# ANALISIS SENTIMEN *MYPERTAMINA* MENGGUNAKAN KLASIFIKASI *NAÏVE BAYES* BERSELEKSI FITUR *GENETIC ALGORITHM*

## SKRIPSI

**WARDATUN SAYYIDAH**

**H051181311**

**PRORAM STUDI STATISTIKA DEPARTEMEN STATISTIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS HASANUDDIN**

**MAKASSAR**

**2023**

# ANALISIS SENTIMEN *MYPERTAMINA* MENGGUNAKAN KLASIFIKASI *NAÏVE BAYES* BERSELEKSI FITUR *GENETIC ALGORITHM*

## SKRIPSI

**Diajukan sebagai salah satu syarat memperoleh gelar Sarjana Sains pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin**

**WARDATUN SAYYIDAH**
**H051181311**

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA**
**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**
**UNIVERSITAS HASANUDDIN**
**MAKASSAR**
**DESEMBER 2023**

# LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

**ANALISIS SENTIMEN *MYPERTAMINA* MENGGUNAKAN KLASIFIKASI *NAÏVE BAYES* BERSELEKSI FITUR *GENETIC ALGORITHM***

**adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun**

**Makassar, 14 Desember 2023**


**Wardatun Sayyidah**

**NIM H051181311**

# ANALISIS SENTIMEN *MYPERTAMINA* MENGGUNAKAN KLASIFIKASI *NAÏVE BAYES* BERSELEKSI FITUR *GENETIC ALGORITHM*

**Disetujui Oleh:**

| Pembimbing Utama | Pembimbing Pertama |
|---|---|
| **Siswanto, S.Si., M.Si.** | **Sri Astuti Thamrin, S.Si., M.Stat., Ph.D.** |
| **NIP. 19920107 201903 1 012** | **NIP. 19740713 199903 2 001** |

**Ketua Program Studi**

**Dr. Anna Islamiyati, S.Si., M.Si.**
**NIP. 19770808 200501 2 002**

Pada 14 Desember 2023

# HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama : Wardatun Sayyidah

NIM : H051181311

Program Studi : Statistika

Judul Skripsi : Analisis Sentimen *MyPertamina* Menggunakan Klasifikasi *Naïve Bayes* Berseleksi Fitur *Genetic Algorithm*

**Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.**

DEWAN PENGUJI

1. Ketua : Siswanto, S.Si., M.Si. (.....)
2. Sekretaris : Sri Astuti Thamrin, S.Si., M.Stat., Ph.D (.....)
3. Anggota : Sitti Sahriman, S.Si., M.Si. (.....)
4. Anggota : Dra. Nasrah Sirajang, M.Si. (.....)

Ditetapkan di : Makassar

Tanggal : 14 Desember 2023

# KATA PENGANTAR

***Assalaamu'alaikum Warahmatullahi Wabarakatuh***

Dengan memanjatkan puja dan puji kehadirat Allah *Subhanallahu Wa Ta'ala* atas segala limpahan rahmat dan hidayah-Nya yang telah diberikan kepada penulis sampai saat ini. Shalawat dan salam senantiasa juga tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa sallam*, yang telah membawa kita dari zaman kegelapan kearah yang lebih terang benderang seperti pada saat ini. *Alhamdulillaahirobbil'aalamiin*, berkat rahmat dan kemudahan yang diberikan oleh Allah *Subhanallahu Wa Ta'ala*, penulis dapat menyelesaikan skripsi yang berjudul "**Analisis Sentimen *MyPertamina* Menggunakan Klasifikasi *Naïve Bayes* Berseleksi Fitur *Genetic Algorithm***" sebagai salah satu syarat memperoleh gelar sarjana pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Penulis menyadari bahwa skripsi ini tidak dapat terselesaikan tanpa adanya bantuan, bimbingan dan nasehat dari berbagai pihak. Oleh karena itu, penulis mengucapkan banyak terima kasih serta penghargaan yang setinggi-tingginya untuk orang tua penulis, Ayahanda **Abdul Asis, S.Pd., M.Pd** dan Ibunda **Saripa, S.Pd** yang telah memberikan dukungan, pengorbanan, kasih sayang dan doa yang tak henti-hentinya dipanjatkan kepada penulis untuk sampai pada tahap ini. Tak lupa pula kepada kedua saudara **Khairul Alam** dan **Asyam Jayadi** yang juga selalu memotivasi penulis untuk tetap semangat mengerjakan skripsi ini serta untuk keluarga besar yang tidak dapat penulis tuliskan satu per satu, terima kasih atas doa dan dukungannya selama ini.

Penghargaan yang tulus dan ucapan terima kasih dengan penuh keikhlasan juga penulis ucapkan kepada:

1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, selaku Rektor Universitas Hasanuddin beserta seluru jajarannya.
2. **Bapak Dr. Eng. Amiruddin.**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Ibu Dr. Anna Islamiyati, S.Si., M.Si.**, selaku Ketua Departemen Statistika, segenap Dosen Pengajar dan Staf yang telah membekali ilmu dan kemudahan

kepada penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.

4. **Bapak Siswanto, S.Si., M.Si.**, selaku Pembimbing Utama dan **Ibu Sri Astuti Thamrin, S.Si., M.Stat., Ph.D**, selaku Pembimbing Pendamping yang dengan penuh kesabaran telah meluangkan waktu dan pemikirannya di tengah berbagai kesibukan dan prioritasnya untuk memberikan arahan, dorongan, dan motivasi kepada penulis mulai dari awal pengerjaan hingga selesainya penulisan tugas akhir ini.
5. **Ibu Sitti Sahriman, S.Si., M.Si.** dan **Ibu Dra. Nasrah Sirajang, M.Si.**, selaku Tim Penguji yang telah memberikan saran dan kritikan dalam penyempurnaan penyusunan tugas akhir ini serta waktu yang telah diberikan kepada penulis.
6. **Ibu Sitti Sahriman, S.Si., M.Si.**, selaku Penasehat Akademik penulis. Terima kasih atas segala bantuan, nasehat serta motivasi yang selalu diberikan kepada Penulis selama menjalani pendidikan di Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.
7. Teman-teman **HIMATIKA 2018, HIMASTAT 2018**, **IPPMP UH 2018, dan AIM (Aksi Indonesia Muda)**, terima kasih atas pengalaman berorganisasi yang luar biasa. Kolaborasi dan kerja keras yang telah menciptakan momen berharga dan pengalaman yang telah membentuk penulis menjadi pribadi yang lebih tangguh dan berkembang.
8. Teman-teman **Statistika 2018**, terkhusus **Muh. Ishak, S.Si., Nehemia Millenium Payung, S.Si., Victor Liman, S,Si., Sonya, S.Si., Fiska Evryan, S.Si., dan Hajratul Aswad K, S.Si.**, terima kasih untuk kebersamaan yang telah dilewati menggapai gelar sarjana sains.
9. Teman-teman **Ubur-Ubur**, terkhusus **Fahira Nurul Ichzza, S.Si., Rael Hofni Tandirerung, S.Si., Amalia Andrianingrum, Nurul Nur Kholifah, dan Gilberty Ruben**, terima kasih telah berjuang bersama dari mahasiswa baru hingga mahasiswa tua dan terima kasih telah menjadi pendengar keluh kesah penulis.
10. Tim **Project Bangkit "Cat Pedigree"** yaitu **Zahrizal Ali, S.Kom., Budi Setiawan, S.Kom., Fina Enno Rizki Oktavia, S.Kom., Robby Ramadhan**

**A, S.Kom., dan Bugi Sulistiyo**, terima kasih untuk segala dukungan dan kebersamaan.

11. Teman-teman **KKN 106 Tamalanrea 6**, terkhusus **Egidiaw dan Dappa**, terima kasih telah menjadi wadah penulis dalam bertukar pikiran dalam berbagai disiplin ilmu.
12. Teman-teman **Nur Azizah Harun, S.Kom dan Nurchairul Akbar Saad, S.Kom., M.T**, terima kasih untuk segala dukungan dan masukan saat penulis menghadapi kesulitan dan hambatan dalam penyelesaian magang.
13. Sahabat penulis, **Eny, Uley, Vira, Egidiaw, Emay, Iis, Bayu, Amal**, yang telah menjadi tempat curhat penulis dalam berbagai hal.
14. Kepada semua pihak yang tidak dapat penulis sebutkan satu-persatu, semoga segala dukungan dan partisipasi yang diberikan kepada penulis bernilai ibadah disisi Allah *Subhanallahu Wa Ta'ala.*

Penulis menyadari bahwa masih banyak kekurangan dalam skripsi ini, untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga tulisan ini memberikan manfaat untuk pembaca.

***Wassalamu'alaikum Warahmatullahi Wabarakatuh***

Makassar, 14 Desember 2023

Wardatun Sayyidah

# PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIK

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

| | |
|---|---|
| Nama | : Wardatun Sayyidah |
| NIM | : H051181311 |
| Program Studi | : Statistika |
| Departemen | : Statistika |
| Fakultas | : Matematika dan Ilmu Pengetahuan Alam |
| Jenis Karya | : Skripsi |

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Non-ekslusif** (***Non-exclusive Royalty- Free Right***) atas tugas akhir saya yang berjudul:

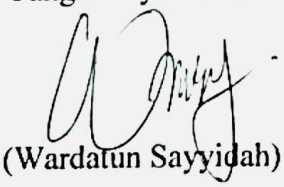
**"Analisis Sentimen *MyPertamina* Menggunakan Klasifikasi *Naïve Bayes* Berseleksi Fitur *Genetic Algorithm*"**

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar pada tanggal, 14 Desember 2023

Yang menyatakan



(Wardatun Sayyidah)

## ABSTRAK

PT. Pertamina menciptakan loyalitas program dengan tujuan memfasilitasi transaksi non-tunai melalui aplikasi *MyPertamina.* Perubahan kebijakan yang mewajibkan pengendara roda empat untuk mendaftar di *MyPertamina* dan melakukan uji coba yang telah memicu perbincangan di media sosial, khususnya *Twitter*. Pengklasifikasian ke dalam kelompok positif dan negatif dalam menggunakan aplikasi *MyPertamina* memungkinkan untuk memahami pandangan publik dengan efektif untuk meningkatkan pelayanan dan pengambilan keputusan yang lebih baik. Penelitian ini bertujuan untuk mengetahui hasil klasifikasi dan kinerja metode klasifikasi pada data opini publik tentang aplikasi *MyPertamina.* Metode klasifikasi yang digunakan adalah metode *Naïve Bayes* dengan seleksi fitur *Genetic Algorithm*. Klasifikasi sentimen mengungkapkan sentimen negatif mendominasi, menunjukkan kritik terhadap aplikasi *MyPertamina* menghasilkan akurasi 91,67%, presisi 94,44%, dan recall 85%. Dari perhitungan kinerja tersebut, dapat disimpulkan bahwa metode klasifikasi *Naïve Bayes* dengan seleksi fitur *Genetic Algorithm* dikategorikan sebagai "*excellent classification*" menurut metode ROC.


**Kata Kunci:** *genetic algorithm, mypertamina, naïve bayes, twitter.*

## ABSTRACT

*PT. Pertamina has established a loyalty program with the aim of facilitating non-cash transactions through the MyPertamina application. The policy change mandating four-wheeled vehicle drivers to register on MyPertamina and undergo a trial has sparked discussions on social media, particularly on Twitter. Classifying opinions into positive and negative groups regarding the use of the MyPertamina application enables an effective understanding of public perspectives to enhance service and make better decisions. This research aims to determine the classification results and the performance of the classification method on public opinions about the MyPertamina application. The classification method employed is the Naïve Bayes method with Genetic Algorithm feature selection. Sentiment classification reveals that negative sentiments dominate, indicating criticism of the MyPertamina application, with an accuracy of 91.67%, precision of 94.44%, and recall of 85%. Based on these performance calculations, it can be concluded that the Naïve Bayes classification method with Genetic Algorithm feature selection is categorized as "excellent classification" according to the ROC method.*


***Keywords:*** *genetic algorithm, mypertamina, naïve bayes, twitter.*

# DAFTAR ISI

**HALAMAN SAMPUL** ..... **i**
**HALAMAN JUDUL** ..... **ii**
**LEMBAR PERNYATAAN KEOTENTIKAN** ..... **iii**
**HALAMAN PENGESAHAN** ..... **v**
**KATA PENGANTAR** ..... **vi**
**PERNYATAAN PERSETUJUAN PUBLIKASI** ..... **ix**
**ABSTRAK** ..... **x**
***ABSTRACT*** ..... **xi**
**DAFTAR ISI** ..... **xii**
**DAFTAR TABEL** ..... **xiv**
**DAFTAR GAMBAR** ..... **xv**
**BAB I PENDAHULUAN** ..... **1**
1.1 Latar Belakang ..... 1
1.2 Rumusan Masalah ..... 3
1.3 Batasan Masalah ..... 4
1.4 Tujuan Penelitian ..... 4
1.5 Manfaat Penelitian ..... 4
**BAB II TINJAUAN PUSTAKA** ..... **5**
2.1 *Text Mining* ..... 5
2.2 *Twitter* ..... 5
2.3 Praproses Teks ..... 6
2.4 Analisis Sentimen ..... 7
2.5 *Naïve Bayes Classifier* ..... 7
2.6 Seleksi Fitur dengan *Genetic Algorithm* ..... 10
2.7 *Word Cloud* ..... 11
2.8 Evaluasi Kinerja Algoritma ..... 12
2.9 *MyPertamina* ..... 14
**BAB III METODOLOGI PENELITIAN** ..... **15**
3.1 Jenis dan Sumber Data Penelitian ..... 15
3.2 Struktur Data ..... 15
3.3 Metode Analisis Data ..... 16

**BAB IV HASIL DAN PEMBAHASAN ..... 18**
4.1 Deskripsi Data ..... 18
4.2 *Preprocessing* Data ..... 20
4.3 Melakukan Seleksi Fitur Menggunakan *Genetic Algorithm* (GA) ..... 28
4.4 Klasifikasi dengan Algoritma *Naïve Bayes* ..... 32
4.4.1 *Data Training dan Data Testing* ..... 32
4.4.2 Pengklasifikasian Manual Sentimen ..... 32
4.4.3 Perhitungan Peluang Kemunculan Kategori Kelas Sentimen ..... 34
4.4.4 Perhitungan Probabilitas Fitur ..... 34
4.4.5 Perhitungan Nilai *Hmap* ..... 35
4.4.6 Hasil Klasifikasi Sentimen ..... 36
4.5 Evaluasi Kinerja Algoritma ..... 36
4.6 Visualisasi ..... 38
4.6.1 Visualisasi Kelas Sentimen ..... 38
4.6.2 Visualisasi *Word Cloud* ..... 39
**BAB V PENUTUP ..... 41**
5.1 Kesimpulan ..... 41
5.2 Saran ..... 41
**DAFTAR PUSTAKA ..... 42**
**LAMPIRAN ..... 45**
Lampiran 1 Data *Tweet* Hasil *Crawling* ..... 46
Lampiran 2 Fitur pada *Data Training* Setiap Kelas ..... 49

# DAFTAR TABEL

**Tabel 2. 1** *Confusion Matrix*..... 12
**Tabel 3. 1** Contoh Struktur Data Penelitian ..... 15
**Tabel 4. 1** Hasil *Crawling* Data..... 18
**Tabel 4. 2** Struktur Data *Tweet* Setelah Melakukan *Labeling* Manual..... 19
**Tabel 4. 3** Struktur Data Sebelum dan Setelah Penghapusan URL..... 21
**Tabel 4. 4** Struktur Data Sebelum dan Setelah Penghapusan *username*, angka, dan tanda baca..... 22
**Tabel 4. 5** Struktur Data Sebelum dan Setelah Proses *Case Folding*..... 23
**Tabel 4. 6** Struktur Data Sebelum dan Setelah Proses *Spelling Normalization*... 24
**Tabel 4. 7** Struktur Data Sebelum dan Setelah Proses *Stemming* ..... 25
**Tabel 4. 8** Struktur Data Sebelum dan Setelah Proses *Stopword Removal*..... 26
**Tabel 4. 9** Struktur Data Sebelum dan Setelah Proses *Tokenizing*..... 26
**Tabel 4. 10** Struktur Data Sebelum dan Setelah *Preprocessing* Data ..... 27
**Tabel 4. 11** Populasi Berukuran 100 Kromosom ..... 28
**Tabel 4. 12** *Crossover* Kromosom Parent menjadi Kromosom *Child*..... 29
**Tabel 4. 13** Mutasi Kromosom *Child* menjadi Kromosom *Mutated Child*..... 30
**Tabel 4. 14** Fitur Terpilih dari Seleksi Fitur Menggunakan GA ..... 31
**Tabel 4. 15** Struktur Data Sebelum dan Seteleh Seleksi Fitur dengan GA ..... 31
**Tabel 4. 16** Proporsi Data *Training* dan Data *Testing*..... 32
**Tabel 4. 17** Struktur Data Hasil Seleksi Fitur ..... 33
**Tabel 4. 18** Frekuensi Kemunculan Fitur *Tweet-1* ..... 33
**Tabel 4. 19** Nilai Peluang Kemunculan Kelas Sentimen ..... 34
**Tabel 4. 20** Probabilitas Tiap Fitur..... 35
**Tabel 4. 21** Klasifikasi Kelas Sentimen ..... 36
**Tabel 4. 22** *Confusion Matrix* Klasifikasi Algoritma *Naïve Bayes*..... 37

# DAFTAR GAMBAR

**Gambar 2. 1** *Conditional Independence* pada *Naïve Bayes*..... 8
**Gambar 4. 1** Visualisasi Kelas Sentimen..... 38
**Gambar 4. 2** Visualisasi *Word Cloud* ..... 42

# BAB I
# PENDAHULUAN

## 1.1 Latar Belakang

Tingginya jumlah penduduk sejalan dengan tingginya jumlah kendaraan yang dibutuhkan, sehingga kebutuhan untuk transportasi yaitu bahan bakar minyak (BBM) juga meningkat. Perusahaan yang menyediakan BBM adalah PT. Pertamina yang merupakan perusahaan BUMN yang bertugas mengelola penambangan minyak dan gas bumi untuk memenuhi kebutuhan masyarakat Indonesia (Saddam, 2022). PT. Pertamina menciptakan *Loyalty Program* yaitu sebuah aplikasi yang dapat diakses melalui *smartphone* yang diharapkan dapat memudahkan konsumen dalam bertransaksi dengan menerapkan pembayaran non-tunai atau melalui aplikasi *MyPertamina* (Heryadi, 2018). Pemerintah melalui PT. Pertamina menetapkan aturan baru yaitu masyarakat pengendara roda empat yang ingin mendapatkan BBM jenis Pertalite dan Solar harus terdaftar di sistem *MyPertamina* dan dilakukan uji coba mulai 1 Juli 2022 (Kompas.com, 2022). Sehingga publik menyoroti peraturan ini dan memberikan komentar terhadap aplikasi *MyPertamina* di media sosial, salah satunya melalui *Twitter* (Kurniawan, 2017).

*Twitter* memiliki keterbukaan untuk data yang dimiliki melalui API (*Applicattion Programming Interface*). Melalui API, *Tweets* yang terdapat dalam *Twitter* dapat diakses sesuai dengan kebutuhan pengguna baik kata kunci (*keyword*) maupun rentang waktu yang dibutuhkan sehingga informasi menjadi mudah didapatkan dan diolah menjadi suatu informasi yang berguna. Hal tersebut yang membuat *Twitter* sebagai *microblog* yang banyak diminati perusahaan, organisasi, maupun individu dalam mengumpulkan opini publik untuk suatu topik atau kasus tertentu. *Tweets* yang dikumpulkan dan dianalisis disebut analisis sentimen (Ikasari dkk., 2020).

Analisis sentimen atau *opinion mining* termasuk ke dalam salah satu bidang dari *Natural Language Processing* (NLP) dan merupakan metode yang digunakan untuk menganalisa pendapat untuk melihat kecenderungan seseorang menilai suatu objek yang bersentimen positif atau negatif (Liu, 2012). Analisis sentimen merupakan bidang penelitian yang popular dan dianggap mampu memberikan keuntungan dalam berbagai aspek, seperti: prediksi harga saham, isu politik,

kepuasan terhadap suatu produk atau layanan, analisis reputasi, dan sebagainya. Adapun beberapa metode klasifikasi dalam menyelesaikan analisis sentimen, yaitu: *K-Nearest Neighbors* (KNN), *Decision Tree, Naïve Bayes, Support Vector Machine* (SVM), dan lain-lain (Fikri dkk., 2020). Salah satu metode yang populer dalam pengklasifikasian pada analisis sentimen adalah metode *Naïve Bayes.* Algoritma *Naïve Bayes* salah satunya digunakan untuk klasifikasi teks serta merupakan metode *machine learning* yang menggunakan perhitungan probabilitas dan statistik yang dikemukakan oleh Thomas Bayes (Chai dkk., 2002). *Naïve Bayes* bekerja dengan menghitung probabilitas dari setiap fitur dalam data yang akan diklasifikasikan untuk tiap kategori. Adapun keuntungan dari metode *Naïve Bayes* yaitu sangat cocok untuk digunakan pada klasifikasi data teks, dapat digunakan pada *dataset* yang besar, dan bekerja sangat baik pada *multi-class prediction* (Pramana dkk., 2018). Namun, *Naïve Bayes* memiliki kekurangan yaitu jika terlalu banyak jumlah fitur, tidak hanya meningkatkan waktu penghitungan tapi juga menurunkan akurasi klasifikasi (Uysal dan Gunal, 2012).

Hal lain yang ditemukan dalam analisis sentimen adalah pemilihan fitur. Suatu algoritma biasanya akan bekerja lebih baik lagi jika dilakukan pemilihan fitur pada datanya (Pushpalata dan Gupta, 2012). Seleksi fitur bisa membuat pengklasifikasi lebih efektif dan efisien dengan menentukan jumlah fitur yang digunakan dan mengurangi fitur yang tidak relevan dalam klasifikasi. Terdapat dua jenis utama metode pemilihan fitur dalam *machine learning*, yaitu: *wrapper* dan *filter*. *Wrapper* mengevaluasi fitur secara berulang dan menghasilkan akurasi klasifikasi yang tinggi (Muthia, 2017). Salah satu metode *wrapper* yang bisa digunakan adalah *Genetic Algorithm* (GA). GA merupakan suatu algoritma optimasi yang dapat digunakan dengan prinsip-prinsip seleksi alam dan genetika alami (Wati, 2016). GA dibangkitkan sebuah populasi yang terdiri dari beberapa individu yang mempresentasikan suatu kombinasi kata yang dianggap sebuah solusi untuk masalah yang akan dipecahkan. Setiap individu berisi beberapa kromosom yang digunakan untuk menghitung *fitness value* untuk menemukan *parent* yang akan digunakan pada tahap *crossover* dan mutasi. Menurut Muthia (2017), GA mudah disejajarkan dan telah digunakan untuk klasifikasi.

Berdasarkan penelitian yang dilakukan oleh Fikri dkk (2020) yang membandingkan metode *Naïve Bayes* dan *Support Vector Machine* pada dataset *tweet* yang berkaitan dengan Universitas Muhammadiyah Malang, mendapatkan hasil metode *Naïve Bayes* memiliki hasil akurasi sebesar 73,65% yang lebih baik dibandingkan dengan SVM dengan akurasi 70,20%. Penelitian lainnya yang dilakukan oleh Hayuningtyas dan Sari (2019), meneliti analisis sentimen ulasan tempat wisata Taman Mini Indonesia Indah menggunakan metode *Naïve Bayes* dan *Particle Swarm Optimazion* (PSO) yang bertujuan untuk meningkatkan nilai akurasi dengan kesimpulan yaitu hasil akurasi yang didapatkan sebesar 94,02%.

Penggunaan pengklasifikasi Naïve Bayes dengan *Genetic Algorithm* dalam analisis sentimen terhadap *MyPertamina* memiliki beberapa alasan, yaitu membantu PT. Pertamina memahami pandangan pengguna dan tingkat kepuasan pengguna terhadap layanan *MyPertamina*. Pengklasifikasi sentimen dari peraturan baru yang diterapkan, memungkinkan perusahaan untuk memantau respon publik dan mengidentifikasi masalah yang mungkin muncul. Selain itu, analisis sentimen juga membantu dalam mengidentifikasi area-area yang perlu perbaikan dalam aplikasi *MyPertamina* yang dapat digunakan untuk mengembangkan perbaikan dan peningkatan layanan. Terakhir, hasil analisis sentimen membantu dalam pengambilan keputusan strategis dan menentukan kebijakan yang lebih baik sesuai dengan kebutuhan pengguna. Hal tersebut memungkinkan perusahaan untuk mengumpulkan, memahami, dan merespons pandangan publik dengan lebih efektif, dengan hasil yang berpotensi meningkatkan pelayanan dan pengambilan keputusan yang lebih baik.

Berdasarkan uraian di atas, maka penulis tertarik melakukan penelitian pengklasifikasi *Naïve Bayes* dengan *Genetic Algorithm* sebagai metode pemilihan fitur akan diterapkan untuk mengklasifikasi penggunaan aplikasi *MyPertamina*.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang, rumusan masalah pada penelitian ini yaitu "Bagaimana hasil klasifikasi dan kinerja klasifikasi *Naïve Bayes* dengan *Genetic Algorithm* sebagai seleksi fitur dalam klasifikasi opini publik tentang aplikasi *MyPertamina*?"

## 1.3 Batasan Masalah

Batasan masalah dari penelitian ini adalah sebagai berikut:

1. *Tweets* yang diambil dan dianalisa hanya *tweets* berbahasa Indonesia yang diunggah pada tanggal 1 Juni – 30 November 2022.
2. Data yang digunakan pada penelitian ini adalah data dari media sosial *Twitter* dari akun yang tidak *private* menggunakan kata kunci "*MyPertamina*".

## 1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, tujuan dari penelitian ini yaitu "memperoleh hasil klasifikasi dan kinerja klasifikasi *Naïve Bayes* dengan *Genetic Algorithm* sebagai seleksi fitur dalam klasifikasi opini publik tentang aplikasi *MyPertamina*"

## 1.5 Manfaat Penelitian

Manfaat yang diharapkan oleh penulis dari hasil penelitian adalah:

1. Bagi peneliti, dapat memberikan kontribusi akademis untuk pengembangan ilmu pengetahuan dalam bidang matematika dan statistika, khususnya yang berkaitan dengan analisis sentimen menggunakan *Naïve Bayes* dan *Genetic Algorithm* sebagai metode seleksi fitur.
2. Bagi pihak-pihak yang terkait, dapat memberikan informasi mengenai opini masyarakat terhadap aturan baru pemerintah untuk terdaftar pada sistem *MyPertamina* sehingga dapat menjadi acuan dalam penetapan kebijakan, pengambilan keputusan, dan perumusan perencanaan program/kegiatan di tahun berikutnya demi mencapai sasaran yang diinginkan dan yang telah ditetapkan.

# BAB II
# TINJAUAN PUSTAKA

## 2.1 *Text Mining*

*Text mining* adalah proses untuk memperoleh informasi berkualitas tinggi dari teks. Informasi berkualitas tinggi biasanya didapatkan karena memperhatikan pola dan tren dengan cara mempelajari pola statistik (Deolika dkk., 2019). *Text mining* juga merupakan teknik yang digunakan untuk mengekstraksi informasi yang berguna dari data teks yang tidak terstruktur (Kurniawan dan Susanto, 2019). Tujuan dari *text mining* yaitu untuk mendapatkan informasi yang berguna dari sekumpulan dokumen. Sehingga, sumber daya yang digunakan dalam *text mining* adalah sekumpulan *text*. Pada dasarnya proses kerja dari *text mining* banyak mengadopsi dari penelitian data mining namun menjadi berbeda karena pola yang digunakan oleh *text mining* diambil dari sekumpulan Bahasa alami yang tidak terstruktur. Tahapan dari *text mining* secara umum adalah praproses teks dan *feature selection* (Kurniawan, 2017).

## 2.2 *Twitter*

*Twitter* adalah sebuah situs web yang dimiliki dan dioperasikan oleh *Twitter Inc.*, yang menawarkan jaringan sosial berupa *microblog* sehingga memungkinkan pengguna untuk mengirim dan membaca pesan *tweets. Microblog* adalah salah satu jenis alat komunikasi *online* yang digunakan untuk memperbarui status tentang mereka yang sedang memikirkan dan melakukan sesuatu. Seperti pendapat tentang suatu objek atau fenomena tertentu (Rizki, 2019).

Berbeda dari sosial media lain, *Twitter* memiliki keterbukaan untuk data yang dimiliki melalui *Applicattion Programming Interface* (API). Menurut Rizki (2019), API merupakan perintah-perintah untuk menggantikan Bahasa yang digunakan dalam *system calls.* Fungsi yang dibuat dengan menggunakan API tersebut kemudian akan memanggil *system calls* sesuai dengan sistem operasinya. *Twitter* API terdiri dari tiga bagian yaitu:

1. *Search* API

*Search* API dirancang untuk memudahkan *user* dalam mengelola *query search* di konten *Twitter*. *User* dapat menggunakannya untuk mencari *tweet*

berdasarkan *keyword* khusus atau mencari *tweet* lebih spesifik berdasarkan *username Twitter*. *Search* API juga menyediakan akses pada data *trending topic.*

2. *Representational State Transfer* (REST API)

REST API memperbolehkan *developer* untuk mengakses inti dari *Twitter* seperti *timeline, status update,* dan informasi *user*. REST API digunakan dalam membangun sebuah aplikasi *Twitter* yang kompleks yang memerlukan inti dari *Twitter.*

3. *Steaming* API

*Steaming* API digunakan *developer* untuk kebutuhan yang lebih intensif seperti melakukan penelitian dan analisa data. *Steaming* API dapat menghasilkan aplikasi yang dapat mengetahui statistik *status update*, *followers*, dan lain sebagainya.

## 2.3 Praproses Teks

Pada tahap praproses teks sebagai awal dari pengolahan setiap data untuk mengubah bentuk dokumen menjadi data yang terstruktur. Praproses teks berguna untuk membersihkan *noise* (kesalahan acak) dokumen teks yang tidak mempunyai makna penting pada dokumen tersebut yang bertujuan untuk meningkatkan akurasi klasifikasi (Kurniawan, 2017). Berikut merupakan tahapan dari praproses teks:

1. *Case Folding*

*Case Folding* adalah proses penyamaan *case* sebuah dokumen, dilakukan untuk mempermudah pencarian. Peran *Case Folding* yaitu mengkonversi keseluruhan teks dalam dokumen menjadi menjadi bentuk standar (*lowercase*).

2. *Spelling Normalization*

*Spelling Normalization* merupakan proses perbaikan atau subtitusi kata-kata yang salah eja atau disingkat dalam bentuk tertentu. Subtitusi kata dilakukan untuk menghindari jumlah perhitungan dimensi kata yang melebar (Khotimah, 2019).

3. *Stopword Removal*

Kata-kata yang terkandung pada daftar *stopword* yang terdapat pada daftar kata *stopword* Bahasa Indonesia berisi kata-kata yang sering muncul namun tidak memiliki makna. Contoh kata “bagaimana”, “juga”, “agar”, dan “jadi” terdapat di tabel kata *stopword* sehingga kata tersebut harus dihilangkan (Ikasari, Fajarmawati, dan Widiastuti, 2020).

4. *Tokenization*

*Tokenization* merupakan proses pemotongan *string* input berdasarkan kata yang menyusunnya serta membedakan karakter-karakter tertentu yang dapat dilakukan sebagai pemisah kata atau bukan (Ikasari, Fajarmawati, dan Widiastuti, 2020). Hasil tokenisasi berguna untuk analisis teks lebih lanjut.

5. *Stemming*

Pada tahap *Stemming* dilakukan untuk mengubah kata ke bentuk dasar dengan cara menghilangkan imbuhan-imbuhan pada kata dalam dokumen (Ikasari, Fajarmawati, dan Widiastuti, 2020).

## 2.4 Analisis Sentimen

Analisis sentimen merupakan cara untuk menilai opini tertulis atau lisan untuk menentukan opini bersifat positif, negatif atau netral (Alsaeedi dan Khan, 2019). Menurut Liu (2012) Analisis sentimen atau yang sering juga disebut sebagai penggalian opini (*opinion mining*) adalah bidang studi yang menganalisa pendapat, sentimen, evaluasi, penilaian dan emosi orang-orang terhadap suatu entitas seperti produk, layanan, organisasi, permasalahan, peristiwa, topik, dan atribut. Analisis sentimen dilakukan untuk mengetahui kecenderungan opini seseorang terhadap sebuah peristiwa atau masalah, apakah cenderung beropini positif atau negatif (Kurniawan dan Susanto, 2019). Pada dasarnya analisis sentimen merupakan tahapan klasifikasi, namun tahapan klasifikasi sentimen pada *Twitter* yang tidak terstruktur menyebabkan sedikit lebih sulit dibanding dengan klasifikasi dokumen terstruktur. Langkah pertama adalah untuk mengklasifikasikan apakah kalimat mengungkapkan pendapat atau tidak. Langkah kedua mengklasifikasin kalimat-kalimat pendapat menjadi positif dan negatif (Rizki, 2019).

## 2.5 *Naïve Bayes Classifier*

*Naïve Bayes Classifier* merupakan pengklasifikasian statistik yang dapat digunakan untuk mempediksi probabilitas keanggotaan suatu *class* (Wijaya dan Dwiasnati, 2020). Keunggulan menggunakan metode *Naïve Bayes Classifier* dibandingkan metode lainnya yaitu metode pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan dari suatu kelas, selain itu terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam

*database* yang besar (Miranda, 2018). Menurut Wibawa dkk (2019) terdapat beberapa keuntungan menggunakan algoritma *Naïve Bayes* diantaranya adalah:

1. Sangat cocok untuk digunakan pada klasifikasi data teks.
2. Dapat digunakan pada *dataset* yang besar.
3. Mudah dan cepat untuk memprediksi klasifikasi *dataset* dan bekerja sangat baik pada *multi-class prediction.*
4. Bekerja sangat cepat sehingga dapat digunakan untuk memprediksi secara *real time.*

Algoritma *Naïve Bayes* digunakan untuk memprediksi probabilitas di masa depan berdasarkan pengalaman dari masa lalu. Teorema Bayes (Zhang, 2004) adalah sebagai berikut:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)}, \qquad P(X) > 0 \tag{2.1}$$

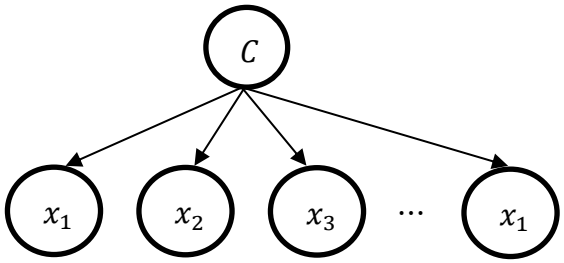
Keterangan:

| | | |
|---|---|---|
| $X$ | : | Atribut |
| $C$ | : | Kelas |
| $P(C\|X)$ | : | Probabilitas kejadian bersyarat $C$ dengan syarat $X$ terjadi |
| $P(X\|C)$ | : | Probabilitas kejadian bersyarat $X$ dengan syarat $C$ terjadi |
| $P(C)$ | : | Probabilitas kejadian $C$ |
| $P(X)$ | : | Probablitas kejadian $X$ |

$X$ dapat ditulis sebagai berikut:

$$X = (x_1, x_2, x_{3,} \cdots, x_n)$$

Hubungan antara $C$ dan $X$ dapat dilihat pada Gambar 2.1.




**Gambar 2. 1** *Conditional Independence* pada *Naïve Bayes*

Pada algoritma *Naïve Bayes Classifier* setiap dokumen dipresentasikan dengan pasangan atribut "$x_1, x_2, \dots, x_n$". $x_1$ merupakan kata pertama, $x_2$ merupakan kata kedua dan seterusnya. Sedangkan $C$ adalah himpunan kategori *tweet.* Pada persamaan (2.1) dapat dituliskan sebagai berikut.

$$P(C|X) = \frac{P(X,C)}{P(C)} \cdot \frac{P(C)}{P(X)} = \frac{P(X,C)}{P(X)}$$

dengan,

$$\left.\begin{aligned} P(C|x_1) &= \frac{P(x_1,C)}{P(x_1)} \\ P(C|x_2) &= \frac{P(x_2,C)}{P(x_2)} \\ &\vdots \\ P(C|x_n) &= \frac{P(x_n,C)}{P(x_n)} \end{aligned}\right\} \quad P(C|X) \propto P(X,C)$$

Selanjutnya, $P(C|X)$ dapat diuraikan menjadi:

$$\begin{aligned} P(C|X) &= P(X,C) \\ &= P(x_1, x_2, \dots, x_n, C) \\ &= P(x_1|x_2, x_3, \dots, x_n, C) \cdot P(x_2, x_3, \dots, x_n, C) \\ &= P(x_1|x_2, x_3, \dots, x_n, C) \cdot P(x_2, x_3, \dots, x_n, C) \\ &= P(x_1|x_2, x_3, \dots, x_n, C) \cdot P(x_2|x_3, x_3, \dots, x_n, C) \cdots P(x_{n-1}|x_n, C) \\ &\qquad \cdot P(x_n|C) \cdot P(C) \end{aligned}$$

Sekarang asumsi *conditional independence* diberlakukan yaitu asumsikan bahwa seluruh fitur di $X$ saling independent yang artinya kemunculan suatu fitur tidak tergantung pada kemunculan kata atau kelompok lainnya, sehingga $P(C|X)$ dapat dituliskan sebagai berikut.

$$P(C|x_1, x_2 \cdots x_n) = P(C) \prod_{i-1}^{n} P\,(x_i|C)$$

Adapun beberapa persamaan yang dapat digunakan dalam *Naïve Bayes Classifier* adalah sebagai berikut.

$$P(v_j) = \frac{|docs|}{|training|} \tag{2.2}$$

Keterangan:

$P(v_j)$ : Probabilitas setiap data terhadap sekumpulan data

$|docs|$ : Frekuensi data pada setiap kelas/kategori

$|training|$ : Jumlah data *training*

$$P(w_k|v_j) = \frac{n_k + 1}{|n + jumlah\ kata\ dalam\ kelas|} \tag{2.3}$$

Keterangan:

$P(w_k|v_j)$ : Probabilitas kemunculan kata $w_k$ pada suatu data

$n_k$ : Frekuensi kata ke$-k$ setiap kategori

Menurut Koller dan Friedman (2009), *Naïve Bayes Classifier* perlu memaksimalkan nilai probabilitas dari setiap kelas yang dinyatakan sebagai *Hypothesis Maximum A Posteriori* $(H_{map})$. Adapun persamaan $H_{map}$ sebagai berikut.

$$H_{map} = \underset{\{positif,negatif\}}{\arg max}\ P(w_k|c)\,P(c) \tag{2.4}$$

Keterangan:

$H_{map}$ : Nilai probabilitas tertinggi data dari masing-masing kelas

$P(w_k|c)$ : Probabilitas kemunculan kata $w_k$ dalam kelas *c*

$P(c)$ : Probabilitas kelas

## 2.6 Seleksi Fitur dengan *Genetic Algorithm*

Seleksi fitur digunakan untuk menghilangkan fitur yang tidak relevan dan berulang yang memungkinkan untuk menyebabkan kekacauan. Menurut Wati (2016) GA merupakan suatu algoritma optimasi yang dapat digunakan dengan prinsip-prinsip seleksi alam dan genetika alami. GA memiliki susunan dari parameter ruang pencarian yang berbentuk kode yang dinamakan kromosom. Cara kerja dari seleksi fitur GA yaitu dengan melakukan proses pencarian kromosom

terbaik diantara kromosom yang lain dengan tujuan memberikan solusi optimal dari fungsi objektif permasalahan optimasi (Suguna dan Thanushkodi, 2010). Terdapat langkah-langkah yang digunakan dalam optimasi pada seleksi fitur GA antara lain:

1. Inisialisasi Populasi

   Inisialisasi populasi merupakan tahap awal dalam melakukan proses GA. Proses ini dilakukan untuk menentukan individu-individu (kromosom) yang akan digunakan dalam menjalankan GA.

2. Seleksi Kromosom

   Seleksi kromosom dilakukan untuk memilih kromosom yang sebelumnya telah ditentukan pada tahap pertama. Pemilihan kromosom berdasarkan pada *fitness value* dari kromosom yang terpilih atau kromosom yang terbaik dari yang lainnya. Kromosom yang terpilih disebut *parent*.

3. *Crossover*

   *Crossover* merupakan salah satu cara yang melakukan pemilihan induk (*parent*) sebanyak dua individu dari populasi secara acak. Agar mendapatkan hasil keturunan atau biasa disebut dengan *child.*

4. Mutasi

   Mutasi merupakan cara reproduksi yang menentukan satu individu dari populasi secara *random.* Mutasi merupakan tahap informasi pada *child* akan dimutasi secara *random* sehingga menghasilkan gen baru yang disebut dengan *mutated child.*

5. Mengembalikan *mutated child* ke populasi

   Mengembalikan *mutated child* ke populasi artinya yaitu dilakukan pengembalian nilai *mutated child* pada populasi sebelumnya dengan syarat bahwa gen *parent* yang terpilih sebelumnya akan digantikan menjadi kromosom baru dari *mutated child.*

6. Proses akan terus berjalan sampai target yang diinginkan terpenuhi. Jika tidak memenuhi persyaratan maka dilakukan perulangan dari langkah ke-2.

## 2.7 *Word Cloud*

*Word cloud* merupakan kumpulan kata-kata yang paling banyak muncul dalam data teks yang dianalisis. Kata-kata tersebut terkumpul seperti sebuah gumpalan awan yang berisi kata-kata yang digunakan, ditunjukkan dengan ukuran

huruf pada kata. Semakin besar huruf dari kata menunjukkan bahwa semakin sering kata tersebut muncul (Adiyana dan Hakim, 2015).

*Word cloud* merupakan salah satu teknik visualisasi yang terdiri dari kata-kata yang paling banyak muncul saat *dataset* dianalisis. Besar kecilnya huruf ditentukan berdasarkan intensitas keseringan kata digunakan. Semakin sering digunakan maka semakin besar ukuran dari kata tersebut. *Word cloud* digunakan agar hasil analisis terlihat lebih menarik dan lebih mudah untuk dimengerti.

## 2.8 Evaluasi Kinerja Algoritma

Pengukuran kinerja algoritma dapat dilakukan menggunakan bantuan *confusion matrix*. *Confusion matrix* merupakan suatu metode yang biasanya digunakan untuk melakukan perhitungan akurasi. *Confusion matrix* digambarkan dengan tabel yang menyatakan jumlah data uji yang benar diklasifikasikan dan jumlah data uji yang salah dikasifikasikan (Rahman dkk., 2017). Berikut tampilan dari tabel *confusion matrix.*

**Tabel 2. 1** *Confusion Matrix*

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

Pada Tabel 2.1 terdapat beberapa informasi yang dihasilkan dari *confusion matrix*. Berikut penjelasan dari masing-masing informasi dari tabel di atas:

1. *True Positive* (TP) merupakan kinerja klasifikasi memprediksi data sebagai ya (TRUE) dan jawaban aktualnya adalah ya (TRUE).
2. *True Negative* (TN) merupakan kinerja klasifikasi memprediksi data sebagai tidak (FALSE) dan jawaban aktualnya adalah tidak (FALSE).
3. *False Positive* (FP) merupakan kinerja klasifikasi memprediksi data sebagai ya (TRUE) dan jawaban aktualnya tidak (FALSE).
4. *False Negative* (FN) merupakan kinerja klasifikasi memprediksi data sebagai tidak (FALSE) dan jawaban aktualnya adalah ya (TRUE).

Setelah mendapatkan empat informasi dari tabel *confusion matrix* maka dapat memperoleh nilai *accuracy*. *Accuracy* merupakan suatu tingkat pengukuran dari rasio benar (positif dan negatif) dengan keseluruhan data (Albert dkk., 2020). Nilai *accuracy* pada *confusion matrix* dapat dihitung performana sebagai berikut (Dellia dan Tjahyanto, 2017):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{2.5}$$

Presisi merupakan jumlah data yang diprediksi benar positif dengan keseluruhan hasil yang diprediksi positif (Albert dkk., 2020). Perhitungan presisi dapat dilihat pada persamaan dibawah:

$$Precision = \frac{TP}{TP+FP} \tag{2.6}$$

*Recall* merupakan pengukuran pada data dengan klasifikasi positif yang benar (Albert dkk., 2020). Perhitungan *recall* dapat dihitung menggunakan persamaan dibawah:

$$Recall = \frac{TP}{TP+FN} \tag{2.7}$$

Menurut Rosandi (2016) Kurva ROC (*Receiver Operating Characteristic*) menunjukkan akurasi dan membandingkan klasifikasi secara visual. ROC mengekspresikan *confusion matrix.* ROC merupakan metode untuk menggambarkan dan mengklasifikasikan beberapa kategori berdasarkan kinerjanya. ROC memiliki tingkat nilai diagnose yaitu:

1. *Accuracy* bernilai $0{,}50 - 0{,}60 = failure$
2. *Accuracy* bernilai $0{,}61 - 0{,}70 = poor\ classification$
3. *Accuracy* bernilai $0{,}71 - 0{,}80 = fair\ classifation$
4. *Accuracy* bernilai $0{,}81 - 0{,}90 = good\ classification$
5. *Accuracy* bernilai $0{,}91 - 1{,}00 = excellent\ classification$

Berdasarkan penjelasan diatas, *confusion matrix* merupakan metode yang digunakan untuk menghitung akurasi pada *data mining*. *Confusion matrix* ditampilkan dalam bentuk tabel yang berisi empat informasi yang dapat diolah dan

mendapatkan nilai akurasi, presisi, dan *recall*. Dari akurasi yang didapatkan dapat menyimpulkan apakah hasil penelitian yang dilakukan termasuk baik atau buruk.

## 2.9 *MyPertamina*

Tingginya jumlah penduduk sejalan dengan tingginya jumlah kendaraan yang dibutuhkan, Bahan Bakar Minyak (BBM) merupakan suatu kebutuhan sehari-hari untuk transportasi. Perusahaan yang menyediakan Bahan Bakar Minyak (BBM) adalah PT. Pertamina yang merupakan perusahaan BUMN yang bertugas mengelola penambangan minyak dan gas bumi untuk memenuhi kebutuhan masyarakat Indonesia (Saddam, 2022).

*MyPertamina* adalah program *loyality* dan *e-payment* yang memberikan *user experiences* dari PT. Pertamina (Persero) dengan mudah untuk seluruh pelanggan Pertamina. Dalam layanan *e-money* ini, telah terdaftar dan diawasi oleh Bank Indonesia. *MyPertamina* berfungsi sebagai *cashless payment* (sistem pembayaran non-tunai). Melalui PT. Pertamina, Pemerintah menepakan aturan baru yaitu masyarakat pengendara roda 4 yang ingin mendapatkan BBM jenis Pertalite dan Solar harus terdaftar di sistem *MyPertamina* dan akan dilakukan uji coba mulai 1 Juli 2022 (Kompas.com, 2022). Sehingga publik menyoroti peraturan ini dan memberikan komentar terhadap aplikasi *MyPertamina* di media sosial, salah satunya melalui *Twitter*. Pengguna *Twitter* mengemukakan pendapatnya terhadap suatu produk atau mengomentari suatu masalah melalui *Twitter* (Kurniawan, 2017).