

# Aplikasi Web Sentimen Analisis Komentar Media Sosial

<sup>1</sup>Setyawan Widyarto, <sup>2</sup>Desti Destiansari Istinabiyah

<sup>1)</sup>Universiti Selangor, Malaysia

<sup>2)</sup>Program Pascasarjana Magister Ilmu Komputer

Universitas Budi Luhur, Indonesia

## ABSTRAK

**Abstrak** - Saat ini semakin banyak pelaku usaha, baik perusahaan besar maupun ritel, beralih atau mengembangkan usaha ke arah digital. Dalam bisnis, website dan media sosial biasanya hanya berfungsi untuk mengenalkan dan memasarkan produk kepada konsumen. Padahal banyak hal lain yang dapat diperoleh melalui website dan media sosial seperti pola sentimen konsumen terhadap suatu produk. Analisis Sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif. Analisis sentimen diberikan oleh pengguna internet pada media sosial untuk memberikan suatu penilaian atau opini pribadi. Namun saat ini, belum ada aplikasi yang dapat melakukan proses analisis sentimen media sosial sekaligus menunjukkan wordcloud secara dinamis. Penelitian ini bertujuan untuk mengembangkan aplikasi yang mampu melakukan analisis sentimen media sosial dengan output berupa diagram jumlah sentimen positif/negatif/netral dan wordcloudnya, dimana machine learning polarity digunakan sebagai metode klasifikasi.

**Kata Kunci:** *Sentimen Analisis, Komentar Media Sosial, Aplikasi Web.*

## I. PENDAHULUAN

Yang menarik dari media sosial adalah kita dapat melakukan analisis terhadap opini pengguna media sosial terhadap suatu isu ataupun produk. di era digital, media sosial memiliki peranan yang sangat penting sebagai sumber pencetak informasi yang diperlukan. Tercatat di tahun 2019, pengguna media sosial di Indonesia terus meningkat yang mana media sosial youtube mengisi deretan pertama sebagai media sosial yang paling banyak digunakan dan dijadikan referensi untuk mendapatkan informasi.

Analisis sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif. Analisis sentimen diberikan oleh pengguna internet pada media sosial untuk memberikan suatu penilaian atau opini pribadi. Analisis komentar media sosial sangat berpotensi untuk meningkatkan popularitas. Diperlukan alat bantu untuk menganalisis sentimen pengguna media sosial (baik itu positif, negatif atau netral), agar influencer dan pelaku bisnis lebih responsif terhadap keadaan yang terjadi dilapangan, dan dapat mendukung keputusan perencanaan inovasi yang kompetitif dimasa depan.

Output dari penelitian ini berupa diagram dan wordcloud yang bertujuan untuk memudahkan pencari informasi terkait komentar instagram mendapatkan pengetahuan berupa persentase sentimen positif, negatif dan netral.

## II. STUDI LITERATUR

### 2.1 Text Mining

Text mining merupakan proses penambangan yang dilakukan oleh komputer untuk mendapatkan sesuatu yang baru, dan tidak diketahui sebelumnya, atau menemukan kembali informasi yang tersirat secara implisit. Temuan tersebut berasal dari informasi yang diekstrak secara otomatis dari sumber-sumber data teks yang berbeda. Text Preprocessing adalah bagian dari Text Mining yang dilakukan untuk menghapus noise pada kalimat. Text Preprocessing bertujuan untuk menghindari data yang kurang sempurna, gangguan pada data dan data yang tidak konsisten.

### 2.2 Sentimen Analisis

Analisis sentimen merupakan sebuah metode yang digunakan untuk mengekstrak data opini, memahami serta mengolah teksual data secara otomatis untuk melihat sentimen yang terkandung dalam sebuah opini.

### 2.3 Web Scraping

Web Scraping adalah metode otomatis untuk mendapatkan data dalam jumlah besar dari sebuah situs web. Sebagian besar data tersebut merupakan data yang tidak terstruktur dalam format HTML, kemudian diubah menjadi data terstruktur dalam spreadsheet atau database sehingga dapat digunakan dalam berbagai aplikasi.

Ada banyak cara berbeda untuk melakukan web scraping agar bisa mendapatkan data dari situs web. Anda dapat menggunakan layanan online, API tertentu, atau bahkan membuat kode web scraping Anda dari awal. Banyak situs web besar seperti Google, Twitter, Facebook, StackOverflow, dan sebagainya memiliki API yang memungkinkan Anda mengakses datanya dalam format terstruktur.

Ini bisa menjadi opsi terbaik, tetapi ada situs lain yang tidak mengizinkan pengguna untuk mengakses data dalam jumlah besar dalam bentuk terstruktur, atau mungkin situs tersebut tidak terlalu maju secara teknologi. Dalam situasi tersebut, sebaiknya gunakan web scraping untuk bisa mengekstrak data dari situs web tersebut.

Web scraping membutuhkan dua komponen, yaitu crawler dan scraper. Crawler adalah algoritma kecerdasan buatan yang menjelajahi web untuk mencari data tertentu yang diperlukan dengan mengikuti tautan di internet. Scraper adalah alat khusus yang dibuat untuk mengekstrak data dari situs web. Desain scraper dapat sangat bervariasi sesuai dengan kompleksitas dan ruang lingkup proyek sehingga Anda dapat mengekstrak data dengan cepat dan akurat.

### III. METODOLOGI PENELITIAN

Berikut adalah uraian dari tahapan dan metode yang digunakan dalam penelitian :

#### 2.1 Studi Kepustakaan

Pengumpulan bahan referensi, seperti jurnal penelitian, prosiding, tesis, buku-buku teori dan sumber-sumber lain termasuk informasi yang diperoleh melalui internet.

#### 2.2 Analisis dan Perancangan Sistem

Proses yang dilakukan dalam analisa dan perancangan sistem terbagi dalam beberapa bentuk diantaranya adalah sebagai berikut:

##### 2.2.1 Analisis Sistem

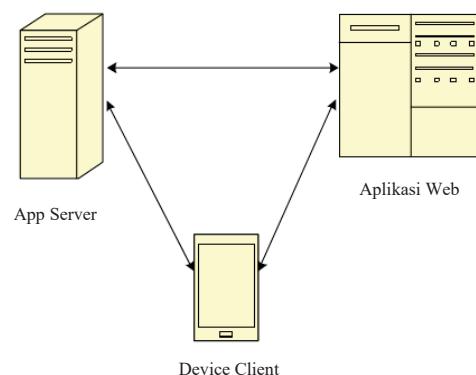
Sistem yang dibangun pada penelitian ini merupakan aplikasi web yang mampu menentukan sentimen ulasan komentar yang diberikan oleh user. Dalam penelitian ini, pendekatan yang digunakan merupakan pendekatan supervised learning. Metode klasifikasi yang digunakan adalah polarity. Proses evaluasi kinerja classifier menggunakan confusion matrix. Informasi dalam confusion matrix diperlukan untuk menentukan kinerja model klasifikasi. Hasil dari confusion matrix digunakan untuk menentukan akurasi dari classifier dengan performance metric accuracy.

##### 2.2.2 Perancangan Sistem Mobile Sentiment Analysis

Sistem yang dibangun memiliki tujuan untuk menentukan polaritas sentimen dari ulasan teksual komentar media sosial instagram/facebook/youtube. Dalam penelitian ini, komentar

diklasifikasikan menjadi tiga buah polaritas sentimen, yaitu sentimen positif(1), negatif(-1) dan netral(0). Adapun perancangan sistem terdiri dari:

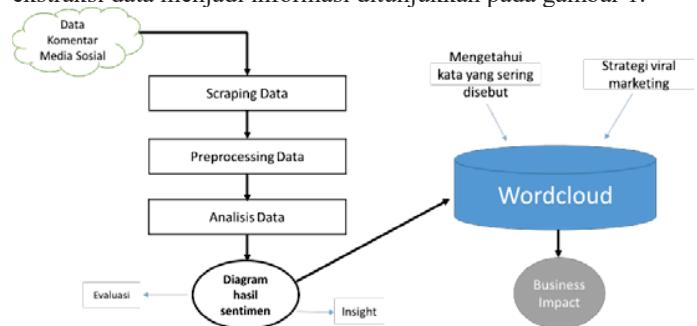
1. Perancangan arsitektur sistem Gambar umum dari arsitektur sistem yang dirancang dapat dilihat pada Gambar 1. Terdapat tiga komponen, komponen tersebut terdiri dari web, app server, dan device client. device client melakukan request analisis komentar ke aplikasi web dengan cara upload data. selanjutnya akan melalui tahap preprocessing kemudian dikirimkan ke app server untuk diklasifikasikan, data yang dikembalikan dari app server ke device client merupakan hasil klasifikasi. Dengan arsitektur sesuai pada Gambar 1, diharapkan dapat meminimalisir penggunaan resource perangkat agar nantinya sistem ini juga dapat digunakan pada perangkat mobile dengan spesifikasi perangkat keras yang rendah.



Gambar 1 Arsitektur sistem sentimen analisis

#### 2. Komponen dan proses pada perangkat lunak

Pada bagian ini akan dijelaskan tentang komponen dan proses dari app server ke device client. Langkah-langkah ekstraksi data menjadi informasi ditunjukkan pada gambar 1.



Gambar 2. Skema pengolahan data

##### a. Scraping Data

Proses pengumpulan data komentar instagram dengan cara,

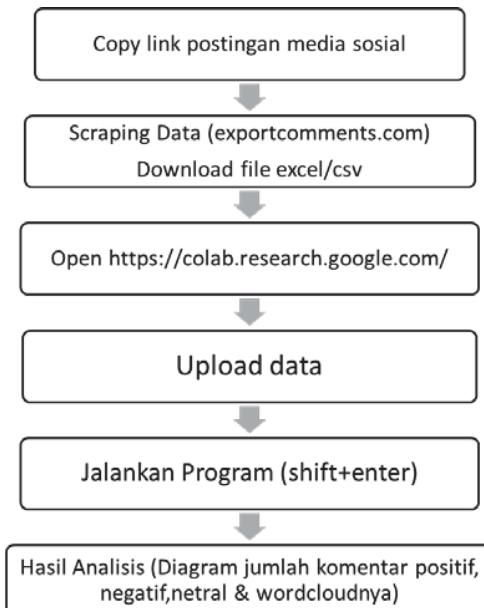
1. Salin link postingan Instagram yang akan di analisis
2. Open web exportcomments.com
3. Salin Link dan download dalam format xls(excel) atau csv.

### b. Preprocessing Data

Dataset yang telah didapatkan seringkali bersifat mentah, misal terdapat nilai yang kosong dan tidak konsisten. Akibatnya perlu dilakukan prapemrosesan data terlebih dahulu. Baris 1 berupa header/atribut, baris kedua dan seterusnya berupa isi.

### c. Analisis Data

Langkah-Langkah melakukan sentimen analisis komentar media sosial ditunjukkan pada gambar 3.



Gambar 3. Langkah analisis

Pada gambar 3, hasil visualisasi data berupa diagram dan wordcloud karena informasi yang ingin diperoleh hasil adalah perbandingan proporsi sentimen positif, negatif dan netral. Kemudian memunculkan wordcloud agar peneliti dapat mengetahui kata kunci yang sering muncul pada komentar postingan.

### d. Evaluasi

Evaluasi menggunakan confusion matrix untuk mengukur tingkat akurasi dalam pengujian metode. Setiap sel pada matriks, terdiri dari angka yang menunjukkan berapa banyak kasus yang diambil untuk direduksi(Hastuti, 2012). Bentuk dari tabel Confusion matrix bisa dilihat pada gambar 4.

		Actual Values	
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	TP (True Positive)	FP (False Positive) <small>Type I Error</small>
	0 (Negative)	FN (False Negative) <small>Type II Error</small>	TN (True Negative)

Gambar 4. Confusion Matrix

Keterangan:

- TP (True Positive) = Kondisi yang menyatakan nilai prediksi sesuai dengan nilai aktualnya
- FN (False Negative) = Kondisi yang menyatakan suatu prediksi bernilai negatif tetapi aktualnya justru bernilai positif
- FP (False Positive) = Kondisi yang menyatakan suatu prediksi bernilai positif tetapi aktualnya justru bernilai negative.
- TN (True Negative) = Kondisi yang menyatakan nilai prediksi sesuai dengan nilai aktualnya

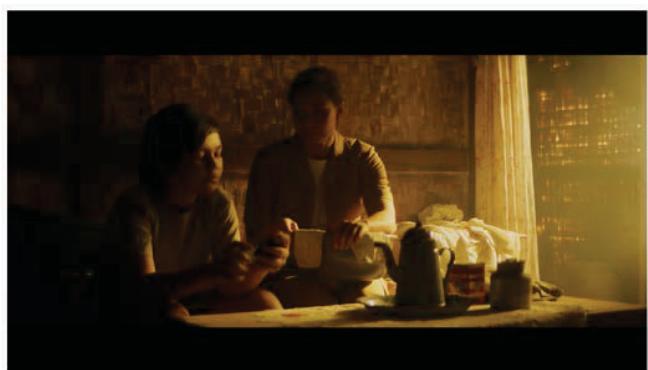
Dari gambar 3 kita dapat menghitung nilai akurasi yang didapat dari menghitung jumlah prediksi yang tepat pada semua kelas dibagi dengan jumlah keseluruhan data, rumusnya dapat dilihat pada Persamaan (1).

$$Akurasi = \frac{TP(A)+TP(B)}{N} \quad (1)$$

## IV. HASIL PENELITIAN

### 4.1 Mengambil Data

Pada proses ini, langkah pertama adalah menentukan ulasan komentar media sosial yang akan di analisa. Pada pengujian ini, akan dianalisis ulasan komentar video youtube milik akun Indonesian Youtubers dengan 1,55 juta subscriber. Video dengan judul Rewind Indonesia 2021 ditunjukkan pada Gambar 5.



#3 DI TRENDING  
REWIND INDONESIA 2021

10.945.900 x ditonton... 1,4 JT TIDAK SUKA BAGIKAN SIMPAN ...

Gambar 5. Video YouTube akun Indonesian Youtubers

Link video pada Gambar 5 adalah <https://www.youtube.com/watch?v=IPZO85HFM4>. Kemudian input link di web exportcomments.com dan download data komentar dalam bentuk file xls atau csv.

## 4.2 Text Preprocessing

Rangkaian kegiatan pra-pemrosesan teks diawali dengan pengubahan bentuk teks menjadi huruf kecil semua. Hal ini telah terbukti mempercepat waktu pemrosesan secara keseluruhan. Kemudian mengilangkan tanda baca agar komentar dapat dianalisis per kata dengan polarity.

### 4.3 Analisa Data

Proses analisis diawali dengan melabeli setiap komentar menggunakan tools textblob.

```
#Calculating the Sentiment Polarity
pol=[] # list which will contain the polarity of the comments
for i in df.Comment.values:
    try:
        analysis =TextBlob(i)
        pol.append(analysis.sentiment.polarity)

    except:
        pol.append(0)
```

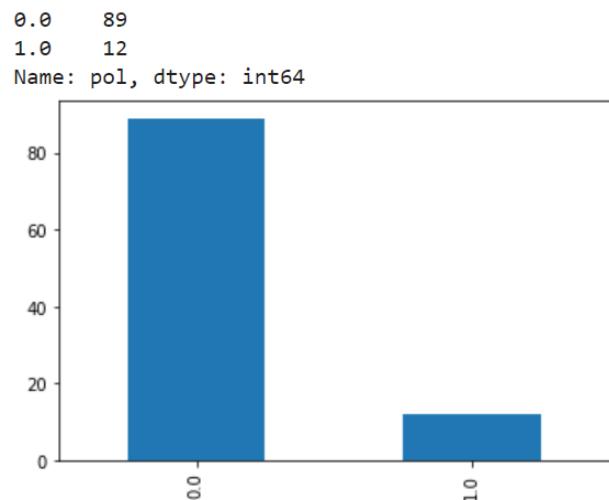
Gambar 6. machine learning polarity python

Textblob merupakan salah satu tools atau lebih tepatnya library untuk pemrosesan dibidang "Natural Language Processing (NLP)" menggunakan bahasa python. Melalui textblob, kita dapat melakukan berbagai proses terhadap data teks mulai dari yang sederhana seperti tokenisasi (pemotongan kata) sampai analisa sentiment.

Hasilnya terlihat pada gambar 7.

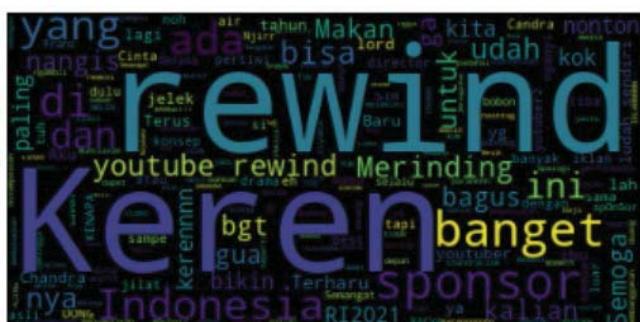
Gambar 7. Sampel hasil labeling data dengan polarity

Langkah berikutnya mengelompokkan hasil sentimen analisis dengan ketentuan, sentimen positif diwakili angka 1, sentimen negatif diwakili angka -1 dan netral diwakili angka 0. Diagram sentimen ulasan komentar video youtube rewind indonesia 2021 ditunjukkan gambar 8.



Gambar 8. Diagram analisis sentimen komentar video rewind Indonesia 2021

Diagram diatas menunjukan bahwa mayoritas komentar postingannya bersifat netral. Kemudian kata yang paling sering disebut pada komentar video tersebut dapat di liat pada Gambar 9.



Gambar 9. Wordcloud analisis sentimen komentar video  
rewind Indonesia 2021

Semakin besar kata, semakin popular kata dalam ulasan komentar video rewind Indonesia 2021. Contoh kata ‘keren banget’ merupakan kata yang sering disebut didalam komentar. Wordcloud dapat dimanfaatkan sebagai evaluasi dan strategi viral marketing diproject berikutnya.

Untuk mendapatkan data hanya komentar positif dapat dilakukan seperti gambar 10.

```
#Displaying the POSITIVE comments  
df_positive = df[df.pol==1]  
df_positive.head(10)
```

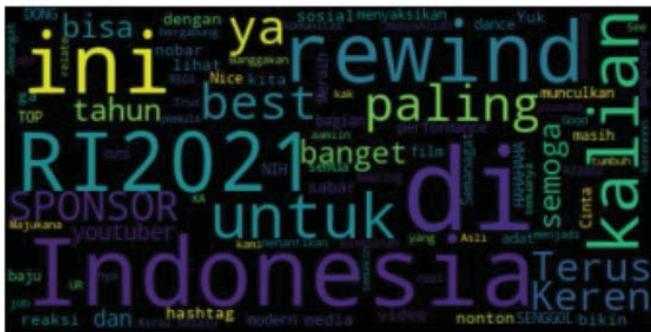
Gambar 10. Memfilter data sentimen positif

Hasilnya terlihat pada gambar 11.

Comment	(view source)
REWIND INDONESIA 2021 #RI2021 Yuk bikin video nobar/montir/reksa kalian di sosial media dengan hastag #RI2021 . nrikita ga sabar untuk lihat kalian menyajikan performance film ini #RI2021 D	<a href="#">view comment</a> 1.0
0.0 bagian paling the best	<a href="#">view comment</a> 1.0
rewind 2014 masih paling best	<a href="#">view comment</a> 1.0
TOP	<a href="#">view comment</a> 1.0
baik adet g d munculkan..... dance modern semua ya	<a href="#">view comment</a> 1.0
SPONSOR NIH, SENGGOAL DONG HAHAHAAA	<a href="#">view comment</a> 1.0
Nice Indonesia Semangat Terus Untuk Meraih Gintis Banggakan Terus Indonesia Majukan Trus Indonesia	<a href="#">view comment</a> 1.0
Amazing	<a href="#">view comment</a> 1.0
KA REGI UR SO CUTE	<a href="#">view comment</a> 1.0
masyallah. inKeren banget kak. Bismillah semoga di tahun 2022 in, kami bisa bergabung menjadi youtuber yang bisa berkembang dan tumbuh di komunitas youtube. aamin. 'n'insyaratkan youtuber pemula ??	<a href="#">view comment</a> 1.0
Asli keren banget Di penghujung tahun salut manantikan ini. Kira2 di 2021 reward nya apa ya dan semuanya relate. Good job all See you di RI2022 Semoga seunik kerennin ??	<a href="#">view comment</a> 1.0
Sponsor is real	<a href="#">view comment</a> 1.0

Gambar 11. Komentar dengan sentimen positif

Kemudian kata yang paling sering muncul pada komentar video tersebut dapat di liat pada Gambar 12.



Gambar 12. Wordcloud sentimen positif

Untuk mendapatkan data hanya komentar negatif dapat dilakukan seperti gambar 13.

```
#Displaying the NEGATIVE comments  
negative = df[df.pol== -1]  
negative.head(10)
```

Gambar 13. Memfilter data sentimen negatif

Hasilnya terlihat pada gambar 14.

no Name (click to view profile) Date Likes isHearted isPinned Comment (view source) pol

Gambar 14. Komentar dengan sentimen negatif

Karena tidak ada komentar negatif pada video rewind tersebut maka tidak ada ulasan komentar yang dimunculkan.

Untuk mendapatkan data hanya komentar netral dapat dilakukan seperti gambar 15.

```
#Displaying the NEUTRAL comments
```

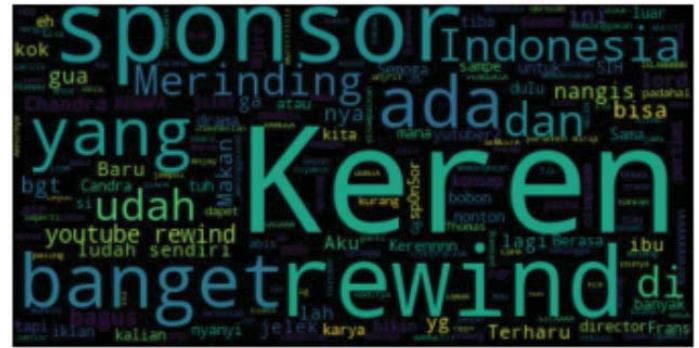
```
netral = df[df.pol==0]  
netral.head(10)
```

Gambar 15. Memfilter data sentimen netral

Hasilnya terlihat pada gambar 16.

Gambar 16. Komentar dengan sentimen netral

Kemudian kata yang paling sering muncul pada komentar video tersebut dapat di liat pada Gambar 17.



Gambar 17. Wordcloud sentimen netral

#### 4.4 Evaluasi Model

Tahap evaluasi model dilakukan dengan menggunakan instrumen Confusion Matrix dengan hasil sebagaimana terlihat di Tabel 1.

Tabel 1. Confusion matrix

		Aktual		
		positif	neutra	negatif
klasifikasi	positif	20		
	neutra	6	52	12
	negatif			0

Tabel 1 menunjukkan 72 data diklasifikasikan benar. Terdapat 18 kesalahan dalam klasifikasi karena NLP textblob tidak bisa menerjemahkan sentimen kata slank atau gaul bahasa indonesia. Berdasarkan tabel 4 diperoleh akurasi sebesar 80%.

Tantangan pada penelitian ini adalah video rewind Indonesia 2021 tersebut lebih banyak komentar yang menyindir dan tidak baku. NLP textblob belum bisa mendeteksi bahasa gaul dan sindiran.

## KESIMPULAN

Kesimpulan yang didapatkan dari hasil penelitian yang telah dilakukan, sebagai berikut:

1. Analisis sentimen komentar media sosial dengan NLP polarity mendapatkan akurasi 80%
2. Analisis sentimen komentar media sosial dapat dimanfaatkan sebagai evaluasi bisnis bagi pelaku usaha atau improvement bagi Influencer.
3. Prototipe dapat dijalankan tanpa aplikasi (dapat dibuka di web). Sehingga memudahkan siapapun melakukan analisis komentar media sosial.

Saran-saran yang diusulkan agar penelitian ini dapat menjadi lebih baik di masa mendatang yaitu melanjutkan ke tahap

pembuatan aplikasi berbasis web dan mobile agar mempermudah pengguna melakukan analisis sentimen.

## DAFTAR PUSTAKA

- [1] Gbashi, S., Adebo, O. A., Doorsamy, W., & Njobeh, P. B. (2021). Systematic Delineation of Media Polarity on COVID-19 Vaccines in Africa: Computational Linguistic Modeling Study. JMIR medical informatics, 9(3), e22916.
- [2] Luqvana, W. A., Cholissodin, I., & Perdana, R. S. (2018). Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN, 2(11), 4704-4713.
- [3] Oyebode, O., & Orji, R. (2019). Social media and sentiment analysis: The Nigeria presidential election 2019. In 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (pp. 0140-0146). IEEE.
- [4] Puspitarani, Y. (2015). Sentimen Analysis Terhadap Nilai Kepercayaan Sebuah Online Shop Di Instagram. Jurnal Ilmiah Teknologi Infomasi Terapan, 2(1).