367, 2015, doi: 10.1017/CBO9781139084789.
- [9] T. Annisa, "Mengenal peran sentiment analysis beserta cara kerjanya," 2020. <https://www.ekrut.com/media/sentiment-analysis-adalah> (accessed Dec. 17, 2020).
- [10] S. Siersdorfer, S. Chelaru, W. Nejdil, and J. San Pedro, "How useful are your comments?," p. 891, 2010, doi: 10.1145/1772690.1772781.
- [11] M. S. Hadna, P. I. Santosa, and W. W. Winarno, "Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2016, no. Sentika, pp. 57–64, 2016, [Online]. Available: <https://fti.uajy.ac.id/sentika/publikasi/makalah/2016/95.pdf>.
- [12] M. Z. Asghar, S. Ahmad, A. Marwat, and F. M. Kundi, "Sentiment Analysis on YouTube: A Brief Survey," pp. 1–11, 2015, [Online]. Available: <http://arxiv.org/abs/1511.09142>.
- [13] S. Choudhury and J. G. Breslin, "User Sentiment Detection: A YouTube Use Case," *21st Natl. Conf. Artif. Intell. Cogn. Sci. (AICS 2010)*, 2010.
- [14] A. Alnawas and N. Arıcı, "The Corpus Based Approach to Sentiment Analysis in Modern Standard Arabic and Arabic Dialects: A Literature Review," *J. Polytech.*, no. November, 2018, doi: 10.2339/politeknik.403975.
- [15] J. W. G. Putra, "Pengenalan Konsep Pembelajaran Mesin dan Deep Learning," no. August, pp. 1–235, 2019, [Online]. Available: <https://www.researchgate.net/publication/323700644>.
- [16] E. Poche, N. Jha, G. Williams, J. Staten, M. Vesper, and A. Mahmoud, "Analyzing User Comments on YouTube Coding Tutorial Videos," *IEEE Int. Conf. Progr. Compr.*, pp. 196–206, 2017, doi: 10.1109/ICPC.2017.26.
- [17] A. Krishna, J. Zambreno, and S. Krishnan, "Polarity trend analysis of public sentiment on YouTube," *19th Int. Conf. Manag. Data*, pp. 125–128, 2013.
- [18] Samsudiney, "Penjelasan Sederhana tentang Apa Itu SVM?itle," 2019. <https://medium.com/@samsudiney/penjelasan-sederhana-tentang-apa-itu-svm-149fec72bd02> (accessed Dec. 17, 2020).
- [19] E. Rinaldi and A. Musdholifah, "FVEC-SVM for opinion mining on Indonesian comments of youtube video," *Proc. 2017 Int. Conf. Data Softw. Eng. ICoDSE 2017*, vol. 2018-January, pp. 1–5, 2018, doi: 10.1109/ICODSE.2017.8285860.
- [20] K. A. B. Permana, M. Sudarma, and W. G. Ariastina, "Analisis Rating Sentimen pada Video di Media Sosial Youtube Menggunakan STRUCT-SVM," *Maj. Ilm. Teknol. Elektro*, vol. 18, no. 1, p. 113, 2019, doi: 10.24843/mite.2019.v18i01.p17.
- [21] B. Purnama, *Pengantar Machine Learning*. Bandung: Informatika Bandung, 2019.
- [22] A. A. L. Cunha, M. C. Costa, and M. A. C. P. Pacheco, "Sentiment Analysis of YouTube Video Comments Using Deep Neural Networks," vol. 2, pp. 358–372, 2019, doi: 10.1007/978-3-030-20912-4.
- [23] V. D. Chaithra, "Hybrid approach: Naive bayes and sentiment VADER for analyzing sentiment of mobile unboxing video comments," *Int. J. Electr. Comput. Eng.*, vol. 9, no. 5, pp. 4452–4459, 2019, doi: 10.11591/ijece.v9i5.pp4452-459.
- [24] R. Amanda, E. S. Negara, A. Info, D. Mining, E. Method, and Y. D. Classification, "Analysis and Implementation Machine Learning for YouTube Data Classification by Comparing the Performance of Classification Algorithms," vol. 5, no. 1, pp. 61–72, 2020, doi: 10.15575/join.v5i1.505.
- [25] P. Sarakit, T. Theeramunkong, C. Haruechaiyasak, and M. Okumura, "Classifying emotion in Thai youtube comments," *2015 6th Int. Conf. Inf. Commun. Technol. Embed. Syst. IC-ICTES 2015*, pp. 2–6, 2015, doi: 10.1109/ICTEmSys.2015.7110808.
- [26] A. S. Aribowo, H. Basiron, N. S. Herman, and S. Khomsah, "An evaluation of preprocessing steps and tree-based ensemble machine learning for analysing sentiment on Indonesian youtube comments," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 5, pp. 7078–7086, 2020, doi: 10.30534/ijatcse/2020/29952020.
- [27] R. Kaur and S. Kautish, "Multimodal sentiment analysis: A survey and comparison," *Int. J. Serv. Sci. Manag. Eng. Technol.*, vol. 10, no. 2, pp. 38–58, 2019, doi: 10.4018/IJSSMET.2019040103.
- [28] V. Perez Rosas, R. Mihalcea, and L. P. Morency, "Multimodal sentiment analysis of spanish online videos," *IEEE Intell. Syst.*, vol. 28, no. 3, pp. 38–45, 2013, doi: 10.1109/MIS.2013.9.