

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Landasan Teori**

##### **2.1.1 DANA**



Gambar 2.1 Logo DANA  
(Wikipedia, 2023)

DANA adalah aplikasi dompet digital yang berasal dari Indonesia yang tersedia di Google Play Store. DANA digunakan untuk pembayaran digital *cashless* dan *cardless* baik online maupun offline. DANA mendesain fitur mulai dari pembelian pulsa, tagihan listrik, hingga *top up*. Tujuannya untuk menciptakan kemudahan bagi pengguna dalam melakukan transaksi dengan cepat, nyaman dan aman (Larasati, Ratnawati and Hanggara, 2022).

##### **2.1.2 Analisis Sentimen**

Analisis sentimen adalah gabungan dari dua pemrosesan, yaitu *text mining* dan *data mining* yang memiliki fungsi untuk mengkategorikan berbagai ulasan, pendapat, opini atau sentimen dari konsumen maupun pakar tentang suatu layanan, produk, organisasi atau aktivitas tertentu (Rizki et al., 2023). Selain itu, analisis sentimen disebut juga proses klasifikasi teks yang menentukan suatu teks tersebut mempunyai makna positif, netral atau negatif. Tujuan analisis sentimen adalah untuk mendistribusikan sentimen publik

secara otomatis melalui proses menafsirkan, mengekstraksi dan mengolah data (Auliya et al., 2023).

Proses analisis sentimen diawali dengan pembuatan data latih, yang selanjutnya diproses algoritma klasifikasi tertentu guna menghasilkan model klasifikasi. Proses pembuatan data latih diawali dari tahap *text preprocessing*. Tahapan *text preprocessing* yaitu *case folding*, *tokenizing*, *spelling correction*, *filtering* dan *stemming*. Kemudian data yang telah di *preprocessing* dilakukan *labeling*. Data kemudian diklasifikasi dan diuji dengan *confusion matrix* guna mengukur performa model klasifikasi (Ilmawan and Mude, 2020).

### **2.1.3 Web Scraping**

*Web Scraping* merupakan metode untuk mengambil data atau informasi dari halaman *web* secara otomatis. Prinsipnya untuk mengeksplorasi informasi dari berbagai *website* yang tidak memiliki struktur tertentu, kemudian merapkannya dan mengubahnya ke dalam format yang lebih teratur dan terstruktur seperti spreadsheet, basis data, atau csv (Parasati, Abdurrachman Bachtiar and Setiawan, 2020).

### **2.1.4 Data Preparation**

*Data Preparation* merupakan proses menghilangkan dan mengatasi *noise data* agar hasil klasifikasi menjadi optimal. Selain itu, data dipersiapkan dan ditransformasikan sebelum tahap permodelan. (Azhar, Surahman and Juliane, 2022).

a. *Case Folding*

*Case Folding* merupakan proses mengubah format teks menjadi huruf kecil atau *lowercase*, penghapusan angka, tanda baca, emoji dan karakter lain yang tidak diperlukan (Baita, Pristyanto and Cahyono, 2021).

b. *Tokenizing*

*Tokenizing* merupakan proses perubahan pada kalimat ulasan jadi bagian-bagian kata (Mas Diyasa et al., 2021). Proses ini dilakukan untuk memberikan kemudahan dalam menghilangkan *stopwords*.

c. *Spelling Correction*

*Spelling Correction* adalah proses untuk memperbaiki kesalahan dalam penulisan. Ketika kesalahan penulisan ini tidak diperbaiki, hasil analisis yang dihasilkan dapat terpengaruh (Khaira et al., 2020).

d. *Filtering*

*Filtering* adalah proses menghilangkan kata-kata yang termasuk *stopword*. *Stopword* merupakan kata-kata yang sering muncul tetapi dianggap tidak memiliki makna (Fitri, 2020).

e. *Stemming*

*Stemming* merupakan proses pengolahan kata-kata pada ulasan untuk membuat kata yang telah mengalami imbuhan dikembalikan pada kata dasarnya dengan aturan-aturan tertentu (Rizki et al., 2023).

### **2.1.5 *Textblob***

*Textblob* adalah *library python* untuk memproses data teks. *Textblob* menyediakan API yang digunakan untuk pemrosesan bahasa alami,

diantaranya ekstraksi frase kata benda, analisis sentimen, klasifikasi, dan terjemahan (Azhar, Surahman and Juliane, 2022).

*Textblob* dapat digunakan untuk proses pelabelan data otomatis. *Textblob* menghitung nilai polaritas (*polarity*) dan subjektivitas (*subjectivity*). Nilai polaritas adalah fungsi yang mendeteksi sentimen suatu kalimat dan nilai subjektivitas adalah fungsi yang mendeteksi nilai suatu kalimat. Jenis nilai suatu kalimat dapat berbentuk opini atau fakta. Semakin tinggi nilai subjektivitas maka termasuk dalam sebuah opini. Semakin tinggi nilai polaritas maka termasuk ke sentimen positif. Nilai *polarity* dan *subjectivity* diperoleh menggunakan persamaan 2.1.

$$Polarity, Subjectivity = \frac{x_1 + x_2 \dots x_n}{n} \quad (2.1)$$

(Aaron, 2015)

Keterangan:

$x$  : nilai polaritas setiap kata penyusun

$n$  : Jumlah seluruh kata penyusun

Terdapat tiga kelas sentimen yang ditentukan berdasarkan nilai *polarity* diantaranya positif, negatif, dan netral (Baita, Pristyanto and Cahyono, 2021).

Tabel 2.1 merupakan ketentuan nilai *Textblob* pada pelabelan data.

Tabel 2.1 Ketentuan Pelabelan *Textblob*

| Sentimen | <i>Polarity</i> (P) |
|----------|---------------------|
| Positif  | $P > 0$             |
| Netral   | $P = 0$             |
| Negatif  | $P < 0$             |

Berdasarkan Tabel 2.1 proses pelabelan data ulasan menggunakan *Textblob* didasarkan nilai polaritas. Jika nilai polaritas di atas 0 maka termasuk sentimen positif. Jika nilai polaritas 0 maka termasuk sentimen netral. Jika nilai polaritas di bawah 0 maka termasuk sentimen negatif.

### 2.1.6 Valence Aware Dictionary and sEntiment Reasoner (VADER)

*Vader* adalah alat analisis sentimen berbasis kosakata dan aturan (*lexicon-and rule-based*) yang secara spesifik disesuaikan dengan sentimen yang diungkapkan di media sosial dan platform ulasan (Ghudafa, Akbar and Srisulistiowati, 2021). *Vader* menggunakan kamus leksikon dalam mengkalkulasikan polaritas kata. Nilai polaritas kata digunakan untuk menghitung *compound score* yang merupakan skor untuk menghitung jumlah nilai semua kosakata (Effendi and Noviana, 2021).

$$\text{Compound Score} = \frac{n}{\sqrt{n^2 + \alpha}} \quad (2.2)$$

(Lee, 2021)

Keterangan:

$n$  : total skor kata penyusun

$\alpha$  : konstanta (*default* = 15)

Tabel 2.2 merupakan ketentuan *compound score* dalam pelabelan *Vader*.

Tabel 2.2 Ketentuan Pelabelan *Vader*

| Sentimen | Skor (x)           |
|----------|--------------------|
| Positif  | $x \geq 0.05$      |
| Netral   | $-0.05 < x < 0.05$ |
| Negatif  | $x \leq -0.05$     |

Berdasarkan Tabel 2.2, proses pelabelan *Vader* mengacu pada *compound score*. Jika *compound score* lebih besar dari 0,05 maka termasuk sentimen positif. Jika *compound score* berkisar antara kurang dari 0,05 hingga lebih dari -0,05 maka termasuk sentimen netral. Jika *compound score* kurang dari -0,05 maka termasuk sentimen negatif.

### 2.1.7 TF-IDF (*Term Frequency - Inverse Document Frequency*)

TF-IDF adalah metode pembobotan setiap kata dalam dokumen untuk menentukan seberapa baik istilah tersebut mewakili suatu ulasan (Zhafira, Rahayudi and Indriati, 2021). TF-IDF terbagi menjadi TF (*Term Frequency*) dan IDF (*Inverse Document Frequency*). TF yaitu frekuensi setiap kata dalam dokumen. IDF merupakan jumlah nilai dokumen setiap kata dan berbanding terbalik (Septian, Fachrudin and Nugroho, 2019).

#### a. TF (*Term Frequency*)

Nilai TF diperoleh menggunakan persamaan 2.3.

$$TF(t_i, d) = \frac{f(t, d)}{n(t, d)} \quad (2.3)$$

(Chouksey, 2020)

Keterangan:

- $t_i$  : kata ke-  $i$
- $d$  : dokumen
- $f(t, d)$  : frekuensi  $t$  dalam  $d$
- $n(t, d)$  : jumlah  $t$  dalam  $d$

b. IDF (*Inverse Document Frequency*)

Nilai IDF diperoleh menggunakan persamaan 2.4.

$$(IDF(t_i) = \ln \frac{|D|}{DF(t_i)} \quad (2.4)$$

(Chouksey, 2020)

Keterangan:

- $\ln$  : logaritma alami  
 $|D|$  : jumlah semua dokumen  
 $DF(t_i)$  : jumlah dokumen yang mengandung kata t

c. TF-IDF (*Term Frequency Inverse Document Frequency*)

Nilai TF-IDF diperoleh menggunakan persamaan 2.5.

$$TFIDF = TF(t_i, d) * IDF(t_i) \quad (2.5)$$

(Chouksey, 2020)

Selanjutnya dilakukan normalisasi *euclidean* menggunakan persamaan 2.6.

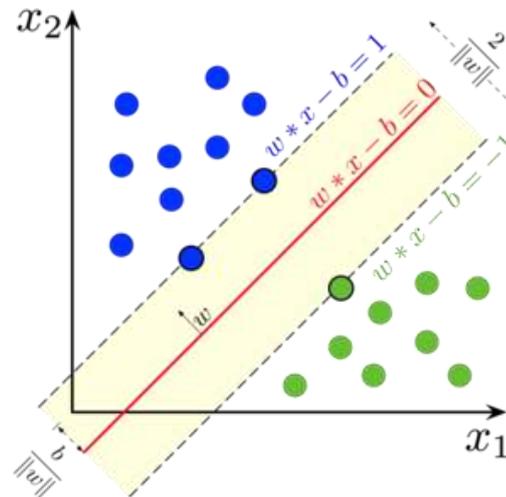
$$V_{norm} = \frac{v}{\sqrt{v_1^2 + v_2^2 + \dots + v_n^2}} \quad (2.6)$$

(Chouksey, 2020)

### 2.1.8 Support Vector Machine (SVM)

SVM adalah algoritma *supervised learning* yang digunakan klasifikasi dengan menentukan *hyperplane*. *Hyperplane* yang sesuai terletak tepat di tengah-tengah kedua kelas dan mempunyai jarak terjauh terhadap data terluar

kedua kelas (Baita, Pristyanto and Cahyono, 2021). Gambar 2.2 merupakan penjelasan kinerja SVM.



Gambar 2.2 Kinerja algoritma SVM

(Tokuç, 2023)

Berdasarkan Gambar 2.2 garis merah merupakan *hyperplane* yang mengklasifikasikan kedua titik tersebut, dimana titik data pada salah satu sisi diberi label kelas negatif yaitu -1, kelas positif diberi label +1, dan kelas netral diberi label dengan 0 berada di tengah garis merah. Data terdekat terhadap *hyperplane* disebut dengan *support vector*. Jarak antara *hyperplane* dan *support vector* disebut dengan Margin. Kernel yang digunakan adalah kernel *linear*. Hal ini karena kernel *linear* merupakan kernel yang sederhana dan sering digunakan untuk kasus klasifikasi teks (Rahman Isnain et al., 2021).

*Support Vector Machine* (SVM) memiliki metode dalam membangun dan melatih model klasifikasi yaitu *Sequential Training*. Tahapan dari metode tersebut dijelaskan sebagai berikut.

a. Inisiasi parameter nilai *Complexity* ( $C$ ), *epsilon* ( $\varepsilon$ ), *alpha* ( $\alpha$ ), *lamda* ( $\lambda$ ), *gamma* ( $\gamma$ )

b. Menghitung fungsi kernel *linear*

$$K(x_a, x_b) = (x_a * x_b) \quad (2.7)$$

(Turmudi and Hadikristanto, 2023)

c. Menghitung matriks *Hessian*

$$D_{ab} = y_a y_b (K(x_a, x_b) + \lambda^2) \quad (2.8)$$

(Turmudi and Hadikristanto, 2023)

Keterangan:

$D_{ab}$  : elemen matriks a dan b

$y_a$  : label data a

$y_b$  : Label data b

d. Menghitung nilai *error*, *delta alpha* dan *alpha* baru berdasarkan iterasi

$$E_i = \sum \alpha_i D_{ab} \quad (2.9)$$

$$\delta\alpha_i = \min \{ \max[\gamma(1 - E_i), \alpha_i], C - \alpha_i \} \quad (2.10)$$

$$\alpha_i = \alpha_i + \delta\alpha_i \quad (2.11)$$

(Turmudi and Hadikristanto, 2023)

Keterangan:

$E_i$  : nilai *error*

$\delta\alpha_i$  : parameter *delta alpha*

e. Menghitung bias

$$b = -\frac{1}{2} * \sum \alpha_i y_i (K(x_i, x^+) + \sum \alpha_i y_i (K(x_i, x^0) + \sum \alpha_i y_i (K(x_i, x^-) \quad (2.12)$$

(Turmudi and Hadikristanto, 2023)

f. Menghitung nilai keputusan

$$\text{sign}(h(x)) = \sum \alpha_i y_i (K(x_a, x_b)) + b \quad (2.13)$$

(Turmudi and Hadikristanto, 2023)

### 2.1.9 Confusion Matrix

*Confusion Matrix* merupakan matriks untuk mengukur performa model klasifikasi melalui perhitungan akurasi, presisi, *recall*, dan *f1-score* (Auliya et al., 2023). Tabel 2.3 merupakan kerangka tabel *confusion matrix*.

Tabel 2.3 *Confusion Matrix*

|        |         | Prediksi |        |         |
|--------|---------|----------|--------|---------|
|        |         | Negatif  | Netral | Positif |
| Aktual | Negatif | TN       | FNeu   | FP      |
|        | Netral  | FN       | TNeu   | FP      |
|        | Positif | FN       | FNeu   | TP      |

Berdasarkan Tabel 2.3 terdapat 4 indikator yang dihitung untuk menguji performa, yaitu:

a. Akurasi

Akurasi yaitu nilai dari seberapa akurat model klasifikasi. Nilai akurasi diperoleh menggunakan Persamaan 2.14.

$$Akurasi = \frac{TP + TNeu + TN}{jumlah\ data\ uji} \quad (2.14)$$

(Rina, 2023)

## b. Presisi

Presisi yaitu nilai dari perbandingan prediksi benar dengan prediksi secara keseluruhan. Nilai presisi diperoleh menggunakan Persamaan 2.15.

$$Presisi = \frac{TP}{TP + FP} \quad (2.15)$$

(Rina, 2023)

## c. Recall

Recall yaitu nilai dari perbandingan prediksi benar dengan hasil yang benar keseluruhan. Nilai *recall* diperoleh menggunakan Persamaan 2.16.

$$Recall = \frac{TP}{TP + FN + FNeu} \quad (2.16)$$

(Rina, 2023)

## d. F1-Score

F1-score yaitu nilai dari rata-rata harmonik presisi dan *recall*. Nilai *f1-score* diperoleh menggunakan persamaan 2.17.

$$F1 - score = \frac{2 * Presisi * Recall}{Presisi + Recall} \quad (2.17)$$

(Rina, 2023)

Keterangan:

*True Positive (TP)* : Jumlah data positif yang diprediksi positif

- True Negative (TN)* : Jumlah data negatif yang diprediksi negatif
- True Neutral (TNeu)* : Jumlah data netral yang diprediksi netral
- False Positive (FP)* : Jumlah data positif yang diprediksi negatif  
atau netral
- False Negative (FN)* : Jumlah data negatif yang diprediksi positif  
atau netral
- False Neutral (FNeu)* : Jumlah data netral yang diprediksi negatif  
atau positif

## 2.2 Penelitian Terkait dan Kebaruan Penelitian

### 2.2.1 *State of The Art (SOTA)*

Tabel 2.4 merupakan perbandingan penelitian yang berkaitan dengan masalah penelitian dan hasil penelitian.

Tabel 2.4 SOTA

| No | Konten            | Deskripsi  |
|----|-------------------|--|
| 1  | Judul paper       | Analisis Sentimen Komentar YouTube TvOne Tentang Ustadz Abdul Somad Dideportasi Dari Singapura Menggunakan Algoritma SVM   |
|    | Penulis           | Desti Mualfah, Ramadhoni, Rahmad Gunawan, Danang Mulyadipa Suratno   |
|    | Jurnal/Konferensi | JURNAL FASILKOM Volume 13 No. 1   April 2023: 72-80  |
|    | URL               | <a href="https://ejurnal.umri.ac.id/index.php/JIK/article/view/4920/2353">https://ejurnal.umri.ac.id/index.php/JIK/article/view/4920/2353</a>  |
|    | Permasalahan      | Channel YouTube Berita TvOne mendapat banyak respon dari publik sehingga menjadikan saluran ini sebagai salah satu saluran terpopuler di Indonesia dengan antusias yang tinggi serta mendapat banyak respon dari masyarakat.   |
|    | Hasil utama       | Algoritma SVM digunakan untuk mengklasifikasikan komentar video YouTube di channel TvOne tentang deportasi Ustad Abdul Somad Singapura dengan akurasi 95,02%, recall 95,02%, presisi 95,18%, dan F1-score 95,01%. Penerapan SVM pada penelitian ini memiliki akurasi yang tinggi dan mampu memberikan solusi terhadap permasalahan analisis sentimen komentar video YouTube. |

| No | Konten            | Deskripsi  |
|----|-------------------|--|
|    | Batasan           | Sebagai pembanding performa yang didapatkan maka diperlukan penambahan algoritma lain untuk melakukan klasifikasi.   |
| 2  | Judul paper       | Analisis Sentimen Masyarakat Indonesia Terhadap Produk Kendaraan Listrik Menggunakan VADER   |
|    | Penulis           | Muchamad Taufiq Anwar, Denny   |
|    | Jurnal/Konferensi | Jurnal Teknik Informatika dan Sistem Informasi Vol. 10, No. 1, Maret 2023, Hal. 783-792  |
|    | URL               | <a href="https://jurnal.mdp.ac.id/index.php/jatisi/article/download/3406/1173">https://jurnal.mdp.ac.id/index.php/jatisi/article/download/3406/1173</a>  |
|    | Permasalahan      | Dalam masa transisi adopsi kendaraan listrik saat ini, beberapa produsen kendaraan listrik telah melepas produknya ke pasaran. Setelah memperkenalkan sebuah produk ke publik, pihak produsen tentu ingin mengetahui respons publik terhadap produk yang diperkenalkan tersebut.   |
|    | Hasil utama       | Pada produk kendaraan listrik yang dibahas, sebagian besar opini masyarakat Indonesia terhadap produk tersebut adalah positif, yaitu terkait dengan fitur, desain, kecanggihan, dan ramah lingkungan. Sementara sentimen yang negatif yaitu tentang harga dan ketiadaan fast charging. Hal ini dapat menjadi masukan bagi perusahaan dan ditindaklanjuti oleh perusahaan dalam pembuatan strategi bisnis ke depan. |
|    | Batasan           | Penelitian selanjutnya diharapkan berfokus pada mengelompokkan sentimen berdasarkan topik-topiknya menggunakan metode topic modelling seperti LDA ( <i>Latent Dirichlet Allocation</i> )   |
| 3  | Judul paper       | Pelabelan Otomatis Lexicon Vader dan Klasifikasi Naive Bayes dalam Menganalisis Sentimen Data Ulasan PLN Mobile  |

| No | Konten            | Deskripsi  |
|----|-------------------|--|
|    | Penulis           | Yessy Asri, Widya Nita Suliyanti, Dwina Kuswardani, Muhamad Fajri  |
|    | Jurnal/Konferensi | PETIR: Jurnal Pengkajian dan Penerapan Teknik Informatika Vol. 15, No. 2, September 2022, P-ISSN 1978-9262, E-ISSN 2655-5018   |
|    | URL               | <a href="https://jurnal.itpln.ac.id/petir/article/view/1733">https://jurnal.itpln.ac.id/petir/article/view/1733</a>  |
|    | Permasalahan      | Ulasan di Google Playstore mempunyai rating 1 hingga 5, namun tidak cukup untuk menggambarkan kualitas aplikasi karena sering kali pengguna memberikan rating yang tidak sesuai dengan ulasan. Teks dalam ulasan lebih menggambarkan reaksi pengguna terhadap aplikasi seluler PLN. Ulasan ini dapat mempengaruhi pengunjung dan pengguna yang baru saja mengunduh aplikasi. |
|    | Hasil utama       | Berdasarkan perbandingan kelas positif, netral, serta negatif dengan menggunakan 1.000 sampel data Vader, kelas positif rating 67%, kelas netral rating 6%, dan kelas negatif rating 27%. Proses klasifikasi juga menggunakan metode NBC. Untuk pendistribusian data uji dan data latih dengan rasio 90:10. Tingkat akurasi proses evaluasi matriks konfusi adalah 70%.      |
|    | Batasan           | Penelitian selanjutnya diharapkan berfokus pada penambahan dataset agar dapat mengoptimalkan tingkat akurasi.  |
| 4  | Judul paper       | Analisis Sentimen Terhadap Cryptocurrency Berbasis Python TextBlob Menggunakan Algoritma Naïve Bayes   |
|    | Penulis           | Rizaldi Azhar, Adi Surahman, Christina Juliane   |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    | Jurnal/Konferensi | Jurnal Sains Komputer & Informatika (J-SAKTI) Volume 6 Nomor 1, Maret 2022, pp. 267-281   |
|    | URL               | <a href="http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/443">http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/443</a>   |
|    | Permasalahan      | Studi ini berfokus pada hashtag kriptografi dan menentukan tingkat sentimen komunitas global terhadap hashtag tersebut. Karena banyaknya opini tentang hashtag ini dari komunitas di seluruh dunia, sangat sulit untuk menentukan sentimen positif, negatif, atau netral secara manual.   |
|    | Hasil utama       | Sebanyak 1032 tweet hasil pelabelan otomatis dengan <i>Textblob</i> menghasilkan 632 tweet atau 61,24% berisi sentimen positif, 296 tweet atau 28,68% tweet berisi sentimen netral, dan 104 tweet atau 10,07% berisi sentimen negatif. Hasil pengujian algoritma NBC pada data latih 0,8 dan data uji 0,2 menghasilkan skor akurasi 71,98%, presisi 83,04%, recall 60,88%, dan f1_score 65,07%. |
|    | Batasan           | Perlu adanya peningkatan jumlah data yang digunakan dalam penelitian.   |
| 5  | Judul paper       | Analisis Sentimen Pengguna Aplikasi Dana Berdasarkan Ulasan Pada Google Play Menggunakan Metode Support Vector Machine  |
|    | Penulis           | Abitdavy Athallah Muhammad, Ermatita, Desta Sandya Prasvita   |
|    | Jurnal/Konferensi | Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA) Jakarta-Indonesia   |
|    | URL               | <a href="https://conference.upnvj.ac.id/index.php/senamika/article/view/2171">https://conference.upnvj.ac.id/index.php/senamika/article/view/2171</a>   |
|    | Permasalahan      | DANA merupakan salah satu <i>financial technology</i> (Fintech) yang menyediakan fitur transaksi melalui <i>smartphone</i> pengguna. Banyaknya fitur yang disediakan DANA, kemungkinan besar mendapatkan  |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    |                   | masukan dari pengguna DANA dalam ulasan di Google Play Store. Contohnya jika fitur isi ulang rusak atau gagal, tidak menerima pengembalian uang atau kesulitan menghubungi layanan pelanggan.   |
|    | Hasil utama       | Hasil evaluasi menunjukkan model SVM dengan akurasi 87,58%, presisi 91,20%, dan recall 90,21% sedangkan chi-square SVM dengan akurasi 89,41%, presisi 93,29%, dan <i>recall</i> 90,76%.   |
|    | Batasan           | Pelabelan data yang dilakukan hanya menggunakan metode manual sehingga tidak ada pembandingan dengan data yang dilakukan pelabelan secara otomatis.   |
| 6  | Judul paper       | Analisis Sentimen Mengenai Vaksin Sinovac Menggunakan Algoritma Support Vector Machine (SVM) Dan K-Nearest Neighbor (KNN)   |
|    | Penulis           | Anna Baita, Yoga Pristyanto, Nuri Cahyono   |
|    | Jurnal/Konferensi | Information System Journal (INFOS)   Vol. 4, No. 2, November 2021   |
|    | URL               | <a href="https://jurnal.amikom.ac.id/index.php/infos/article/view/687">https://jurnal.amikom.ac.id/index.php/infos/article/view/687</a>   |
|    | Permasalahan      | Vaksinasi sedang hangat diperbincangkan di berbagai platform media sosial, termasuk Twitter. Dengan menganalisis sentimen dari berbagai jenis reaksi netizen di Twitter, dapat mengategorikan jenis opini tersebut menjadi sentimen positif, negatif, dan netral. |
|    | Hasil utama       | <i>Textblob</i> digunakan untuk proses pelabelan teks. Hasil penelitian mendeskripsikan SVM mempunyai performa yang lebih baik jika dibandingkan dengan KNN. Akurasi SVM 0,7, sedangkan akurasi KNN 0,56.   |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    | Batasan           | Nilai akurasi SVM atau KNN relatif rendah, disebabkan oleh penggunaan <i>Textblob</i> dalam proses pelabelan sentimen. Oleh karena itu, penelitian selanjutnya juga dapat dicoba menggunakan pelabelan manual atau juga menggunakan platform lain seperti <i>vader</i> atau <i>spacy</i> .  |
| 7  | Judul paper       | Optimasi TextBlob Menggunakan Support Vector Machine untuk Analisis Sentimen (Studi Kasus Layanan Telkomsel)  |
|    | Penulis           | Kevin Perdana, Titania Pricillia, Zulfachmi   |
|    | Jurnal/Konferensi | Bangkit Indonesia, Vol. X, No. 01, Bulan Maret 2021   |
|    | URL               | <a href="https://journal.sttindonesia.ac.id/index.php/bangkitindonesia/article/view/120">https://journal.sttindonesia.ac.id/index.php/bangkitindonesia/article/view/120</a>   |
|    | Permasalahan      | Twitter merupakan salah satu media sosial yang memungkinkan penggunanya untuk menyampaikan pendapatnya mengenai produk, layanan, dan hal lainnya. Twitter efektif dan efisien karena hanya bisa menulis maksimal 140 karakter.  |
|    | Hasil utama       | Diketahui bahwa <i>unsupervised learning</i> tidak menghasilkan nilai akurasi yang sangat baik untuk analisis sentimen yang dilakukan oleh TextBlob. Pada studi kasus layanan telekomunikasi menggunakan dataset Twitter, namun hasil analisis menggunakan TextBlob hanya menunjukkan nilai akurasi sebesar 58,59%. Optimalisasi dilakukan dengan menambahkan <i>support vector machine</i> yang termasuk dalam kategori supervised learning. Hasil terbaik dari penelitian ini adalah nilai akurasi sebesar 75%. |
|    | Batasan           |   |

| No          | Konten  | Deskripsi  |
|-------------|---|--|
|             |   | Untuk penelusuran lebih lanjut, analisis sentimen ini dapat dilakukan secara real time (berdasarkan data tweet terbaru)  |
| 8           | Judul paper   | Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine  |
|             | Penulis   | Evita Fitri, Yuri Yuliani, Susy Rosyida, Windu Gata  |
|             | Jurnal/Konferensi   | TRANSFORMTIKA, Vol.18, No.1, July 2020, pp. 71 - 80  |
|             | URL   | <a href="https://www.semanticscholar.org/paper/Analisis-Sentimen-Terhadap-Aplikasi-Ruangguru-Naive-Fitri/9961d28fec2f9076107905d8e287653097a9d552">https://www.semanticscholar.org/paper/Analisis-Sentimen-Terhadap-Aplikasi-Ruangguru-Naive-Fitri/9961d28fec2f9076107905d8e287653097a9d552</a>  |
|             | Permasalahan  | Terdapat lebih dari satu aplikasi pembelajaran online yang mudah digunakan saat ini. Dengan banyaknya aplikasi pembelajaran online yang tersedia, pengguna juga bisa lebih selektif menentukan aplikasi mana yang cocok untuk digunakan. Salah satu hal yang mempengaruhi pilihan pengguna terhadap suatu aplikasi adalah dengan memeriksa review pada halaman download aplikasi tersebut. |
| Hasil utama | Penerapan beberapa model algoritma klasifikasi untuk menentukan nilai akurasi terbaik dan menghasilkan nilai AUC untuk setiap model klasifikasi. Nilai ini digunakan bersama dengan kumpulan data yang diambil dan digunakan sepanjang tahap <i>text preprocessing</i> untuk menginformasikan hasilnya pada kisaran akurasinya >94% dan setiap pengujian menggunakan metode <i>cross validation</i> dengan nilai K-fold sebesar 10. |  |

| No          | Konten   | Deskripsi  |
|-------------|--|--|
|             | Batasan  | Untuk penelitian selanjutnya, dilakukan perbandingan dengan metode atau algoritma klasifikasi lain.  |
| 9           | Judul paper  | Twitter Sentiment Analysis as an Evaluation and Service Base On Python Textblob  |
|             | Penulis  | I Gede Susrama Mas Diyasa, Ni Made Ika Marini Mandenni, Mohammad Idham Fachrurrozi, Sunu Ilham Pradika, Kholilul Rachman Nur Manab, Nyoman Rahadi Sasmita  |
|             | Jurnal/Konferensi  | Workshop on Environmental Science, Society, and Technology (WESTECH 2020) IOP Conf. Series: Materials Science and Engineering  |
|             | URL  | <a href="https://iopscience.iop.org/article/10.1088/1757-899X/1125/1/012034">https://iopscience.iop.org/article/10.1088/1757-899X/1125/1/012034</a>  |
|             | Permasalahan   | Sejauh ini, TELKOM belum melakukan analisis mendalam untuk melihat bagaimana pendapat para pengguna twitter tersebut terhadap produk dan layanannya. Sayang sekali karena pendapat mereka terkait dengan brand TELKOM sehingga diperlukan analisis yang mendalam untuk mencari informasi yang mungkin berguna bagi perusahaan. |
| Hasil utama | Hasil analisis sentimen diperoleh. Dari 3324 tweet yang diambil, terdapat 34,4% tweet positif, 16,1% tweet negatif, dan 49,6% tweet netral. Dari data yang diolah tersebut, matriks konfusi menghitung nilai presisi klasifikasi saraf sebesar 80%, nilai presisi klasifikasi positif sebesar 79,3%, dan nilai presisi klasifikasi negatif sebesar 86%. Sedangkan untuk nilai akurasi program analisis sentimen dihitung memiliki nilai akurasi sebesar 77,2%. |  |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    | Batasan           | Perlu adanya pembandingan metode analisis sentimen berbasis leksikal yang lain untuk mngeteahui metode mana yang lebih baik.  |
| 10 | Judul paper       | Analisis Sentimen Aplikasi Novel Online Di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)   |
|    | Penulis           | Selva Indah Nurhafida <sup>1</sup> , Falentino Sembiring  |
|    | Jurnal/Konferensi | Jurnal Sains Komputer & Informatika (J-SAKTI) Volume 6 Nomor 1, Maret 2022, pp. 317-327   |
|    | URL               | <a href="http://tunasbangsa.ac.id/ejurnal/index.php/jsakti/article/view/447/424">http://tunasbangsa.ac.id/ejurnal/index.php/jsakti/article/view/447/424</a>   |
|    | Permasalahan      | Aplikasi Wattpad dan Dreame memiliki kelebihan dan kekurangan, dan ada beberapa komentar di situs Google Play store. Ulasan aplikasi di Google Play store dapat digunakan untuk menentukan apakah suatu aplikasi baik atau buruk, serta untuk mencari masalah yang dimiliki orang-orang dengan aplikasi tersebut.   |
|    | Hasil utama       | Akurasi aplikasi Wattpad 88,60% berdasarkan data latih dan data uji menggunakan SVM. Artinya dari 342 data uji, terdapat 303 data yang terdiri dari 230 data positif dan 73 data negatif. Sebanyak 39 data terdiri dari 23 data negatif dan 16 data positif salah klasifikasi. Program Dreame mempunyai akurasi sebesar 87,45%, artinya dari 263 data yang dievaluasi, 230 data (181 data kelas positif dan 49 data kelas negatif) teridentifikasi dengan benar dan 33 data salah (23 data kelas negatif dan 9 data kelas positif) teridentifikasi dengan benar dan diklasifikasikan salah. |

| No | Konten            | Deskripsi  |
|----|-------------------|--|
|    | Batasan           | Disarankan untuk melakukan pelabelan secara otomatis lain sbagai pembanding agar terlihat metode mana yang lebih unggul.   |
| 11 | Judul paper       | Analisis Sentimen Terhadap Sekolah Saat Covid-19 Pada Twitter Menggunakan Metode Lexicon Based   |
|    | Penulis           | Achmad Hari Mulyadi, Sri Lestari   |
|    | Jurnal/Konferensi | Jurnal J-COM (Jurnal Informatika dan Teknologi Komputer) Vol. 03 No. 01 (2022) 17 – 23   |
|    | URL               | <a href="https://www.ejurnalunsam.id/index.php/jicom/article/view/4064/3168">https://www.ejurnalunsam.id/index.php/jicom/article/view/4064/3168</a>  |
|    | Permasalahan      | Pada 30 Maret 2020, Menteri Pendidikan mengusulkan dimulainya kembali kelas tatap muka secara bertahap karena pembelajaran daring dinilai kurang efektif selama pandemi virus corona. Tak heran, langkah tersebut menuai banyak komentar masyarakat di media sosial, khususnya Twitter.  |
|    | Hasil utama       | Analisis sentimen kelas tatap muka di Twitter pada masa pandemi dengan menggunakan metode vader sentiment menghasilkan hasil sentimen sangat positif sebanyak 5,4%, positif 6,04%, sangat negatif 0,54%, negatif 3,9%, netral 88,23%. Jika dilihat dari persentase pendapat pengguna Twitter mengenai kelas tatap muka di masa pandemi virus corona, lebih banyak orang yang memberikan komentar netral. |
|    | Batasan           | Belum adanya <i>confusion matrix</i> sehingga hasil yang didapatkan belum terbukti validitasnya.   |
| 12 | Judul paper       | Penggunaan Algoritma Klasifikasi Terhadap Analisa Sentimen Pemindahan Ibukota Dengan Pelabelan Otomatis  |
|    | Penulis           | Jananto Watori, Riska Aryanti, Agus Junaidi  |

| No | Konten            | Deskripsi  |
|----|-------------------|--|
|    | Jurnal/Konferensi | JURNAL INFORMATIKA, Vol.7 No.1 April 2020, Halaman 85~90   |
|    | URL               | <a href="https://ejournal.bsi.ac.id/ejurnal/index.php/ji/article/view/7528/pdf">https://ejournal.bsi.ac.id/ejurnal/index.php/ji/article/view/7528/pdf</a>  |
|    | Permasalahan      | Internet memungkinkan publik menyebarkan ide dan opini dalam kelompok atau individu di situs berita dan media sosial. Penelitian ini menjelaskan bagaimana melakukan analisis sentimen menggunakan tweet pengguna Twitter mengenai pemindahan ibu kota Indonesia. Berdasarkan banyaknya gagasan dan opini masyarakat di Twitter, bisa menganalisis suasana terkait pemindahan ibu kota Indonesia setidaknya secara global. |
|    | Hasil utama       | Untuk memberi label pada setiap tweet di Twitter, dapat menggunakan Vader untuk membuat sentimen. Selain itu, berdasarkan algoritma yang digunakan, algoritma SVM terbukti memberikan nilai akurasi tertinggi yaitu mencapai 76,40% dengan nilai AUC sebesar 0,771 menjadikannya klasifikasi yang wajar dibandingkan dengan algoritma Naive Bayes.   |
|    | Batasan           | Tidak ada kejelasan terkait jumlah data yang digunakan dalam penelitian.   |
| 13 | Judul paper       | Sentiments Analysis of Covid-19 Vaccine Tweets Using Machine Learning and Vader Lexicon Method   |
|    | Penulis           | Vishakha Arya, Amit Kumar Mishra, Alfonso González-Briones   |
|    | Jurnal/Konferensi | ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal Regular Issue, Vol. 11 N. 4 (2022), 507-518  |
|    | URL               | <a href="https://revistas.usal.es/cinco/index.php/2255-2863/article/view/27349/28712">https://revistas.usal.es/cinco/index.php/2255-2863/article/view/27349/28712</a>  |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    | Permasalahan      | Dalam proses memantau sentimen masyarakat di situs media sosial membantu pemerintah dan lembaga lain untuk memahami perilaku masyarakat terhadap vaksin dan membuat kebijakan yang efektif untuk meluncurkan program imunisasi guna memitigasi pandemi virus corona.  |
|    | Hasil utama       | Model klasifikasi machine learning yang diimplementasikan ke dataset yang telah diproses sebelumnya menggunakan Regresi Logistik dengan akurasi 92%. Random Forest Classifier dengan TF-IDF dengan akurasi 90% dan metode leksikon VADER mencapai akurasi 64%. Regresi logistik menunjukkan hasil yang baik pada tweet vaksinasi. |
|    | Batasan           | Pekerjaan di masa depan, kumpulan data perlu dimasukkan ke dalam model klasifikasi yang berbeda untuk menganalisis pendirian sentimen. Bahasa lain harus digunakan saat menganalisis sentimen dari media sosial.  |
| 14 | Judul paper       | Aspect-based Sentiment Analysis on Car Reviews Using SpaCy Dependency Parsing and VADER   |
|    | Penulis           | Muchamad Taufiq Anwar, Dedy Trisanto, Ahmad Juniar, Fitra Aprilindo Sase  |
|    | Jurnal/Konferensi | Advance Sustainable Science, Engineering and Technology (ASSET) Vol. 5, No.1, April 2023, pp. 0230109-01 ~ 0230109-07   |
|    | URL               | <a href="https://journal.upgris.ac.id/index.php/asset/article/view/14897/pdf">https://journal.upgris.ac.id/index.php/asset/article/view/14897/pdf</a>   |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    | Permasalahan      | Semua pelaku bisnis, termasuk produsen mobil, perlu memahami aspek apa dari produk mereka yang dianggap positif dan negatif berdasarkan ulasan pengguna sehingga mereka dapat melakukan perbaikan pada aspek negatif dan mempertahankan aspek positif dari produk mereka  |
|    | Hasil utama       | Hasil penelitian menunjukkan bahwa sebagian besar temuan berada pada aspek positif: kenyamanan berkendara, penghematan bahan bakar/jarak tempuh yang baik, keandalan, kelapangan, nilai uang, kamera belakang yang membantu, pengendalian yang tenang, akselerasi yang baik, desain yang baik, sistem suara yang bagus, dan bangunan kokoh. Akibat dari aspek negatifnya adalah jarak tempuh yang rendah, kualitas yang buruk, desain yang buruk, sulit dikendarai, dan akselerasi yang sulit. Namun hasil pada aspek negatif bertolak belakang dengan hasil pada aspek positif dan hanya mempunyai frekuensi yang sangat rendah. |
|    | Batasan           | Penelitian di masa depan dapat meningkatkan identifikasi sentimen terkait aspek ketika tidak ada frasa kata benda eksplisit yang ada dalam ulasan. Penelitian di masa depan juga dapat menggabungkan ABSA dengan LDA untuk mengidentifikasi kelompok aspek secara otomatis.   |
| 15 | Judul paper       | Metode Naïve Bayes Classifier Dengan Textblob Untuk Analisis Sentimen Terhadap Pelayanan Indihome dan First Media   |
|    | Penulis           | Navi Atri Lestari, Tubagus Mohammad Akhriza dan Eka Yuniar  |
|    | Jurnal/Konferensi | Seminar Nasional Teknologi Informasi dan Komunikasi STI&K (SeNTIK) Volume 4 Nomor 1, 23 September 2020, ISSN : 2581-2327  |

| No | Konten            | Deskripsi  |
|----|-------------------|--|
|    | URL               | <a href="https://ejournal.jak-stik.ac.id/index.php/sentik/article/view/3106">https://ejournal.jak-stik.ac.id/index.php/sentik/article/view/3106</a>  |
|    | Permasalahan      | Jejaring sosial yang banyak digunakan adalah Twitter, dimana pengguna dapat menulis berbagai ulasan tentang produk, layanan, selebriti, dll. Hal ini memungkinkan perusahaan untuk mengetahui opini dan sentimen masyarakat mengenai produk dan layanannya, yang dapat digunakan sebagai bahan evaluasi untuk meningkatkan layanan dan produknya.                            |
|    | Hasil utama       | Dengan menggunakan model yang dibuat untuk masing-masing vendor untuk mengklasifikasikan sentimen positif dan negatif, Indihome menerima lebih banyak tanggapan negatif (135 tanggapan) dibandingkan tanggapan positif (58 tanggapan); Demikian pula, First Media menerima 54 tanggapan positif dengan 141 tanggapan negatif, skor akurasi 74%, recall 66%, dan presisi 83%. |
|    | Batasan           | Perlu adanya peningkatan jumlah data yang digunakan dalam penelitian.  |
| 16 | Judul paper       | ANALYSIS SENTIMENT ON THE ACCEPTANCE OF CPNS 2021 ON TWITTER SOCIAL MEDIA USING TEXTBLOB   |
|    | Penulis           | Widi Astuti, Elly Firasari, F. Lia Dwi Cahyanti, Fajar Sarasati  |
|    | Jurnal/Konferensi | Techno Nusa Mandiri: Journal of Computing and Information Technology Vol. 19, No. 1 Maret 2022   |
|    | URL               | <a href="https://ejournal.nusamandiri.ac.id/index.php/techno/article/view/2980/918">https://ejournal.nusamandiri.ac.id/index.php/techno/article/view/2980/918</a>  |
|    | Permasalahan      | Untuk mengkaji antusiasme masyarakat Indonesia, penelitian ini menggunakan data review aplikasi Twitter Persetujuan CPNS 2021.   |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    | Hasil utama       | Temuan penelitian ini berdasarkan data media sosial Twitter mengenai penerimaan CPNS tahun 2021 di kalangan WNI. Artinya, 42,5% masyarakat mempunyai sikap positif, 44,4% masyarakat mempunyai sikap netral, dan 13,1% masyarakat mempunyai sikap netral. Kesimpulan tersebut diperkuat dengan akurasi pengujian teknik data mining “regresi logistik” yaitu sebesar 89,70%. Dari hasil tersebut dapat disimpulkan bahwa masyarakat Indonesia cenderung memiliki sikap positif terhadap pemberitaan mengenai penerimaan CPNS. |
|    | Batasan           | Untuk penelitian selanjutnya, dilakukan perbandingan dengan metode atau algoritma analisis sentimen yang lain.  |
| 17 | Judul paper       | Sentiment Analysis of Tweets About Allowing Outdoor Mask Wear Using Naïve Bayes and TextBlob  |
|    | Penulis           | Ilham Firman Ashari, Fadhillah A., M. Daffa, Sekar Ali  |
|    | Jurnal/Konferensi | Indonesian Journal of Computer Science Vol.12, No.3, Ed.2023  |
|    | URL               | <a href="http://3.8.6.95/ijcs/index.php/ijcs/article/view/3238/162">http://3.8.6.95/ijcs/index.php/ijcs/article/view/3238/162</a>   |
|    | Permasalahan      | Disampaikan langsung oleh Presiden Joko Widodo pada 17 Mei 2022, pemerintah memberikan kelonggaran penggunaan masker di ruang terbuka dengan tetap menjaga jarak sosial. Padahal masker berperan penting dalam menghambat penyebaran virus karena sistem pernapasan bagian luar (mulut dan hidung) terhalang dari udara bebas sebagai media penularan. Hal ini menimbulkan pro dan kontra serta   |

| No | Konten            | Deskripsi  |
|----|-------------------|--|
|    |                   | sentimen di masyarakat, dimana masih ada sebagian masyarakat yang masih trauma dengan virus Covid-19).   |
|    | Hasil utama       | Berdasarkan penelitian, mayoritas pengguna twitter memberikan respon netral. Hal ini ditunjukkan dengan jumlah sentimen netral sebesar 75,76% atau sekitar 757 tweet. Data penelitian ini yaitu 1000 data berbahasa Indonesia dengan <i>keywords</i> 'topeng jokowi'. Data pengujian sebesar 20% menghasilkan model yang lebih akurat, yaitu menghasilkan akurasi sekitar 85%, sedangkan model yang menggunakan data pengujian sebesar 30% hanya menghasilkan akurasi sekitar 83%. |
|    | Batasan           | Penelitian selanjutnya diharapkan berfokus pada penambahan dataset agar dapat mengoptimalkan tingkat akurasi.  |
| 18 | Judul paper       | TWITTER SENTIMENT ANALYSIS TOWARDS QATAR AS HOST OF THE 2022 WORLD CUP USING TEXTBLOB  |
|    | Penulis           | Syarafina Dewi, Dede Brahma Arianto  |
|    | Jurnal/Konferensi | Vol. 2 No. 2 (2023): Journal of Social Research  |
|    | URL               | <a href="https://ijsr.internationaljournallabs.com/index.php/ijsr/article/view/615">https://ijsr.internationaljournallabs.com/index.php/ijsr/article/view/615</a>  |
|    | Permasalahan      | Pada tanggal 2 Desember 2010, Qatar terpilih menjadi tuan rumah Piala Dunia FIFA 2022. Terpilihnya Qatar menjadi tuan rumah Piala Dunia 2022 rupanya mendapat reaksi beragam dari berbagai kalangan di dunia.  |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    | Hasil utama       | Penelitian ini menggunakan tiga tahap, tahap pertama sebelum Qatar terpilih menjadi tuan rumah terdapat 88,46% sentimen positif dan 11,54% sentimen negatif, tahap kedua setelah Qatar terpilih menjadi tuan rumah terdapat 79,38% sentimen positif dan 20,62% sentimen negatif, tahap ketiga Sentimen pada saat Piala Dunia 2022 berlangsung di Qatar sebesar 83,72% sentimen positif serta 16,28% sentimen negatif. Berdasarkan penelitian skor akurasi 83% yang berarti model mampu memprediksi secara akurat 83% dari total data pengujian. |
|    | Batasan           | Penelitian selanjutnya diharapkan untuk menggunakan pembandingan metode atau algoritma lain.  |
| 19 | Judul paper       | ANALISIS SENTIMEN MASYARAKAT TERHADAP VAKSINASI COVID-19 PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)   |
|    | Penulis           | Herwinsyah, Arita Witanti   |
|    | Jurnal/Konferensi | Jurnal Sistem Informasi dan Informatika (Simika) Vol 5 No 1 Tahun 2022  |
|    | URL               | <a href="http://ejournal.lppm-unbaja.ac.id/index.php/jsii/article/view/1411/926">http://ejournal.lppm-unbaja.ac.id/index.php/jsii/article/view/1411/926</a>   |
|    | Permasalahan      | Pesatnya penyebaran virus ini ke seluruh dunia, termasuk Indonesia, memerlukan banyak langkah kebijakan dari pemerintah. Pemerintah meyakini vaksinasi virus corona bisa menjadi langkah awal menuju masyarakat yang lebih sehat. Namun, beberapa isu terkait kebijakan pemerintah ini beredar di masyarakat. Banyak juga perbincangan di media sosial mengenai dampak positif dan negatif vaksinasi terhadap tubuh manusia. Bahkan Kementerian Komunikasi dan Informatika menghitung jumlah kejadian   |

| No | Konten            | Deskripsi  |
|----|-------------------|--|
|    |                   | yang muncul di pemberitaan harian dan menemukan jumlah rumor yang beredar di masyarakat berjumlah 1.387.   |
|    | Hasil utama       | Dari hasil penelitian, terdapat 4.708 indeks data tweet yang diolah, dan analisis sentimen menghasilkan 3.937 (83,6%) tanggapan positif dan 771 (16,4%) tanggapan negatif. Hasil akurasi diperoleh dengan membandingkan 80% (3766) data latih dan 20% (942) data uji, memberikan skor akurasi 89%, skor F1 93%, skor presisi 88%, dan skor recall 99%. Hasil pengecekan data indeks Confusion Matrix 942 teridentifikasi 756 true positif, 82 true negative, 99 false positif, dan 5 false negative. |
|    | Batasan           | Dalam penelitian yang dilakukan, masih terdapat kesenjangan pada beberapa proses analisis, seperti belum adanya perpustakaan yang mendukung perhitungan polaritas pada data teks bahasa Indonesia. Oleh karena itu, beberapa kali menemui masalah pengolah kata selama penelitian ini. Harapannya kedepannya tersedia <i>library</i> yang mendukung polaritas dalam proses pelabelan data.   |
| 20 | Judul paper       | ANALISIS SENTIMEN PENGGUNA MEDIA SOSIAL TERHADAP APLIKASI M-HEALTH PEDULI LINDUNGI DENGAN METODE LEXICON BASED DAN NAÏVE BAYES   |
|    | Penulis           | Riky Iskandar Syah, Hoiriyah, Miftahul Wali  |
|    | Jurnal/Konferensi | Indonesian Journal of Business Intelligence Volume 6   Issue 1   June (2023)   |
|    | URL               | <a href="https://ejournal.almaata.ac.id/index.php/IJUBI/article/view/3275/1975">https://ejournal.almaata.ac.id/index.php/IJUBI/article/view/3275/1975</a>  |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    | Permasalahan      | Aplikasi Peduli Lindungi dirilis guna mengatasi Covid-19 dan menggalakkan vaksinasi. Meskipun pelanggaran data menurunkan kepercayaan publik, masih banyak yang menggunakannya dan menganjurkan peningkatan keamanan. Namun sebagian masyarakat merasa tidak puas karena sering mengalami kesalahan registrasi dan kesulitan membuka sertifikat vaksinasi.  |
|    | Hasil utama       | Analisis Sentimen Media Sosial menggunakan Lexicon Based dan Naive Bayes yang dipadukan dengan aplikasi Peduli Lindungi pada Youtube, Tiktok, dan Twitter menunjukkan akurasi 91%, presisi 94%, recall 82%, dan f1_scores 86%, untuk kasus Youtube akurasi 90%, presisi 93%, recall 81%, dan f1_scores 84%. Namun model ini tidak memberikan hasil yang baik pada data dari Twitter, dengan akurasi yang relatif rendah dan nilai presisi, recall, dan f1 skor yang buruk, dengan presisi 70%, presisi 23%, dan recall, 33% menunjukkan pencapaian F1-Skor. |
|    | Batasan           | Untuk penelitian selanjutnya, merekomendasikan pengumpulan data secara otomatis dalam jangka waktu yang lebih lama dan dengan jumlah data yang lebih besar. Selain itu, merekomendasikan pengujian data menggunakan metode klasifikasi yang berbeda.  |
| 21 | Judul paper       | Unmasking the Sentiments of Labuan Bajo: An Instagram-based Analysis for Tourism Insights through VADER Sentiment Analysis  |
|    | Penulis           | Johan Setiawan, Vegeterrikin Gousander, Iwan Prasetiawan  |
|    | Jurnal/Konferensi | G-Tech : Jurnal Teknologi Terapan Volume 7, No. 3, Juli 2023, hal. 967-976  |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    | URL               | <a href="https://ejournal.uniramalang.ac.id/index.php/g-tech/article/view/2615/1827">https://ejournal.uniramalang.ac.id/index.php/g-tech/article/view/2615/1827</a>   |
|    | Permasalahan      | Banyaknya informasi yang beredar di internet menimbulkan tantangan dalam menyaring informasi yang diperlukan, dan membaca seluruh ulasan yang tersedia secara online dapat memakan waktu lama. Jika jumlah resensi yang dibaca terbatas, maka bisa timbul keraguan atau bias.   |
|    | Hasil utama       | Hasil eksperimen menunjukkan akurasi yang mengesankan sebesar 72% saat menggunakan metode sentimen leksikal VADER. Berdasarkan hasil riset, destinasi wisata di Labuan Bajo cenderung menciptakan suasana positif, dengan 58,55% dari 3.351 data yang terkumpul tergolong positif sehingga totalnya menjadi 1.962 kali. Destinasi populer yang sering dikunjungi wisatawan antara lain Pulau Komodo, Pulau Padar, Pantai Pink, Pulau Kelor, Pulau Rinca, Pulau Kanawa, dan Desa Werebo. |
|    | Batasan           | Penelitian selanjutnya memerlukan penerapan machine learning, untuk mengidentifikasi pola secara efektif dan membedakan antara konten yang relevan dan tidak relevan.   |
| 22 | Judul paper       | Analisis Sentimen Berbasis Aspek Ulasan Aplikasi Mobile JKN dengan Lexicon Based dan Naïve Bayes  |
|    | Penulis           | Salsabila Roiqoh, Badrus Zaman, Kartono   |
|    | Jurnal/Konferensi | JURNAL MEDIA INFORMATIKA BUDIDARMA Volume 7, Nomor 3, Juli 2023, Page 1582-159  |
|    | URL               | <a href="http://www.ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/6194/3614">http://www.ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/6194/3614</a>   |

| No | Konten       | Deskripsi   |
|----|--------------|---|
|    | Permasalahan | Pada tahun 2017, BPJS Kesehatan resmi merilis aplikasi Mobile JKN yang mendukung berbagai aktivitas pelayanan medis dengan menggunakan perangkat apapun. Mobile JKN harus mampu memberikan pelayanan terbaik kepada masyarakat. Namun aplikasi mobile JKN belum sepenuhnya diterima masyarakat dan masih memiliki banyak kekurangan.  |
|    | Hasil utama  | Berdasarkan hasil analisis sentimen, metode Naive Bayes mengungguli Lexicon Based (Inset Lexicon). Berdasarkan evaluasi, Naive Bayes dengan 70% data pelatihan dan 30% data pengujian memiliki performa 30% lebih baik dibandingkan data pelatihan lainnya untuk menguji rasio data. Naive Bayes dengan 70% data latih dan 30% data uji memperoleh skor akurasi tertinggi sebesar 94,75%, rata-rata presisi sebesar 64,32%, rata-rata recall sebesar 70,3%, dan rata-rata skor f1 sebesar 64,55%. Sedangkan berbasis leksikon dengan leksikon inset memperoleh akurasi sebesar 59,99%, rata-rata presisi sebesar 45,96%, rata-rata recall sebesar 52,16%, dan skor f1 sebesar 43,36%. |
|    | Batasan      | Berdasarkan kajian tersebut, penelitian dan pengembangan dapat dilakukan dengan menggunakan lebih banyak data review dan teknik pemodelan topik dengan pendekatan lain seperti Latent Semantic Analysis (LSA) dan Non-Negative Matrix Factorization (NMF). Selain itu, keakuratan analisis sentimen menggunakan kosakata yang digunakan dalam penelitian ini kurang tinggi. Oleh karena itu, jika implementasi memerlukan terjemahan kata ke dalam bahasa Inggris, dapat mengembangkannya menggunakan kamus kosakata lain seperti VADER yang memiliki 7500 fitur kosakata.  |
| 23 | Judul paper  | Support Vector Machine untuk Sentiment Analysis Bakal Calon Presiden Republik Indonesia 2024  |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    | Penulis           | Boby Pranata, Susanti, Susi Erlinda, Hadi Asnal   |
|    | Jurnal/Konferensi | Indonesian Journal of Computer Science ISSN 2302-4364(print) dan 2549-7286(online)  |
|    | URL               | <a href="http://3.8.6.95/ijcs/index.php/ijcs/article/view/3231/180">http://3.8.6.95/ijcs/index.php/ijcs/article/view/3231/180</a>   |
|    | Permasalahan      | Salah satu media sosial yang digunakan untuk menyampaikan pendapat adalah Twitter. Menelaah perasaan terhadap seorang calon presiden dapat memberi tahu bagaimana reaksi masyarakat terhadap orang tersebut, sentimen apa yang cenderung dimiliki orang tersebut, dan kata-kata apa yang memicu perasaan tersebut. Pengguna Twitter mengekspresikan berbagai emosi, dan ada banyak opini serta pesan di Twitter.                              |
|    | Hasil utama       | <i>Labeling</i> menggunakan kamus VADER menghasilkan 57,8% sentimen positif atau 1.022 ulasan, 16,6% sentimen netral atau 295 ulasan, dan 25,4% sentimen negatif atau 450 ulasan. Hasil pengujian menggunakan SVM dan perbandingan 80% pelatihan dan 20% pengujian menunjukkan akurasi sebesar 60%. Perbandingan 70% pelatihan dan 30% pengujian, akurasinya adalah 59%. Perbandingan 60% pelatihan dan 40% pengujian, akurasinya adalah 58%. |
|    | Batasan           | Penelitian kedepannya diharapkan berfokus pada penambahan dataset agar dapat mengoptimalkan tingkat akurasi.  |
| 24 | Judul paper       | Vader Lexicon and Support Vector Machine Algorithm to Detect Customer Sentiment Orientation   |
|    | Penulis           | Vivine Nurcahyawati, Zuriani Mustaffa   |

| No | Konten            | Deskripsi   |
|----|-------------------|---|
|    | Jurnal/Konferensi | Journal of Information Systems Engineering and Business Intelligence Vol.9, No.1, April 2023  |
|    | URL               | <a href="https://e-journal.unair.ac.id/JISEBI/article/download/43377/24459">https://e-journal.unair.ac.id/JISEBI/article/download/43377/24459</a>   |
|    | Permasalahan      | Ada beberapa indikator orientasi pelanggan, salah satunya adalah opini atau selera pelanggan, yang memberikan masukan yang berharga bagi bisnis. Dengan pesatnya perkembangan media sosial, pelanggan dapat mengungkapkan emosi, pikiran, dan pendapatnya mengenai suatu jasa atau produk yang mungkin tidak mudah tersampaikan di dunia nyata.   |
|    | Hasil utama       | Temuan menunjukkan bahwa penggunaan Vader dalam proses anotasi menghasilkan nilai akurasi yang lebih unggul dibandingkan dengan anotasi manual. Secara khusus, nilai akurasi meningkat dari 86% menjadi 88,57%, menunjukkan bahwa Vader dapat menjadi alat yang andal untuk membuat anotasi teks. Oleh karena itu, penelitian selanjutnya perlu pertimbangan penggunaan Vader sebagai pengklasifikasi atau mengintegrasikannya ke dalam proses anotasi untuk lebih meningkatkan kinerjanya. |
|    | Batasan           | Diharapkan penelitian ini dapat dilanjutkan dengan menggunakan metode yang lain untuk dapat membandingkan hasil yang didapatkan.  |
| 25 | Judul paper       | Analisis Sentimen Ciutan Twitter Terkait Penerapan Permendikbudristek Nomor 30 Tahun 2021 Menggunakan TextBlob dan Support Vector Machine   |
|    | Penulis           | Yesy Afrillia, Lidya Rosnita, Deassy Siska, Muzaffar Rigayatsyah, Nurqamarina   |
|    | Jurnal/Konferensi | G-Tech Jurnal Teknologi Terapan Volume 6, No. 2, Oktober2022, hal. 387-394  |

| No | Konten       | Deskripsi  |
|----|--------------|--|
|    | URL          | <a href="https://ejournal.uniramalang.ac.id/index.php/g-tech/article/view/1778/1158">https://ejournal.uniramalang.ac.id/index.php/g-tech/article/view/1778/1158</a>  |
|    | Permasalahan | Di era digital, seluruh aspek implementasi kebijakan pemerintah menjadi topik yang banyak dibicarakan di jejaring sosial, khususnya Twitter. Salah satu topiknya terkait implementasi Peraturan Menteri Pendidikan dan Kebudayaan Nomor 30 Tahun 2021. |
|    | Hasil utama  | Berdasarkan hasil pengujian dengan Textblob yang didukung penggunaan algoritma <i>support vector machine</i> , menggunakan 48 data pengujian yang terdiri dari 34 data positif dan 14 data negatif, dan mencapai akurasi sebesar 70,8%.                |
|    | Batasan      | Penelitian ini dapat dilanjutkan dengan topik dan metode lain untuk membandingkan hasil akurasi sistem dari masing-masing metode.  |

Berdasarkan Tabel 2.4 terdapat persamaan dan perbedaan latar belakang masalah dan hasil penelitian yang dilakukan. Perbedaan latar belakang masalah, algoritma dan metode yang digunakan, ruang lingkup, dan lain-lain yang dapat mempengaruhi hasil penelitian. Atas dasar tersebut, penelitian yang dilakukan juga mempunyai faktor pembeda yang dapat mempengaruhi hasil penelitian, seperti perbedaan ruang lingkup, jumlah data dan beberapa indikator variabel yang digunakan dalam penelitian.

Penelitian terkait dengan judul “Analisis Sentimen Pengguna Aplikasi Dana Berdasarkan Ulasan Pada Google Play Menggunakan Metode *Support Vector Machine*” yang dilakukan oleh Abitdavy Athallah Muhammad, Ermatita, dan Desta Sandya

Prasvita. Masalah yang ditemukan yaitu DANA merupakan salah satu *financial technology* (Fintech) yang menyediakan fitur transaksi melalui *smartphone* pengguna. Banyaknya fitur yang disediakan DANA, kemungkinan besar mendapatkan masukan dari pengguna DANA dalam ulasan di Google Play Store. Contohnya jika fitur isi ulang rusak atau gagal, tidak menerima pengembalian uang atau kesulitan menghubungi layanan pelanggan. Hasil penelitian menunjukkan Hasil evaluasi menunjukkan model SVM dengan akurasi 87,58%, presisi 91,20%, dan recall 90,21% sedangkan chi-square SVM dengan akurasi 89,41%, presisi 93,29%, dan *recall* 90,76%. Penelitian ini memberikan saran pelabelan data yang dilakukan hanya menggunakan metode manual sehingga tidak ada pembandingan dengan data yang dilakukan pelabelan secara otomatis. (Athallah Muhammad et al., 2022) .

Penelitian terkait selanjutnya dengan judul “Analisis Sentimen Aplikasi Novel Online Di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)”. Masalah yang ditemukan yaitu Aplikasi Wattpad dan Dreame memiliki kelebihan dan kekurangan, dan ada beberapa komentar di situs Google Play store. Ulasan aplikasi di Google Play store dapat digunakan untuk menentukan apakah suatu aplikasi baik atau buruk, serta untuk mencari masalah yang dimiliki orang-orang dengan aplikasi tersebut. Hasil penelitian menunjukkan Akurasi aplikasi Wattpad 88,60% berdasarkan data latih dan data uji menggunakan SVM. Artinya dari 342 data uji, terdapat 303 data yang terdiri dari 230 data positif dan 73 data negatif. Sebanyak 39 data terdiri dari 23 data negatif dan 16 data positif salah klasifikasi. Program Dreame mempunyai akurasi sebesar 87,45%, artinya dari 263 data yang

dievaluasi, 230 data (181 data kelas positif dan 49 data kelas negatif) teridentifikasi dengan benar dan 33 data salah (23 data kelas negatif dan 9 data kelas positif) teridentifikasi dengan benar dan diklasifikasikan salah. Disarankan melakukan pelabelan secara otomatis lain sbagai pembanding agar terlihat metode mana yang lebih unggul (Nurhafida and Sembiring, 2022).

Penelitian terkait selanjutnya dengan judul “Analisis Sentimen Komentar YouTube TvOne Tentang Ustadz Abdul Somad Dideportasi Dari Singapura Menggunakan Algoritma SVM”. Masalah yang ditemukan yaitu channel YouTube Berita TvOne mendapat banyak respon dari publik sehingga menjadikan saluran ini sebagai salah satu saluran terpopuler di Indonesia dengan antusias yang tinggi serta mendapat banyak respon dari masyarakat. Hasil penelitian menunjukkan Algoritma SVM digunakan untuk mengklasifikasikan komentar video YouTube di channel TvOne tentang deportasi Ustad Abdul Somad Singapura dengan akurasi 95,02%, recall 95,02%, presisi 95,18%, dan F1-score 95,01%. Penerapan SVM pada penelitian ini memiliki akurasi yang tinggi dan mampu memberikan solusi terhadap permasalahan analisis sentimen komentar video YouTube. Disarankan sebagai pembanding performa yang didapatkan maka diperlukan penambahan algoritma lain untuk melakukan klasifikasi (Mualfah, Gunawan and Suratno, 2023).

### **2.2.2 Matriks Penelitian**

Tabel 2.5 merupakan matriks penelitian berisi judul dan metode yang digunakan pada penelitian sebelumnya.

Tabel 2.5 Matriks Penelitian

| No | Judul  | Metode / Algoritma |               |     |             |       |          |       |
|----|--|--------------------|---------------|-----|-------------|-------|----------|-------|
|    |  | SVM                | Random Forest | KNN | Naïve Bayes | Vader | Textblob | Inset |
| 1  | Analisis Sentimen Komentar YouTube TvOne Tentang Ustadz Abdul Somad Dideportasi Dari Singapura Menggunakan Algoritma SVM | √                  |               |     |             |       | √        |       |
| 2  | Analisis Sentimen Masyarakat Indonesia Terhadap Produk Kendaraan Listrik Menggunakan VADER                               |                    |               |     |             | √     |          |       |
| 3  | Pelabelan Otomatis Lexicon Vader dan Klasifikasi Naive Bayes dalam Menganalisis Sentimen Data Ulasan PLN Mobile          |                    |               |     | √           | √     |          |       |
| 4  | Analisis Sentimen Terhadap Cryptocurrency Berbasis Python TextBlob Menggunakan Algoritma Naïve Bayes                     |                    |               |     | √           |       | √        |       |

| No | Judul   | Metode / Algoritma |               |     |             |       |          |       |
|----|---|--------------------|---------------|-----|-------------|-------|----------|-------|
|    |   | SVM                | Random Forest | KNN | Naïve Bayes | Vader | Textblob | Inset |
| 5  | Analisis Sentimen Pengguna Aplikasi Dana Berdasarkan Ulasan Pada Google Play Menggunakan Metode Support Vector Machine    | √                  |               |     |             |       |          |       |
| 6  | Analisis Sentimen Mengenai Vaksin Sinovac Menggunakan Algoritma Support Vector Machine (SVM) Dan K-Nearest Neighbor (KNN) | √                  |               | √   |             |       | √        |       |
| 7  | Optimasi TextBlob Menggunakan Support Vector Machine untuk Analisis Sentimen (Studi Kasus Layanan Telkomsel)              | √                  |               |     |             |       | √        |       |
| 8  | Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine | √                  | √             |     | √           |       |          |       |

| No | Judul   | Metode / Algoritma |               |     |             |       |          |       |
|----|---|--------------------|---------------|-----|-------------|-------|----------|-------|
|    |   | SVM                | Random Forest | KNN | Naïve Bayes | Vader | Textblob | Inset |
| 9  | Twitter Sentiment Analysis as an Evaluation and Service Base On Python Textblob                                 |                    |               |     |             |       | √        |       |
| 10 | Analisis Sentimen Aplikasi Novel Online Di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM) | √                  |               |     |             | √     |          |       |
| 11 | Analisis Sentimen Terhadap Sekolah Saat Covid-19 Pada Twitter Menggunakan Metode Lexicon Based                  |                    |               |     |             | √     | √        |       |
| 12 | Penggunaan Algoritma Klasifikasi Terhadap Analisa Sentimen Pemindahan Ibukota Dengan Pelabelan Otomatis         | √                  |               |     | √           | √     |          |       |

| No | Judul   | Metode / Algoritma |               |     |             |       |          |       |
|----|---|--------------------|---------------|-----|-------------|-------|----------|-------|
|    |   | SVM                | Random Forest | KNN | Naïve Bayes | Vader | Textblob | Inset |
| 13 | Sentiments Analysis of Covid-19 Vaccine Tweets Using Machine Learning and Vader Lexicon Method                    |                    | √             |     |             | √     |          |       |
| 14 | Aspect-based Sentiment Analysis on Car Reviews Using SpaCy Dependency Parsing and VADER                           |                    |               |     |             | √     |          |       |
| 15 | Metode Naïve Bayes Classifier Dengan Textblob Untuk Analisis Sentimen Terhadap Pelayanan Indihome dan First Media |                    |               |     | √           |       | √        |       |
| 16 | ANALYSIS SENTIMENT ON THE ACCEPTANCE OF CPNS 2021 ON TWITTER SOCIAL MEDIA USING TEXTBLOB                          |                    |               |     |             |       | √        |       |

| No | Judul   | Metode / Algoritma |               |     |             |       |          |       |
|----|---|--------------------|---------------|-----|-------------|-------|----------|-------|
|    |   | SVM                | Random Forest | KNN | Naïve Bayes | Vader | Textblob | Inset |
| 17 | Sentiment Analysis of Tweets About Allowing Outdoor Mask Wear Using Naïve Bayes and TextBlob  |                    |               |     | √           |       | √        |       |
| 18 | TWITTER SENTIMENT ANALYSIS TOWARDS QATAR AS HOST OF THE 2022 WORLD CUP USING TEXTBLOB   |                    |               |     |             |       | √        |       |
| 19 | ANALISIS SENTIMEN MASYARAKAT TERHADAP VAKSINASI COVID-19 PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) | √                  |               |     |             |       | √        |       |
| 20 | ANALISIS SENTIMEN PENGGUNA MEDIA SOSIAL TERHADAP  |                    |               |     | √           | √     |          |       |

| No | Judul  | Metode / Algoritma |               |     |             |       |          |       |
|----|--|--------------------|---------------|-----|-------------|-------|----------|-------|
|    |  | SVM                | Random Forest | KNN | Naïve Bayes | Vader | Textblob | Inset |
|    | APLIKASI M-HEALTH PEDULI LINDUNGI DENGAN METODE LEXICON BASED DAN NAÏVE BAYES  |                    |               |     |             |       |          |       |
| 21 | Unmasking the Sentiments of Labuan Bajo: An Instagram-based Analysis for Tourism Insights through VADER Sentiment Analysis |                    |               |     |             | √     |          |       |
| 22 | Analisis Sentimen Berbasis Aspek Ulasan Aplikasi Mobile JKN dengan Lexicon Based dan Naïve Bayes                           |                    |               |     | √           |       |          | √     |
| 23 | Support Vector Machine untuk Sentiment Analysis Bakal Calon Presiden Republik Indonesia 2024                               | √                  |               |     |             | √     |          |       |

| No | Judul   | Metode / Algoritma |               |     |             |       |          |       |
|----|---|--------------------|---------------|-----|-------------|-------|----------|-------|
|    |   | SVM                | Random Forest | KNN | Naïve Bayes | Vader | Textblob | Inset |
| 24 | Vader Lexicon and Support Vector Machine Algorithm to Detect Customer Sentiment Orientation   | √                  |               |     |             | √     |          |       |
| 25 | Analisis Sentimen Ciutan Twitter Terkait Penerapan Permendikbudristek Nomor 30 Tahun 2021 Menggunakan TextBlob dan Support Vector Machine | √                  |               |     |             |       | √        |       |

### 2.2.3 Relevansi Penelitian

Tabel 2.6 merupakan perbandingan relevansi penelitian sebelumnya dengan penelitian yang dilakukan.

Tabel 2.6 Relevansi Penelitian

| Peneliti | (Athallah Muhammad et al., 2022)  | (Rahmi Intari, 2023)  |
|----------|---|---|
| Judul    | Analisis Sentimen Pengguna Aplikasi DANA Berdasarkan Ulasan Pada Google Play Menggunakan Metode <i>Support Vector Machine</i> | Komparasi Metode Pelabelan Otomatis <i>Textblob</i> dan <i>Vader</i> Pada Analisis Sentimen Ulasan Aplikasi Dana dengan <i>Support Vector Machine</i> |

| <b>Peneliti</b>           | (Athallah Muhammad et al., 2022)   | (Rahmi Intari, 2023)  |
|---------------------------|--|---|
| <b>Masalah Penelitian</b> | Banyaknya fitur yang disediakan DANA, kemungkinan besar mendapatkan masukan dari pengguna DANA dalam ulasan di Google Play Store. Contohnya jika fitur isi ulang rusak atau gagal, tidak menerima pengembalian uang atau kesulitan menghubungi layanan pelanggan | Ulasan aplikasi DANA harus dipertimbangkan karena ada banyak aplikasi dompet digital lain yang memiliki total <i>download</i> dan <i>rating</i> hampir sama dengan aplikasi DANA, sehingga tidak relevan dengan penilaian terhadap aplikasi. Selain itu, rating dan ulasan tersebut tidak dapat dijadikan acuan untuk menilai aplikasi karena ketidaksesuaian rating dan ulasan yang diberikan. |
| <b>Objek Penelitian</b>   | Aplikasi DANA  | Aplikasi DANA   |
| <b>Metode</b>             | <i>SVM, Rating-based</i>   | <i>SVM, Textblob, Vader</i>   |
| <b>Dataset</b>            | Dataset yang digunakan sebanyak 1.366 ulasan aplikasi DANA terbaru berbahasa indonesia.  | Dataset yang digunakan sebanyak sekitar 10.683 ulasan aplikasi DANA terbaru.  |

Berdasarkan Tabel 2.6 terdapat penelitian yang mempunyai keterkaitan dengan penelitian yang dilakukan. Penelitian yang dilakukan mempunyai latar belakang masalah serta objek penelitian yang relevan, namun perlu adanya pengembangan dari penelitian sebelumnya yaitu dengan membuat perubahan pada metode pelabelan dengan menggunakan metode pelabelan *lexicon-*

*based* serta membandingkan perfroma kedua metode pelabelan *lexicon-based* tersebut pada algoritma *Support Vector Machine*.

Selain itu, adanya penambahan jumlah dataset yang digunakan menjadi sekitar 10.683 ulasan dari Google Play Store