# BAB II

# LANDASAN TEORI

## 2.1 Sentiment analysis

Analisis sentimen merupakan proses untuk mengekstraksi dan menganalisis sentimen, opini, atau perasaan yang terkandung di dalam teks atau dokumen tertentu. Tujuan dari analisis sentimen adalah untuk memahami pandangan dan opini pengguna terhadap suatu produk, merek, atau topik tertentu. Dalam dunia bisnis, analisis sentimen dapat digunakan untuk mengukur kepuasan pelanggan dan meningkatkan kualitas produk atau layanan.

Menurut (Vinodhini & Chandrasekaran, 2012), Analisis sentimen adalah cabang NLP yang mengukur perasaan orang tentang merek, ide, atau konten dari waktu ke waktu. Membangun sistem yang mengumpulkan dan memvalidasi konten buatan pengguna seperti entri *blog*, komentar, ulasan, dan *tweet* adalah inti dari analisis sentimen, juga dikenal sebagai penambangan opini. Ada beberapa aplikasi untuk analisis sentimen. Dalam pemasaran, misalnya, ini dapat digunakan untuk menilai kemanjuran kampanye iklan atau peluncuran produk baru, menemukan iterasi produk populer, dan mempelajari demografi mana yang menyukai atau membenci fitur tertentu.

Pendapat, emosi, penilaian, penilaian, sikap, dan sikap tentang barang, jasa, organisasi, orang, situasi, peristiwa, subjek, dan atributnya adalah fokus analisis sentimen, cabang studi (Liu, 2012).

## 2.2 Pengertian Bank

Bab 1 dan Pasal 1 Ayat 2 Undang-Undang Republik Indonesia No. 10 Tahun 1998 tanggal 10 November 1998 mengubah Undang-undang No. 7 Tahun 1992 mendefinisikan bank menjelaskan bahwa bank adalah organisasi komersial yang menerima simpanan dari masyarakat umum dan meminjamkan dana kembali ke masyarakat umum dalam bentuk kredit dan produk keuangan lainnya. Tujuan keseluruhannya adalah untuk meningkatkan kondisi kehidupan masyarakat.

### 2.2.1 Fungsi Bank

(Triandaru & Budisantoso, 2006) mengklasifikasikan peran bank dalam perekonomian sebagai berikut:

1. Fungsi bank Sebagai *Agent of Trust*

   Untuk mengumpulkan dan mendistribusikan uang, kepercayaan adalah prinsip utama aktivitas perbankan. Jika ada dasar kepercayaan, orang akan mau menyimpan uangnya di bank. Lingkungan memiliki keyakinan bahwa bank akan memperlakukan uang pelanggannya dengan adil, tidak akan bangkrut, dan akan membiarkan mereka menarik simpanan mereka ketika dijanjikan.

2. Fungsi bank Sebagai *Agent of Development*

   Tidak mungkin membedakan antara kegiatan ekonomi masyarakat di sektor keuangan dengan sektor aktual. Selalu berinteraksi dan mempengaruhi satu sama lain adalah dua sektor. Jika sektor moneter tidak berfungsi dengan baik, maka sektor riil juga tidak dapat berfungsi dengan baik. Agar kegiatan

ekonomi di sektor riil dapat berjalan lancar, kegiatan bank seperti penghimpunan dana dan distribusi menjadi sangat penting.

3. Fungsi Bank sebagai *Agent of service*

   Pelayanan yang diberikan oleh bank memiliki keterkaitan yang erat dengan kegiatan ekonomi masyarakat secara keseluruhan. Solusi tersebut meliputi pembayaran tagihan, penyimpanan barang berharga, bank garansi, dan layanan pengiriman uang.

## 2.3 Branchless Banking

Menurut (Yusharto, 2014), Kantor cabang tidak diharuskan menggunakan layanan keuangan yang dikenal sebagai *branchless banking*. Untuk menjangkau nasabah di pedesaan, bank tidak perlu lagi membuka cabang baru. Untuk meniadakan keharusan bank mempertimbangkan strategi pengembangan cabang untuk melayani lokasi pedesaan, Untuk lebih melayani nasabahnya, Bank Indonesia sangat menganjurkan agar bank bekerja sama dengan agen dan pihak ketiga lainnya. Melalui infrastruktur dan TI yang dimilikinya, akan dilakukan Program *Branchless Banking*. Perbankan dapat menjangkau tempat-tempat yang lebih terpencil dengan menggunakan skema *Branchless Banking*. Selain itu, jika bank ingin bekerjasama dengan koperasi dan lembaga keuangan mikro (LKM) untuk bertindak sebagai agen.

Pemerintah Indonesia telah mengidentifikasi inklusi keuangan sebagai inisiatif strategis, dan salah satu komponen dari program ini adalah gagasan *Branchless Banking*. Untuk itu, Strategi Nasional Keuangan Inklusif (SNKI) memasukkan program tersebut di atas, serta Peraturan Otoritas Jasa Keuangan

(OJK) Nomor 19/POJK.03/2014 tentang jasa keuangan tanpa kantor, yang keduanya diterbitkan pada 22 Juli. 2014 oleh Bank Indonesia. Dengan adanya konsep *Branchless Banking*, diharapkan masyarakat yang sebelumnya sulit mengakses layanan keuangan konvensional dapat memperoleh akses ke layanan keuangan melalui berbagai kanal alternatif yang disediakan, seperti melalui agen-agen atau jaringan mitra.

### 2.4 BRILink

Dengan menggunakan Brilink, BRI dan kliennya dapat bekerja sama untuk menyediakan lebih banyak layanan perbankan satu sama lain. Dengan menggunakan *mobile* app BRILink dan kapabilitas EDC (mini ATM), nasabah BRI dapat menjadi agen dan memberikan layanan transaksi perbankan langsung dan *online* kepada masyarakat umum. Gagasan di balik kemitraan ini adalah agar BRI dan klien agensi membagi pendapatan yang dihasilkan dari upaya bersama mereka. Nasabah BRI dapat melakukan berbagai transaksi keuangan, antara lain isi ulang listrik prabayar, bayar biaya pembiayaan, dan isi ulang pulsa, menggunakan Brilink, setoran uang tunai, penarikan uang tunai, dan berbagai layanan lainnya. Brilink juga merupakan bagian dari program *Branchless Banking* yang didukung oleh pemerintah, di mana layanan perbankan dapat dilakukan tanpa perlu melalui kantor fisik bank. Seorang agen Brilink dapat memperoleh manfaat dari menjadi agen, seperti penghasilan tambahan dan meningkatkan keterampilan dalam berbisnis.

## *2.5* **Text mining**

Ekstraksi intelijen yang dapat ditindak lanjuti dari teks atau dokumen melalui penggunaan metode komputer dikenal sebagai penambangan teks. Tujuan dari *Text mining* adalah untuk mengekstraksi informasi yang berguna dari sejumlah besar teks yang tidak terstruktur, sehingga dapat memberikan wawasan baru atau informasi yang berharga bagi pengguna. teknik *text mining* melibatkan pemrosesan bahasa alami, analisis statistik, dan *machine learning*. Contoh aplikasi *text mining* adalah analisis sentimen, klasifikasi dokumen, pengelompokan dokumen, dan ekstraksi informasi. Dalam era informasi digital yang semakin berkembang, *text mining* menjadi semakin penting untuk membantu organisasi dalam mengelola dan memahami data teks yang melimpah.

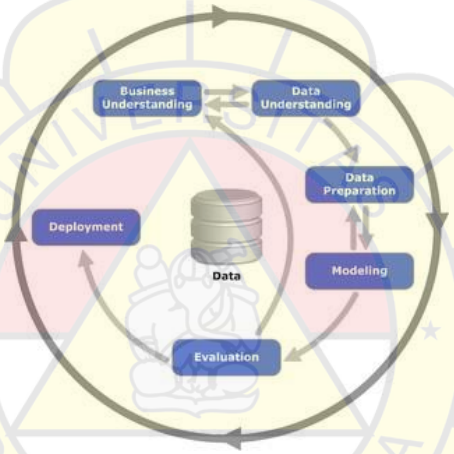
Proses intensif data di mana orang berinteraksi dengan sekumpulan dokumen dari waktu ke waktu menggunakan alat analisis dikenal sebagai *text mining. Text mining,* mirip dengan data *mining*, bertujuan untuk mengekstraksi informasi bermanfaat dari sumber data dengan menemukan dan menyelidiki pola yang menarik. Namun, sumber data *Text mining* adalah sekumpulan dokumen, dan pola yang menarik ditemukan dalam informasi tekstual yang tidak terstruktur dari kumpulan dokumen, bukan dalam kumpulan catatan basis data formal (Feldman & Sanger, 2007).

## 2.6 CRISP-DM

*Cross-Industry Standard Process for Data Mining (CRISP-DM)* adalah standar yang dikembangkan pada tahun 1996 oleh analis dari beberapa industri seperti *Daimler Chrysler, SPSS* dan *NCR. CRISP-DM* juga digunakan sebagai

panduan untuk melakukan analisis data dari berbagai industri sebagai strategi untuk memecahkan masalah dalam bisnis atau unit penelitian (Ncr et al., 2000).

*CRISP-DM (Cross-Industry Standard Process for Data Mining)* adalah sebuah metodologi yang umum digunakan dalam proses data mining dan analisis data. Metodologi ini terdiri dari enam tahap yang saling terkait dan dapat digunakan sebagai panduan dalam melakukan proyek analisis data. Berikut adalah penerapan tahapan CRISP-DM pada penelitian:




Gambar 2. 1 Fase *CRISP-DM*

1. Tahap Pemahaman Bisnis (*Business Understanding*):

   Pada tahap ini, peneliti akan mengidentifikasi tujuan utama dari analisis sentimen terhadap aplikasi Brilink Mobile. Tujuan ini mungkin termasuk meningkatkan kepuasan pengguna, mengidentifikasi masalah aplikasi, atau mendapatkan wawasan strategis untuk pengembangan selanjutnya. Selain itu, peneliti juga akan mengumpulkan persyaratan dan informasi lain yang relevan untuk penelitian.

2. Tahap Pemahaman Data (*Data Understanding*):

Tahap ini berfokus pada eksplorasi dan pemahaman data yang akan digunakan dalam analisis sentimen. Peneliti akan mengumpulkan data ulasan dan komentar pengguna tentang aplikasi Brilink Mobile dari berbagai sumber seperti media sosial, situs ulasan aplikasi, atau platform lainnya. Data ini kemudian akan dieksplorasi untuk memahami struktur, kualitas, dan karakteristiknya.

3. Tahap Persiapan Data (*Data Preparation*):

   Pada tahap ini, data akan dipersiapkan dan dibersihkan untuk analisis selanjutnya. Proses ini meliputi pemilihan data yang relevan, penghilangan data yang tidak lengkap atau tidak valid, penggabungan data dari berbagai sumber, dan transformasi data ke dalam format yang sesuai untuk metode analisis yang akan digunakan.

4. Tahap Pemodelan (*Modelling*):

   Tahap ini melibatkan penggunaan metode *Naïve Bayes* dan *Support Vector Machine (SVM)* untuk membangun model klasifikasi sentimen berdasarkan data ulasan dan komentar pengguna. Model ini akan dilatih dengan data yang telah dipersiapkan sebelumnya untuk memprediksi sentimen *positive*, *negative*, atau netral dari teks ulasan.

5. Tahap Evaluasi (*Evaluation*):

   Tahap evaluasi bertujuan untuk mengevaluasi kinerja model klasifikasi yang telah dibangun. Metrik evaluasi seperti akurasi, presisi, recall, dan F1 score akan digunakan untuk mengukur sejauh mana model dapat mengklasifikasikan sentimen dengan benar.

6. Tahap Penyajian (*Deployment*):

   Tahap akhir adalah penyajian hasil analisis sentimen kepada pemangku kepentingan, seperti Bank Rakyat Indonesia (BRI) atau tim pengembangan aplikasi Brilink Mobile. Hasil ini disajikan dalam bentuk laporan, visualisasi, atau presentasi, yang memberikan wawasan dan rekomendasi berdasarkan temuan dari analisis sentimen.

## 2.7 Preprocessing

*Preprocessing* merupakan langkah yang dilakukan sebelum melakukan proses klasifikasi, yang bertujuan untuk membersihkan, menghapus, dan memodifikasi data sumber yang terdiri dari karakter non-alfabet dan kata-kata yang tidak relevan. Tujuan dari *Preprocessing* adalah untuk memastikan bahwa data yang digunakan dalam proses klasifikasi lebih optimal dan sesuai dengan kebutuhan. Dengan melakukan *Preprocessing*, data dapat disederhanakan, *noise* dapat dihilangkan, dan fitur yang penting dapat ditingkatkan, sehingga memudahkan proses klasifikasi dan meningkatkan kualitas hasil yang diperoleh (F. A. Muttaqin & Bachtiar, 2016). Berikut adalah penjelasan mengenai tahapan *Preprocessing*:

1) *Case folding* Pada tahap ini, dilakukan konversi huruf besar menjadi huruf kecil dan penghapusan seluruh tanda baca dari kalimat.
2) *Tokenizing* Pada tahapan ini, setiap kata akan dipisahkan berdasarkan spasi yang ditemukan.

3) *Stemming* dalam tahapan ini, dilakukan pengubahan kata-kata yang memiliki imbuhan menjadi bentuk dasarnya.
4) *Filtering* pada tahap ini, dilakukan eliminasi kata-kata yang dianggap tidak penting dari hasil tokenisasi.

## 2.8 Pembobotan TF-IDF

TFIDF (*Term Frequency Inverse Document Frequency*) adalah metode pembobotan berupa integrasi antara *term frequency* dan *document inverse frequency* (Widyasanti et al., 2018). TF-IDF terdiri dari dua komponen utama:

1) Frekuensi kata (TF): Metrik ini mengukur seberapa sering sebuah kata muncul dalam dokumen. Semakin sering sebuah kata muncul dalam sebuah dokumen, semakin penting kata itu dalam dokumen tersebut.
2) *Inverse Document Frequency* (IDF): Metrik ini mengukur seberapa umum atau jarang sebuah kata di seluruh dokumen. Semakin jarang sebuah kata muncul di seluruh dokumen, semakin besar bobotnya dalam sebuah dokumen.

Nantinya, kata yang sering muncul akan diberi nilai terendah. Kemudian, hasil dari pembobotan TF-IDF dapat diimplementasikan menggunakan algoritma yang telah dipersiapkan untuk dibandingkan performanya.

Rumus untuk TF-IDF:

$$tf = 0{,}5 + 0{,}5\ x\ \frac{tf}{max(tf)}$$

$$idf_t = log\left(\frac{D}{df_t}\right)$$

$$W_{d,t} = tf_{d,t}\ x\ IDF_{d,t}$$

penjelasan:

$tf$ = jumlah kemunculan kata yang dicari dalam sebuah dokumen.

max$(tf)$ = jumlah kata (term) yang paling sering muncul dalam dokumen yang sama.

D = total jumlah dokumen.

$df_t$ = jumlah dokumen yang berisi term t.

IDF = Inversed Document Frequency ( log2(D/df)).

d = merujuk pada dokumen ke-d.

t = merujuk pada kata ke-t dari kata kunci yang digunakan.

W = bobot dokumen ke-d dengan term ke-t.

### 2.9 Naïve Bayes Classifier

*Naïve Bayes* merupakan algoritma yang umum digunakan dalam situasi di mana prediksi peristiwa diperlukan atau sebagai metode untuk mengukur probabilitas terjadinya suatu peristiwa. Dalam konteks analisis sentimen, algoritma ini mengasumsikan bahwa munculnya suatu kata tidak memiliki pengaruh terhadap kata lainnya, dengan kata lain, kata-kata dianggap saling independen. Algoritma *Naïve Bayes* digunakan sebagai pengklasifikasi dalam berbagai masalah dunia nyata, seperti analisis sentimen, deteksi spam email, pengelompokan email secara otomatis, penentuan prioritas email, dan klasifikasi dokumen (Andika et al., 2019).

*Naïve Bayes* mengasumsikan bahwa setiap fitur independen dari yang lainnya, sehingga dapat dihitung dengan menggunakan rumus:

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A|B) * P(A)}{P(B)}$$

Dimana:

- P(A|B) = Probabilitas posterior dari A pada kondisi B (*posterior probability*).
- P(B|A) - Probabilitas posterior dari B pada kondisi A (*like hood*).
- P(A) - Probabilitas prior dari A (*class prior probability*).
- P(B) = Probabilitas prior dari B (*predictor prior probability*).

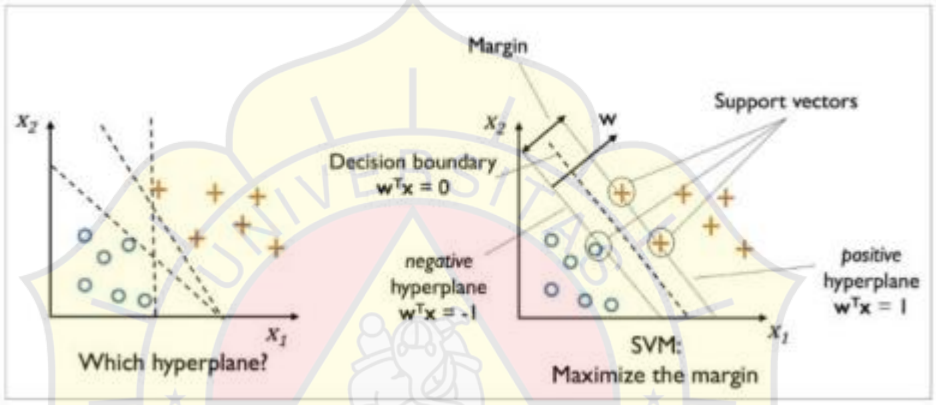
## 2.10 Support Vector Machine (SVM)

*Support Vector Machine* (SVM) adalah sebuah metode dalam pembelajaran mesin yang menggunakan *supervised learning* dan umumnya digunakan untuk melakukan klasifikasi, seperti dalam *Support Vector Classification* (SVC), maupun regresi, seperti dalam *Support Vector Regression* (SVR). Prinsip kerja SVM adalah dengan mencari *hyperplane* yang optimal dengan margin terbesar. *Hyperplane* ini berfungsi sebagai batas keputusan untuk membagi data menjadi dua jenis. (Goh & Lee, 2019).

Dalam konteks pemodelan klasifikasi, SVM memiliki konsep matematis yang lebih solid dan terdefinisi dengan jelas dibandingkan dengan teknik klasifikasi lainnya. SVM memiliki kemampuan untuk menyelesaikan permasalahan klasifikasi dan regresi baik secara linear maupun non-linear. Dengan kata lain, SVM dapat mengatasi berbagai jenis data yang memiliki pola linier maupun pola yang lebih kompleks. Menurut (M. N. Muttaqin & Kharisudin, 2021), meskipun prinsip kerja SVM dapat digunakan untuk kasus klasifikasi *linier*, SVM diciptakan untuk menangani masalah *non-linier* dengan memasukkan konsep carnel ke dalam ruang kerja besar. Prinsip dasar SVM adalah menemukan *hyperplane* terbaik untuk

memisahkan dua kelas di input space. Sebuah *hyperplane* dapat berupa garis dua dimensi dan dapat berupa *flat plane* pada berbagai *multiple plane*.

Dengan menggunakan *support vector* dan *margin score*, klasifikasi dilakukan dengan memeriksa *hyperplane* atau garis batas (*decision boundary*) yang memisahkan satu kelas dengan kelas lainnya yang dapat dipahami lebih lanjut menggunakan gambar




Gambar 2. 2 *Hyperplane* and SVM

Dalam SVM, terdapat dua jenis kernel function yang digunakan, yaitu *kernel linear* dan *kernel non-linear*. *Kernel function* digunakan untuk memetakan data input ke dalam ruang dimensi yang lebih tinggi, sehingga data *input* dapat dipisahkan secara *linear* atau *non-linear*. Macam - macam tipe fungsi *kernel*: (Aggarwal, 2018)

1. *Kernel Linear*

   $K(x, y) = x.y$

2. *Kernel Gaussian Radial Basic Function* (RBF)

   $$\boldsymbol{K}(x_i, x_j) = \exp(-\frac{||x_i - x_j||^2}{2\sigma^2})$$

3. *Kernel Polynomial*

$$K(x_i, x_j) = ((x_i \cdot x_j) + c)^d$$

dengan $x_i$ dan $x_j$ Ini adalah pasangan dari dua data *Training*. Parameter $\sigma, c, d > 0$ Ini adalah sebuah konstanta.

Dalam hal ini, SVM dapat memetakan data input ke dalam ruang fitur yang lebih kompleks dan menghasilkan keputusan klasifikasi yang lebih akurat. Terdapat data *learning* dengan data *points* $x_i(i = 1,2, \dots, m)$ memiliki dua kelas $y_i = \pm 1$ ialah kelas *positive* (+1) dan kelas *negative* (−1) sehingga akan diperoleh *decision function* berikut.

$$f(x) = sign\ (w \cdot x + b)$$

Di mana:

x adalah vektor fitur input

w adalah vektor bobot yang ditemukan selama pelatihan model

b adalah konstanta bias yang ditemukan selama pelatihan model

sign () adalah fungsi yang mengembalikan tanda bilangan (-1 atau 1) sesuai dengan prediksi kelas data input x.

Fungsi *Margin*

Margin = $\frac{1}{||w||}$ Di mana ||w|| adalah norma Euclidean dari vektor bobot w.

Fungsi Objektif

$$\min \tau(w) = \frac{1}{2}\|w\|^2$$

subject to $y_i(w.x_i + b) \geq 1 \quad i = 1,2, \dots, n$

Di mana:

$||w||^2$ adalah kuadrat norma *Euclidean* dari vektor bobot w

$y_i$ adalah label kelas data input x_i, yang bernilai -1 atau 1

n adalah jumlah data pelatihan

$x_i$ adalah vektor fitur dari data pelatihan ke-i

b adalah konstanta bias

subject to menunjukkan bahwa batasan harus dipenuhi untuk semua data pelatihan, yaitu data harus terpisah oleh *hyperplane* dengan margin yang maksimum.

## 2.11 Confusion matrix

*Confusion matrix* merupakan sebuah metode evaluasi performa model *machine learning* yang dapat digunakan untuk mengukur keakuratan prediksi model. *Confusion matrix* digunakan untuk menunjukkan jumlah *True positive* (TP), *False positive* (FP), *True negative* (TN), dan *False negative* (FN) dari prediksi model. Berikut adalah penjelasan singkat mengenai konsep dasar dari *confusion matrix*:

- *True Positive* (TP): Jumlah data *positive* yang berhasil diprediksi dengan benar oleh model.
- *False Positive* (FP): Jumlah data *negative* yang salah diprediksi sebagai *positive* oleh model.
- *True Negative* (TN): Jumlah data *negative* yang berhasil diprediksi dengan benar oleh model.
- *False Negative* (FN): Jumlah data *positive* yang salah diprediksi sebagai *negative* oleh model.

Berikut adalah contoh *confusion matrix:*

Tabel 2. 1 *Confusion Matrix*

| | *Predicted Class* (*Positive*) | *Predicted Class* (*Negative*) |
|---|---|---|
| *Actual Class* (*Positive*) | TP (*True Positive*) | FN (*False Negative*) |
| *Actual Class* (*Negative*) | FP (*False Positive*) | TN (*True Negative*) |

Dengan persamaan berturut-turut sebagai berikut :

Akurasi = $\frac{TP+TN}{TP+TN+FP+FN} \times 100\%$

*Recall* = $\frac{TP}{FN+TP} \times 100\%$

*Precision* = $\frac{TP}{FP+TP} \times 100\%$

*F1 Score* = $2 \times \frac{Precision\ x\ Recall}{Precision+Recall} \times 100\%$

## 2.12 Jurnal Penelitian Sebelumnya

Dalam penelitian yang dilakukan oleh M. Nurul Muttaqin dan Iqbal Kharisudin dengan judul "Analisis Sentimen Aplikasi Gojek Menggunakan Metode *Support Vector Machine* dan *K Nearest Neighbour*", tujuan utamanya adalah mengklasifikasikan ulasan Pengguna Aplikasi Gojek menerapkan metode SVM (*Support Vector Machine*) dan *K Nearest Neighbour* (KNN) dalam penggunaannya. Untuk dataset penelitian yang menganalisis sentimen ini, sumber data terbaik adalah Google Playstore, karena datanya lebih murni dan tidak mengandung iklan atau promosi. Hasil dari penelitian ini menunjukkan bahwa metode KNN mencapai akurasi, presisi, dan *Recall* masing-masing sebesar 82,14%, 82,28%, dan 95,43%, sedangkan metode SVM mencapai akurasi, presisi, dan *Recall* masing-masing sebesar 87,98%, 88,55%, dan 95,43%. Berdasarkan penelitian ini, dalam

melakukan analisis sentimen terhadap ulasan aplikasi Gojek. di *Google Play store*, metode SVM menghasilkan klasifikasi yang lebih baik daripada metode KNN. (M. N. Muttaqin & Kharisudin, 2021).

Dalam penelitian yang kedua yang dilakukan oleh Aan Rohanah, Dwi Latifah Rianti, dan Betha Nurina Sari dengan judul "Perbandingan *Naïve Bayes* dan *Support Vector Machine* untuk Klasifikasi Ulasan Pelanggan Indihome", tujuan utamanya adalah untuk melakukan klasifikasi terhadap ulasan yang diberikan oleh pelanggan Indihome. Penelitian ini mengikuti tahapan penelitian *CRISP-DM* dan menggunakan dua algoritma, yaitu *Naïve Bayes* Classifier dan *Support Vector Machine Kernel Linear*. Data ulasan pelanggan diperoleh dari *platform Twitter* dengan total 1000 tweet menggunakan alat bantu *Rapid Miner* dan perpustakaan R. Hasil penelitian menunjukkan bahwa penerapan algoritma *Support Vector Machine* Kernel Linear menghasilkan performa yang lebih baik daripada algoritma *Naïve Bayes* Classifier. Nilai akurasi yang dicapai adalah sebesar 82,11%, presisi sebesar 76,44%, *Recall* sebesar 88,01%, dan nilai AUC (Area Under Curve) mencapai 0,909. Hal ini menunjukkan bahwa SVM Kernel Linear memberikan kinerja yang lebih optimal dalam melakukan klasifikasi terhadap ulasan pelanggan Indihome berdasarkan data yang diperoleh dari *Twitter* (Rohanah et al., 2021).

Dalam penelitian ketiga yang dilakukan oleh Mujaddid Izzul Fikri, Trifebi Shina Sabrila, dan Yufis Azhar dengan judul "Perbandingan Metode *Naïve Bayes* dan *Support Vector Machine* dalam Analisis Sentimen *Twitter*", tujuan utamanya adalah untuk memahami opini masyarakat tentang Universitas Muhammadiyah Malang dengan menganalisis sentimen tweet yang terkait. Klasifikasi dilakukan

menggunakan metode *Naïve Bayes* dan *Support Vector Machine* (SVM). Hasil penelitian menunjukkan bahwa metode *Naïve Bayes* menghasilkan akurasi yang lebih baik dibandingkan SVM, dengan akurasi mencapai 73,65% (Fikri et al., 2020).

Dalam penelitian terakhir yang dilakukan oleh Elisabet Sinta Romaito, M. Khairul Anam, Rahmaddeni, dan Aniq Noviciate Ulfah dengan judul "Perbandingan Algoritma SVM dan NBC dalam Analisa Sentimen Pilkada di *Twitter*", tujuan utamanya adalah untuk membandingkan algoritma *Support Vector Machine* (SVM) dan naive bayes classifier (NBC) dalam analisis sentimen terkait PILKADA berdasarkan data *Twitter*. Hasil penelitian menunjukkan bahwa algoritma NBC mencapai akurasi 81,7%, *Recall* 81,7%, dan presisi 80%. Sementara itu, algoritma SVM mencapai akurasi 80,7%, *Recall* 80,7%, dan presisi 84%. Dari penelitian ini dapat disimpulkan bahwa algoritma NBC lebih unggul dalam hal akurasi dan recall, sedangkan algoritma SVM lebih unggul dalam hal presisi (Br.Situmorang et al., 2021).

# TEKNOLOGI INFORMASI

# UNIVERSITAS DARMA PERSADA