

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen merupakan bagian dari studi *text mining* yang dapat disebut sebagai studi komputasi untuk mengklasifikasikan pendapat, emosi dan sikap seseorang terhadap suatu entitas (Hudaya, Fakhrurroja and Alamsyah, 2019).

Analisis sentimen seringkali dianggap setara dengan *opinion mining* karena metode ini berfokus pada klasifikasi pendapat yang bersifat positif, netral atau negatif (Samsir *et al.*, 2021). Terdapat beberapa kelas dalam analisis sentimen, yaitu diantaranya :

1. Sentimen Positif : Menurut Kamus Besar Bahasa Indonesia atau disingkat menjadi KBBI, sentimen positif merupakan sebuah reaksi atau sikap yang meningkatkan nilai seseorang atau terhadap sesuatu .
2. Sentimen Negatif : Menurut Kamus Besar Bahasa Indonesia atau disingkat menjadi KBBI, sentimen negatif merupakan sebuah reaksi atau sikap yang bisa menurunkan nilai seseorang atau terhadap sesuatu hingga membuat tren menjadi down. Pada umumnya kalimat yang bersentimen negatif ditandai dengan penggunaan kata yang negasi. Biasanya negasi ini digunakan untuk mengubah pola dari suatu pernyataan (Ardiani, Sujaini and Tursina, 2020).
3. Sentimen Netral : Menurut Kamus Besar Bahasa Indonesia atau disingkat menjadi KBBI, sentimen negatif merupakan sebuah reaksi yang tidak

berpihak. Biasanya kalimat sentimen netral ini tidak memiliki ekspresi kalimat yang tidak bersifat positif maupun kalimat negatif (Ardiani, Sujaini and Tursina, 2020).

2.2 Kepuasan Pengguna

Kepuasan adalah perbandingan antara ekspektasi (harapan) dengan realita (kenyataan). Semakin sejalan harapan dengan kenyataan maka tingkat kepuasan akan ikut meningkat dan begitu pun sebaliknya. Kepuasan pengguna merupakan hal yang penting dalam pengembangan sebuah sistem informasi. Tingkat kepuasan pengguna suatu sistem dapat dijadikan acuan dalam proses pengembangan sistem itu sendiri, dan untuk mengetahui kelebihan dan kekurangan dari sistem yang sedang berjalan atau diterapkan (Saputra and Kurniadi, 2019).

2.3 Text Mining

Text mining merupakan sebuah bidang yang mencakup pencarian informasi, pertambahan data, statistik, pembelajaran mesin dan komputasi linguistik. *Text mining* juga digunakan untuk menangani klasifikasi, pengelompokan dan ekstraksi informasi. Metode ini bertujuan menghasilkan informasi dari dokumen yang tidak terstruktur melalui tahapan pemrosesan, pengelompokan, dan analisis data (Hudaya, Fakhurroja and Alamsyah, 2019). *Text mining* ini memiliki beberapa tahapan, yaitu :

1. *Tokenizing*

Tahapan *Tokenizing* merupakan proses pemotongan string berdasarkan tiap kata yang menyusunnya. Proses ini dilakukan untuk menghasilkan kumpulan kata dengan menghilangkan delimiter seperti titik (.), koma (,), spasi dan karakter angka

yang terdapat dalam kata-kata tersebut. Tujuan dari Tokenizing adalah untuk mengeksplorasi kata-kata dalam suatu kalimat (Ardiani, Sujaini and Tursina, 2020).

2. *Filtering*

Tahapan Filtering dilakukan setelah tahap Tokenizing, dan merupakan proses seleksi data dimana sebagian data yang tidak relevan dibuang (Pratama, Andrean and Nugroho, 2019). Pada tahap ini, kata-kata yang tidak akan digunakan yang disebut sebagai *stopword* dihapus dari *bag of words*. *Stopword* adalah daftar kata yang tidak merepresentasikan isi dokumen teks. Proses ini bertujuan untuk menghilangkan kata-kata yang dianggap tidak memiliki makna signifikan (Ardiani, Sujaini and Tursina, 2020).

3. *Stemming*

Tahap Stemming merupakan proses pemetaan dan penguraian berbagai varian suatu kata menjadi bentuk kata dasarnya (Pratama, Andrean and Nugroho, 2019). Tujuan dari tahap Stemming adalah menghilangkan imbuhan baik itu prefiks, sufiks atau pun konfiks yang terdapat pada setiap kata. Proses ini biasanya menggunakan library yang terdapat pada bahasa pemrograman python, yaitu sastrawi (Ardiani, Sujaini and Tursina, 2020).

4. *Tagging*

Tahapan Filtering ini merupakan sebuah proses yang menyusun sebuah informasi berdasarkan kategorinya (Pratama, Andrean and Nugroho, 2019).

5. *Analyzing*

Tahapan Analyzing ini merupakan proses untuk mencari seberapa jauh hubungan antara kata dari setiap dokumen (Pratama, Andrean and Nugroho, 2019).

2.4 Naïve Bayes Classifier

Naïve *bayes classifier* merupakan salah satu metode yang banyak digunakan untuk melakukan *sentiment analysis* dengan menggunakan teknik prediksi yang berbasis probalistik berdasar teorema *Bayes*. Metode ini melakukan pengklasifikasian menggunakan metode pendekatan *supervised learning* berarti memiliki label sebagai acuan dan dibangun oleh data latih. Tujuan dari algoritma NBC untuk memperkirakan probabilitas berdasarkan kategori yang ada pada data latih. Klasifikasi ini menggabungkan pengetahuan sebelumnya dengan pengetahuan baru (Gunawan, Fauzi and Adikara, 2017).

Metode Naïve Bayes ini bisa mengambil keputusan berdasarkan data yang diperoleh sebelumnya. Walaupun independensi pada tiap kata pada dokumen tidak dapat bekerja secara maksimal, tetapi kinerja Naïve Bayes ini dalam melakukan klasifikasi sangat baik. Metode Naïve Bayes ini memiliki keunggulan, yaitu diantaranya sangat mudah untuk diimplementasikan, cepat dalam melakukan klasifikasi data, serta memiliki akurasi yang cukup tinggi (Rahat, Kahir and Masum, 2020). Secara umum rumus Naïve Bayes ini pada dasarnya adalah sebagai berikut (Tuhenay and Mailoa, 2021) :

$$P(C_i|X) = \frac{P(X|C_i) P(C_i)}{P(X)}$$

Keterangan :

$P(C_i|X)$ = Probabilitas C_i dengan bukti X

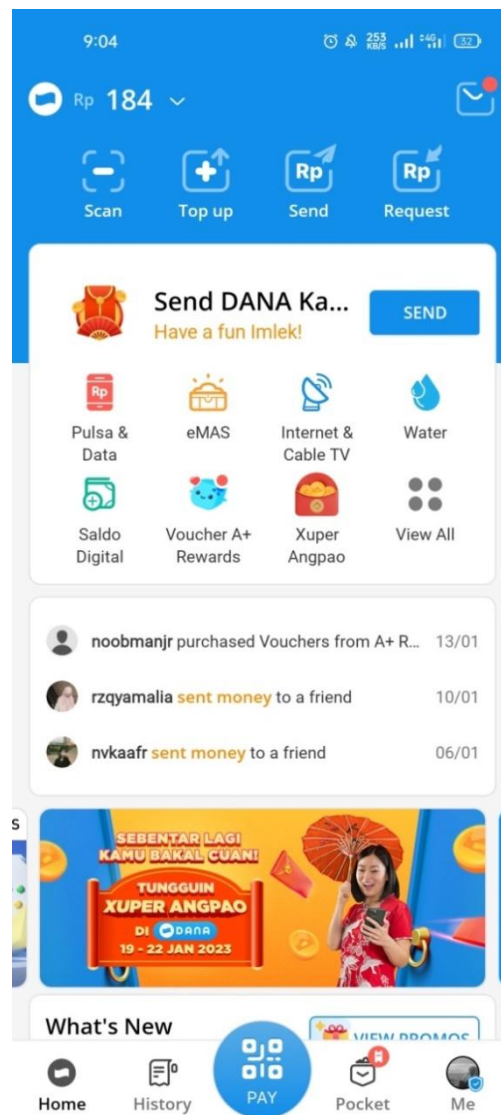
$P(C_i)$ = Probabilitas C_i

$P(X|C_i)$ = Probabilitas X berpengaruh terhadap C_i

$P(X)$ = Probabilitas X

2.5 Digital Payment DANA

Sebuah *startup* bernama DANA atau Dompet Digital Indonesia pada tanggal 21 Maret 2018 secara resmi telah diperkenalkan sebagai salah satu layanan pembayaran digital di Indonesia. Dengan berbasis *open-platform*, DANA diharapkan bisa menjadi solusi bagi kegiatan transaksi digital non-tunai, baik secara *online* maupun *offline*. DANA merupakan dompet digital yang sudah terdaftar di Bank Indonesia dengan memiliki empat lisensi yaitu sebagai uang elektronik, dompet digital, kirim uang, dan Likuiditas Keuangan Digital (LKD). DANA juga sudah mendapat izin untuk beroperasi sebagai perusahaan *fintech* di seluruh Indonesia pada tanggal 5 Desember 2018. DANA dapat digunakan untuk melakukan transaksi mulai dari pembayaran tagihan, *top up game online*, tarik saldo tunai, kirim uang ke sesama pengguna DANA atau ke Bank lain, kredit, dan DANA *PayLater*.



Gambar 2.1 Tampilan Awal Aplikasi DANA Versi 2.27.1

2.6 Dataset

Dataset adalah istilah informal yang merujuk pada kumpulan data dan digunakan untuk klasifikasi dengan metode data mining. Dataset merupakan kumpulan data yang berasal dari informasi-informasi pada masa lalu dan siap untuk dikelola menjadi informasi baru (Yusa *et al.*, 2021). Dataset merupakan

representasi data yang disimpan di memori dalam kondisi tidak terhubung (kumparan, 2021).

2.7 Adaptive Boosting (Adaboost)

Adaptive Boosting (Adaboost) merupakan salah satu dari beberapa varian pada algoritma *boosting*. *Adaboost* merupakan *ensemble learning* yang sering digunakan pada algoritma *boosting* (Qadrini L, Sepperwali A and Aina A, 2021). *Adaboost* merupakan salah satu algoritma klasifikasi yang ditemukan oleh Yoav Freund dan Robert Schapire. Algoritma ini membangun *strong classifier* dengan menggabungkan beberapa *weak classifier*. Algoritma ini juga dapat menyesuaikan diri dengan data dan algoritma *classifier* lainnya, sehingga disebut *adaptive*. Algoritma ini juga dapat memperkecil *error* dari *weak classifier* sehingga dapat menaikkan akurasi dari setiap algoritma pembelajaran, sehingga algoritma ini bernama *boosting* (Abdurrahman, 2022). Langkah-langkah pada algoritma *Adaboost* adalah sebagai berikut (Freund & Schapire, 1999) :

1. Lakukan *input* suatu kumpulan sampel pelatihan $\{(x_1, y_1, \dots, x_m, y_m)\}$ dengan label $y_i \in Y = \{1, \dots, k\}$, algoritma pembelajaran dasar (*base learning algorithm*) dan jumlah iterasi atau perulangan T .
2. Lakukan inisialisasi nilai bobot suatu sampel pelatihan, dengan persamaan berikut :

$$D_1(i) = \frac{1}{m} \text{ untuk } i = 1, \dots, m \quad (2.3)$$
3. Untuk $t = 1, \dots, T$
4. Kemudian latih *base learn* h_t dari sampel pelatihan menggunakan distribusi D_t .

5. Hitung kesalahan dari $h_t : \epsilon_t = Pr_{i \sim D_i} [h_t(x_i) \neq y_i]$, menggunakan persamaan berikut :

$$\epsilon_t = \sum_i 1_{h_t(x_i) \neq y_i} D_t(i) \quad (2.4)$$

Jika $\epsilon_t \geq 1/2$, maka set $T = t - 1$, batalkan perulangan dan langsung menuju *output*.

6. Tetapkan bobot dari h_t , dengan persamaan berikut :

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right) \quad (2.5)$$

7. Lakukan *update* bobot sampel pelatihan, dengan persamaan berikut :

$$D_{t+1}(i) = \frac{D_t(i)}{z_t} \times \begin{cases} e^{-\alpha_t} & \text{if } h_t(x_i) = y_i \\ e^{\alpha_t} & \text{if } h_t(x_i) \neq y_i \end{cases} \quad (2.6)$$

Dimana z_t sebuah faktor normalisasi yang mengaktifkan $D_{t+1}(i)$ menjadi distribusi.

8. *Output*, dengan persamaan berikut :

$$H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right) \quad (2.7)$$

2.8 Evaluasi dan Model Klasifikasi

Mengevaluasi satu atau model yang digunakan dan menetapkan apakah terdapat model yang memenuhi tujuan pada tahap awal. Kemudian menentukan apakah ada permasalahan yang tidak dapat tertangani dengan baik serta mengambil keputusan hasil penelitian.

Evaluasi kinerja model klasifikasi sangat penting dilakukan. Untuk melakukan evaluasi terhadap model klasifikasi berdasarkan perhitungan objek testing mana yang diprediksi benar dan tidak benar. Perhitungan ini ditabulasikan kedalam tabel yang disebut dengan *confusion matrix*. *Confusion Matrix* merupakan

alat pengukuran yang dapat digunakan untuk menghitung kinerja atau tingkat kebenaran suatu klasifikasi. *Confusion Matrix* merupakan dataset yang hanya memiliki dua kelas, kelas yang satu sebagai positif dan kelas yang lain sebagai negatif. *Confusion Matrix* ditunjukkan pada tabel berikut :

Tabel 2.1 *Confusion Matrix*

<i>Confusion Matrix</i>		Nilai Prediksi	
		Positif	Negatif
Nilai Sebenarnya	Positif	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
	Negatif	<i>False Positive (FP)</i>	<i>True Negative (TN)</i>

Keterangan :

- TP (*True Positive*) adalah jumlah *record* positif yang diklasifikasikan sebagai positif.
- FP (*False Positive*) adalah jumlah *record* negatif yang diklasifikasikan sebagai positif.
- FN (*False Negative*) adalah jumlah *record* positif yang diklasifikasikan sebagai negatif.
- TN (*True Negative*) adalah jumlah *record* negatif yang diklasifikasikan sebagai negatif.

Hasil *confusion matrix* digunakan untuk mengukur *accuracy*, *precision*, *recall* dan *F1-Score* untuk menganalisa kinerja algoritma dalam melakukan klasifikasi.

- a. *Accuracy* merupakan rasio prediksi benar (positif dan negatif) dengan keseluruhan data. *Accuracy* dapat dihitung dengan rumus berikut :

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

- b. *Precision* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. *Precision* dapat dihitung dengan rumus berikut :

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

- c. *Recall* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. *Recall* dapat dihitung dengan rumus berikut:

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

- d. *F1-Score* merupakan perbandingan rata-rata *precision* dan *recall* yang dibobotkan. *F1-Score* dapat dihitung dengan rumus berikut :

$$F1-Score = \frac{2 (precision \times recall)}{(precision + recall)}$$

2.9 Penelitian Terkait (*State of The Art*)

Penelitian terkait digunakan untuk mengetahui bagaimana metode dan hasil-hasil penelitian yang dilakukan sebagai tolak ukur peneliti untuk menulis dan menganalisis suatu penelitian dan untuk menambah referensi atau teori. Hasil penelitian terkait disajikan pada tabel 2.2 *State Of The Art*.

Tabel 2.2 *State of The Art*

No	Nama Penulis	Tahun	Judul	Metode	Hasil
1.	Artanti Inez Tanggraeni, Melkior N.N Sitokdana	2022	Analisis Sentimen Aplikasi <i>E-Government</i> Pada <i>Google Play</i> Menggunakan Algoritma <i>Naïve Bayes</i>	<i>Naïve Bayes</i>	Pada penelitian ini penulis melakukan beberapa tahapan yaitu salah satunya dengan pembobotan TF-IDF. Dari aplikasi Sentuh Tanahku mendapatkan ulasan berupa 407 ulasan positif dan 235 ulasan negatif. Hasil dari pengujian analisis tersebut menghasilkan akurasi sebesar 89%, <i>precision</i> sebesar 83%, dan <i>recall</i> sebesar 87%.
2.	Nuraeni Herlinawati, Yuri Yulianti, Siti Faizah, Windu Gata, Saudi	2020	Analisis Sentimen <i>Zoom Cloud Meetings</i> Di <i>Play Store</i> Menggunakan <i>Naïve Bayes</i> dan <i>Support Vector Machine</i>	<i>Naïve Bayes</i> dan <i>Support Vector Machine</i>	Pada penelitian ini dataset yang diperoleh setelah tahap <i>preprocessing</i> adalah 1007 data dengan 546 ulasan positif dan 461 ulasan negatif. Dengan menggunakan 10 <i>fold cross validation</i> diperoleh nilai akurasi dan nilai AUC. Dari algoritma <i>Naïve Bayes</i> diperoleh nilai akurasi sebesar 74,37% dan nilai AUC sebesar 0,659. Sedangkan untuk algoritma SVM memperoleh nilai akurasi sebesar 81,22% dan nilai AUC sebesar 0,886.
3.	Moh Khoirul Insan, Umi Hayati, Odi Nurdiawan	2023	Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna di <i>Google Play</i> Menggunakan	<i>Naïve Bayes Classifier</i>	Pada penelitian ini didapatkan dataset sebanyak 1550 data, dengan sentimen positif sebanyak 1012 data dan sentimen negatif sebanyak 894 data. Maka dari itu didapatkan hasil klasifikasi menggunakan

			Algoritma <i>Naïve Bayes</i>		metode <i>naïve bayes</i> dengan nilai akurasi sebesar 84,52%, nilai <i>precision</i> sebesar 82,51%, dan nilai <i>recall</i> sebesar 87,62%.
4.	Surohman, Sopian Aji, Rousyati, Fanny Fatma Wati	2020	Analisis Sentimen Terhadap <i>Review Fintech</i> Dengan Metode <i>Naïve Bayes Classifier</i> dan <i>K-Nearest Neighbor</i>	<i>Naïve Bayes Classifier</i> dan <i>K-Nearest Neighbor</i>	Pada penelitian ini didapatkan <i>dataset</i> sebanyak 232 data, 116 data <i>review</i> positif dan 116 data <i>review</i> negatif. Akurasi yang dihasilkan dengan Algoritma <i>Naïve Bayes</i> menghasilkan nilai <i>accuracy</i> sebesar 84,76% +/-3,93% dengan rata-rata mikro 84,85%. Kemudian dilakukan pengujian kembali dengan menggunakan Algoritma <i>K-Nearest Neighbors</i> (KNN) yang menghasilkan akurasi sebesar 82,92% +/-4,87% dengan rata-rata mikro 82,96%.
5.	Mochamad Anjasmoros, Fitri Marisa, Tri Istiadi,	2020	Analisis Sentimen Aplikasi Go-Jek Menggunakan Metode SVM dan NBC (Studi Kasus : Komentar Pada <i>Play Store</i>)	<i>Support Vector Machine</i> dan <i>Naïve Bayes Classifier</i>	Pada penelitian ini didapatkan dataset sebanyak 2000 data yang terdiri dari 1000 ulasan positif dan 1000 ulasan negatif. <i>Score accuracy, recall</i> dan <i>precision</i> paling tinggi adalah 0,897 dengan <i>data testing</i> sebanyak 50% dan 50% <i>data training</i> menggunakan metode SVM kernel linear. <i>Score accuracy, recall</i> dan <i>precision</i> yang paling rendah adalah 0,48625 dengan <i>data testing</i> sebanyak 40% dan 60% <i>data training</i> menggunakan metode SVM kernel RBF, polynomial dan sigmoid.

6.	Hartati, Deni Hermawan, M. Akhsanal, Zailani Wahyudi, Angga Ariyanto, Dedi Dwi Saputra	2022	Optimasi Analisis Sentimen Pada Twitter Olshop Tokopedia Menggunakan <i>Text Mining</i> dengan Algoritma <i>Naïve Bayes & Adaboost</i>	<i>Naïve Bayes Classifier, Adaboost</i>	Berdasarkan penelitian yang dilakukan didapatkan 1000 <i>dataset</i> berupa tweet yang berkaitan dengan tokopedia. Maka didapatkan hasil dari proses analisis data tersebut dengan nilai akurasi 94.95%, presisi 90.86%, <i>recall</i> 100,00% dan AUC 0,950.
7.	Ristyani Slamet, Windu Gata, Annisa Novtariany, Khairunisa Hilyati, Febri Ainun Jariyah	2022	Analisis Sentimen Terhadap Penggunaan Artis Korea Selatan Sebagai <i>Brand Ambassador</i> Produk Kecantikan Lokal	<i>Support Vector Machine</i>	Pada penelitian yang dilakukan didapatkan 317 <i>dataset</i> yang terdiri dari 266 komentar positif dan 21 komentar negatif. Maka hasil yang didapatkan dari penelitian tersebut dengan menggunakan SVM dengan nilai akurasi sebesar 83.60%, <i>precision</i> sebesar 83.86% dan <i>recall</i> sebesar 99.62%.
8.	Evita Fitri, Yuri Yuliani, Susy Rosyida, Windu Gata	2020	Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma <i>Naïve Bayes, Random Forest</i> dan <i>Support Vector Machine</i>	<i>Naïve Bayes Classifier, Random Forest dan Support Vector Machine</i>	Pada penelitian yang dilakukan, didapatkan <i>dataset</i> sebanyak 1629 yang didapat melalui <i>review google play store</i> . Hasil nilai akurasi yang diperoleh dari berbagai metode yang digunakan adalah sebesar 97,16% untuk <i>Random Forest</i> serta nilai AUC nya sebesar 0,996. Kemudian sebesar 96,01% dengan nilai AUC nya sebesar 0,543 untuk <i>Support Vector Machine</i> , dan sebesar 94,16% serta nilai AUC sebesar 0,999 untuk <i>Naïve Bayes</i> . Maka dapat disimpulkan bahwa hasil performa dari penelitian tersebut menunjukkan bahwa menggunakan

					<i>Random Forest</i> memiliki nilai akurasi tertinggi daripada algoritma yang lainnya.
9.	Christian Cahyaningtyas, Yessica Nataliani, Indrastanti Ratna Widiarsari	2021	Analisis Sentimen Pada Rating Aplikasi Shopee Menggunakan Metode <i>Decision Tree</i> Berbasis SMOTE	<i>Decision Tree</i> , SMOTE	Pada penelitian yang dilakukan didapatkan dataset hasil dari rating <i>Google Play Store</i> dengan data yang diambil sebanyak 2705 data dengan skor penilaian bintang satu sampai dengan bintang lima. Hasil penelitian dengan menggunakan algoritma <i>Decision Tree</i> dengan SMOTE (<i>Synthetic Minority Oversampling Technique</i>) nilai <i>accuracy</i> -nya menghasilkan 99,91 persen, <i>AUC (Area Under Curve)</i> 0,999, <i>recall</i> 99,88 persen dan nilai <i>precision</i> 99,98 persen. Hasil menggunakan algoritma <i>Decision Tree</i> tanpa SMOTE nilai <i>accuracy</i> -nya menghasilkan 99,89 persen, <i>AUC (Area Under Curve)</i> 0,950, <i>recall</i> 99,88 persen dan nilai <i>precision</i> 99,98 persen.
10.	Muhammad Rezki, Desiana Nur Kholifah, Muhammad Faisal, Priyono, Rachmat Suryadithia	2020	Analisis <i>Review</i> Pengguna <i>Google Meet</i> dan <i>Zoom Cloud Meeting</i> Menggunakan Algoritma <i>Naïve Bayes</i>	<i>Naïve Bayes</i> , PSO, SMOTE	Pada penelitian yang dilakukan didapatkan dataset berupa 406 ulasan dari <i>google play</i> untuk aplikasi <i>google meet</i> dan 308 ulasan untuk <i>zoom cloud meeting</i> . Dengan demikian hasil pengujian tersebut mendapatkan hasil akurasi sebesar 85,76 % yang ditambah dengan Feature SMOTE dan PSO untuk <i>review Zoom Cloud Meeting</i> , sedangkan untuk pengguna <i>Google Meet</i>

					yang ditambah dengan <i>Feature SMOTE</i> dan PSO hanya mampu mendapat tingkat akurasi sebesar 79,33 %
11.	Muhammad Yusuf Hidayatulloh, Anto Sunanto, Armansyah, Muhammad Farrell Afelino Gevin, Dedi Dwi Saputra	2023	Optimasi Sentimen Analisis Informatif dan Tidak Informatif dari Tweet di BMKG Menggunakan Algoritma <i>Naïve Bayes</i> dan Metode Teknik Pengambilan Sampel Minoritas Sintetis	<i>Naïve Bayes</i> , Adaboost, SVM, PSO	Dari penelitian tersebut, <i>dataset</i> diambil pada <i>tweet</i> dalam bahasa Indonesia dengan data yang diambil sebanyak 1000 <i>tweet</i> dari akun <i>twitter @infoBMKG</i> . Hasil pengujian dengan menggunakan algoritma <i>Naïve Bayes</i> didapatkan nilai akurasi sebesar 78.79%, presisi 76.93%, <i>recall</i> 82.06% dan AUC 0.779. Sedangkan pengujian algoritma <i>naïve bayes + adaboost</i> didapatkan nilai akurasi sebesar 76.60%, presisi 79.00%, <i>recall</i> 72.72% dan AUC 0.827. Menggunakan algoritma SVM mendapatkan nilai akurasi sebesar 79.25%, presisi 74.35%, <i>recall</i> 89.83% dan AUC 0.845. SVM + PSO mendapatkan nilai akurasi sebesar 78.52%, presisi 76.72%, <i>recall</i> 82.44% dan AUC 0.835. Maka hasil pengujian dari beberapa pengujian tersebut terbukti bahwa optimasi sentiment analisa informatif dan tidak informatif dari <i>tweet</i> pada <i>twitter BMKG</i> mendapatkan hasil yang baik menggunakan metode SVM.
12.	Fanka Angelina Larasati, Dian Eka	2022	Analisis Sentimen Ulasan Aplikasi DANA dengan	<i>Random Forest</i>	Pada penelitian ini digunakan metode <i>random forest</i> dengan membagi tiga kelas sentimen yaitu positif, negatif maupun

	Ratnawati, Buce Trias Hanggara		Metode <i>Random Forest</i>		netral juga mengambil indikator evaluasi yaitu akurasi, <i>recall</i> , presisi dan <i>f-measure</i> . Pengujian dilakukan berdasarkan jumlah <i>tree</i> dan kedalaman <i>tree</i> terhadap 1354 data dengan membagi data 250 data per kelasnya. Berdasarkan hasil pengujian dan analisis yang sudah dilakukan, dengan perbandingan data latih dan data uji 80% : 20% diperoleh nilai <i>precision</i> 84%, <i>recall</i> 84%, <i>F1-Score</i> 84% dan <i>accuracy</i> sebesar 84% dengan kedalaman <i>tree</i> 65 dan jumlah <i>tree</i> 400.
13.	Wahyu Eko Saputro, Haris Yuana, Wahyu Dwi Puspitasari	2023	Analisis Sentimen Pengguna Dompot Digital Dana Pada Kolom Komentar Google Play Store Dengan Metode Klasifikasi <i>Support Vector Machine</i>	<i>Support Vector Machine</i>	Pada penelitian ini didapatkan sebanyak 35% pengguna aplikasi Dana memiliki sentimen positif sedangkan sebanyak 65% pengguna aplikasi Dana memiliki sentimen negatif berdasarkan pengujian klasifikasi SVM memiliki akurasi sebesar 80%, <i>precision</i> sebesar 84.06% untuk sentimen negatif dan 74.08% untuk sentimen positif, serta <i>recall</i> sebesar 87.02% untuk sentimen negatif, serta <i>recall</i> sebesar 69.21% untuk sentimen positif.
14.	Muhamad Kabirur Rifa, Moch. Hafid Totohendarto, M. Rafi Muttaqin	2023	Analisis Sentimen Pengguna <i>E-Wallet</i> Dana dan Gopay Pada <i>Twitter</i> Menggunakan	<i>Support Vector Machine</i>	Pada penelitian ini, jenis kernel SVM yang digunakan adalah kernel linear. Hasil perhitungan aplikasi Dana memiliki <i>accuracy</i> sebesar 92%, dan nilai presisi sebesar 91%, dengan tingkat <i>recall</i> sebesar

			Metode <i>Support Vector Machine</i> (SVM)		92%. Sedangkan aplikasi Gopay memiliki <i>accuracy</i> sebesar 90%, dan nilai presisi sebesar 91% dengan tingkat <i>recall</i> sebesar 90%. Dari hasil tersebut membuktikan sentiment pengguna <i>twitter</i> mengenai aplikasi Gopay lebih baik daripada aplikasi Dana.
15.	Andriani Nurian, Betha Nurina Sari	2023	Analisis Sentimen Pengguna Aplikasi <i>Google Play</i> Menggunakan <i>Naïve Bayes</i>	<i>Naïve Bayes Classifier</i>	Pada penelitian ini didapatkan dataset yang bersumber dari komentar <i>google play store</i> sebanyak 3000 data. Dengan <i>data testing</i> yang digunakan sebanyak 200 data atau 20% dari 3000 data. Hasil dari penerapan algoritma <i>Naïve Bayes Classifier</i> pada aplikasi DANA menunjukkan bahwa nilai akurasi sebesar 85%, presisi sebesar 79%, <i>recall</i> sebesar 85%, dan <i>f1-score</i> sebesar 80%.
16.	Abitdavy Athallah Muhammad, Ermatita, Desta Sandya Prasvita	2022	Analisis Sentimen Pengguna Aplikasi Dana Berdasarkan Ulasan <i>Google Play</i> Menggunakan Metode <i>Support Vector Machine</i>	<i>Support Vector Machine</i>	Berdasarkan penelitian yang dilakukan, didapatkan data sebanyak 1.366 data yang bersumber dari ulasan <i>google play store</i> , dengan komentar positif sebanyak 883 data dan komentar negatif sebanyak 483 data. Hasil evaluasi menunjukkan performa model klasifikasi SVM tanpa seleksi fitur teroptimal diperoleh dengan kernel rbf, nilai $C = 0.5$, dan nilai $\text{Gamma} = 1$ dengan fitur sebanyak 1357 memperoleh akurasi sebesar 87,58%, presisi sebesar 90,21%, <i>recall</i>

					sebesar 91,20%. Sedangkan hasil evaluasi model klasifikasi SVM dengan seleksi fitur <i>chi square</i> dengan kernel rbf, nilai $C = 1$, dan nilai $\Gamma = 1$ dengan fitur sebanyak 540 memperoleh akurasi sebesar 89,41%, presisi sebesar 93,29%, dan <i>recall</i> sebesar 90,76%.
17.	Mohammad Rizal Ramli, Heni Sulastri, Rianto	2023	Analisis Sentimen Terhadap Opini Mahasiswa Terkait Pembelajaran Daring Dengan Menggunakan Algoritma <i>Naïve Bayes Classifier</i> dan SVM Dengan Adaboost Pada Media Sosial Twitter	<i>Naïve Bayes Classifier, Support Vector Machine, Adaboost</i>	Pada penelitian ini dilakukan optimalisasi boosting dengan menggunakan Adaboost. Tujuan dari penelitian ini adalah membandingkan performa algoritma sebelum dan sesudah menggunakan Adaboost. Hasil analisis sentimen terhadap pembelajaran daring ini didapatkan hasil akurasi tertinggi oleh algoritma <i>Naïve Bayes Classifier</i> ditambah dengan Adaboost sebesar 99.26%, dengan presisi sebesar 99.39% dan <i>recall</i> 99.20%.

2.10 Matriks Penelitian

Penelitian terdekat selanjutnya dibuatkan model matriks ruang lingkup penelitian. Adapun matriks ruang lingkup penelitian dipaparkan dalam Tabel 2.3 :

Tabel 2.3 Matriks Penelitian

No.	Peneliti/Tahun	Sumber Data	Lingkup Penelitian						
			Metode				Optimasi		
			SVM	Naïve Bayes	Decision Tree	Random Forest	SMOTE	PSO	Adaboost
1.	(Tanggraeni and Sitokdana, 2022)	<i>Review Google Play Store</i>	-	√	-	-	-	-	-
2.	(Hartati <i>et al.</i> , 2022)	<i>Tweet Twitter</i>	-	√	-	-	-	-	√
3.	(Rezki <i>et al.</i> , 2020)	<i>Review Google Play Store</i>	-	√	-	-	√	√	-
4.	(Khoirul Insan, Hayati and Nurdiawan, 2023)	<i>Review Google Play Store</i>	-		-	-	-	-	-

5.	(Fitri, 2020)	<i>Review Google Play Store</i>			-		-	-	-
6.	(Cahyaningtyas, Nataliani and Widiasari, 2021)	<i>Review Google Play Store</i>	-	-		-		-	-
7.	(Nurian and Nurina Sari, 2023)	<i>Review Google Play Store</i>	-		-	-	-	-	-
8.	(Charlieta, 2024)	<i>Review Google PlayStore</i>	-	√	-	-	-	-	√

