

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan adalah badan hukum yang bertanggung jawab langsung kepada Presiden dan memiliki tugas untuk menyelenggarakan Kesehatan Nasional bagi seluruh rakyat Indonesia terutama untuk Pegawai Negeri Sipil dan TNI/POLRI, Veteran, Perintis Kemerdekaan beserta keluarganya dan Badan Usaha lainnya ataupun rakyat biasa. Dengan perkataan lain BPJS Kesehatan merupakan badan hukum nirlaba yang bertanggung jawab kepada Presiden. Sebagai Badan hukum nirlaba maka BPJS Kesehatan adalah perusahaan yang tidak mencari keuntungan. BPJS Kesehatan merupakan bagian dari sistem jaminan nasional [1].

Beragam konten berformat video dan musik disajikan oleh para content creator di TikTok, dari konten hiburan, edukasi, hingga dakwah dapat dilakukan. Para content creator tentu lebih leluasa dalam membuat konten berdurasi 15-60 detik sesuai ciri khas mereka. Dan bagi para pengguna TikTok, kemudahan akses dan segmentasi video dapat dipilih sesuai minat dan ketertarikan terhadap konten tertentu. Memasuki abad informasi, menjadikan sumber daya informasi dari old media (media lama) menuju ke new media (media baru) [2].

Penelitian ini mengusulkan pembelajaran naïve bayes classifier untuk menyelesaikan penganalisaan sentimen pada BPJS Kesehatan. Banyak masyarakat yang beropini terhadap layanan dan pelayanan dari BPJS Kesehatan. Opini-opini yang telah diberikan masyarakat untuk layanan BPJS Kesehatan memiliki

beragam sifat, ada yang memberikan opini positif, ada yang memberikan opini negatif. Opini-opini dari masyarakat tersebut terkumpul dan dapat digali lagi untuk menghasilkan suatu kumpulan sentimen ini biasanya diperoleh melalui jumlah like dan tweet yang terpadat dalam suatu postingan dari layanan media sosial instansi pemerintah. Sentimen tersebut berguna sebagai evaluasi untuk pemerintahan untuk membuat kebijakan-kebijakan yang dibutuhkan sesuai hasil dari sentimen masyarakat yang sudah di olah dan di kumpulkan [3].

Kesehatan merupakan hak asasi manusia dan sekaligus investasi untuk keberhasilan pembangunan bangsa. Untuk itu perlu diselenggarakan suatu sistem perencanaan pembangunan nasional secara menyeluruh dan berkesinambungan, dengan tujuan meningkatkan kesadaran, kemauan dan kemampuan hidup sehat bagi setiap orang, agar terwujud derajat kesehatan masyarakat yang setinggi-tingginya. Kesinambungan dan keberhasilan pembangunan kesehatan ditentukan oleh tersedianya pedoman penyelenggaraan pembangunan Kesehatan [4].

Pedoman penyelenggaraan pembangunan kesehatan merupakan suatu Program Indonesia Sehat yang dilaksanakan untuk meningkatkan derajat kesehatan masyarakat melalui upaya kesehatan dan pemberdayaan masyarakat yang didukung dengan perlindungan finansial dan pemerataan pelayanan kesehatan. Dalam rangka penyelenggaraan Program Indonesia Sehat salah satunya dengan Pendekatan Keluarga. Proses pendekatan keluarga ditetapkan sebagai penanda status kesehatan keluarga dengan menjadi anggota Jaminan Kesehatan Nasional (JKN) [5].

Guna membantu menyelesaikan masalah Kesehatan diperlukannya sebuah riset atau analisis terkait seberapa besar Fungsi BPJS Kesehatan bagi masyarakat yang dilihat dari tanggapan masyarakat melalui jejaring sosial Tiktok. *Text mining* adalah salah satu teknik yang dapat digunakan untuk melakukan klasifikasi dokumen, *clustering*, *information extraction*, analisis sentimen dan *information retrieval* dimana *text mining* merupakan variasi dari data mining yang berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar [6].

Analisis sentimen atau bisa disebut juga *opinion mining* merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat apakah beropini positif atau negatif. Analisis sentimen dapat sebagai solusi dari masalah yang sudah dipaparkan sebelumnya. Pada penelitian ini algoritma yang akan digunakan adalah NBC (*Naive Bayes Classifier*). Algoritma *Naive Bayes Classifier* dirasa cocok digunakan pada analisis sentimen dikarenakan algoritma ini bertujuan sebagai metode klasifikasi kedalam kategori positif dan negatif [7].

Penelitian sebelumnya tentang *text mining* dan sentiment Analysis menggunakan metode *Naive Bayes* dengan hasil dari 50 data uji yang klasifikasi menghasilkan 12 data bersentimen positif dan sebanyak 38 data sentimen negatif. Dapat diartikan bahwa 50 data uji yang diklasifikasi masuk dalam kategori bersentimen negatif dikarenakan data positif lebih kecil daripada data yang bersentimen negatif. Adapun nilai akurasi algoritma *Naive Bayes Classifier* memberikan nilai akurasi sebesar 84%. Maka pada penelitian ini peneliti akan

menggunakan metode text mining dan sentiment Analysis untuk mengetahui tanggapan dari masyarakat melalui jejaring sosial *Tiktok* tentang dampak kematian BPJS Gratis yang nantinya bisa dijadikan rekomendasi bagi pemerintah untuk mengatasi permasalahan dampak tersebut [8].

Penelitian sebelumnya diteliti oleh Fira Fathonah tentang *text mining* dan sentiment Analysis menggunakan metode *Naïve Bayes* dengan judul Penerapan *Text Mining* Analisis Sentimen Mengenai Vaksin Covid - 19 Menggunakan Metode *Naïve Bayes* dengan Hasil dari 500 data uji yang klasifikasi menghasilkan 180 data bersentimen positif dan sebanyak 320 data sentimen negatif. Dapat diartikan bahwa 500 data uji yang diklasifikasi masuk dalam kategori bersentimen negatif dikarenakan data positif lebih kecil daripada data yang bersentimen negatif. Adapun nilai akurasi algoritma *Naïve Bayes Classifier* memberikan nilai akurasi sebesar 80%. Maka pada penelitian ini peneliti akan menggunakan metode *text mining* dan sentiment Analysis untuk mengetahui tanggapan dari masyarakat melalui jejaring sosial *Tiktok* tentang dampak matinya BPJS gratis yang nantinya bisa dijadikan rekomendasi bagi pemerintah untuk mengatasi permasalahan dampak dari banjir tersebut [8].

Naive Bayes merupakan suatu bentuk klasifikasi data dengan menggunakan metode probabilitas dan statistik. Algoritma *Naive Bayes* dapat diartikan sebagai sebuah metode yang tidak memiliki aturan, *Naive Bayes* menggunakan cabang matematika yang dikenal dengan teori probabilitas untuk mencari peluang terbesar dari kemungkinan klasifikasi dengan cara melihat frekuensi tiap klasifikasi pada data training [8].

Teorema Bayes menerangkan hubungan antara probabilitas terjadinya peristiwa A dengan syarat peristiwa B telah terjadi dan probabilitas terjadinya peristiwa B dengan syarat peristiwa A telah terjadi. *Teorema* ini didasarkan pada prinsip bahwa tambahan informasi dapat memperbaiki probabilitas. *Teorema Bayes* ini bermanfaat untuk mengubah atau memutakhirkan (*mengupdate*) probabilitas yang dihitung dengan tersedianya data dan informasi tambahan [9].

Berdasarkan latar belakang yang telah diuraikan, maka diangkat sebuah judul “Implementasi Metode *Naïve Bayes Classifier* Untuk Analisis Sentimen Publik Di Tiktok Terhadap Penonaktifan Kartu Bpjs Gratis”.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, yang menjadi pokok permasalahan adalah Bagaimana menerapkan implementasi metode *naïve bayes classifier* untuk analisis sentimen publik di tiktok terhadap penonaktifan kartu bpjs gratis?

1.3 Tujuan Penelitian

Menerapkan implementasi metode *naïve bayes classifier* untuk analisis sentimen publik di tiktok terhadap penonaktifan kartu bpjs gratis web.

1.4 Batasan Masalah

Batasan masalah dalam penelitian ini adalah

1. Membuat implementasi metode *naïve bayes classifier* untuk analisis sentimen publik di tiktok terhadap penonaktifan kartu bpjs gratis.
2. Metode yang digunakan dalam penelitian ini adalah *text mining*, *sentiment analysis*, dan *Naïve Bayes Classifier*.

3. Penelitian Menggunakan 1000 data komentar.

1.5 Manfaat Penelitian

Manfaat yang di peroleh dalam penelitian ini adalah:

1. Menambah referensi ilmiah mengenai penerapan metode *naïve bayes classifier* dalam analisis sentiment pada media sosial khususnya Tiktok
2. Menjadi referensi peneliti selanjutnya yang ingin mengembangkan penelitian serupa dengan metode atau objek yang berbeda.

1.6 Sistematika Penulisan

Sistematika penulisan dari skripsi ini terdiri dari lima bagian utama sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah tujuan penelitian, batasan masalah, manfaat penelitian dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab ini berisi teori-teori yang digunakan pada penelitian ini. Teori-teori yang berhubungan dengan landasan dalam pembuatan aplikasi.

BAB 3 METODOLOGI PENELITIAN

Bab ini berisi tahapan–tahapan dalam pengumpulan data, perancangan sistem perumusan masalah dan analisa.

BAB 4 ANALISA DAN PERANCANGAN

Bab ini berisi analisa dan perancangan media pembelajaran yang akan dibangun atau dikembangkan.

BAB 5 IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi implementasi dari analisa dan perancangan dan pengujian pada aplikasi yang berhasil dibangun.

BAB 6 PENUTUP

Bab ini berisi rangkuman dari hasil penelitian yang telah dilakukan dan saran–saran untuk pengembangan aplikasi atau penelitian selanjutnya.

BAB 2

LANDASAN TEORI

2.1 Analisis Sentimen

Sentiment analysis (analisis sentimen) atau sering disebut juga dengan *opinion mining* (penambangan opini) adalah studi komputasi untuk mengenali dan mengekspresikan opini, sentimen, evaluasi, sikap, emosi, subjektivitas, penilaian atau pandangan yang terdapat dalam suatu teks [10].

Analisis sentimen atau *opinion mining* merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat *opini*. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung berpandangan atau beropini negatif, positif atau netral. Salah satu contoh penggunaan analisis sentimen dalam dunia nyata adalah identifikasi kecenderungan pasar dan *opini* pasar terhadap suatu objek barang. Besarnya pengaruh dan manfaat dari analisis sentimen menyebabkan penelitian dan aplikasi berbasis analisis sentimen berkembang pesat. Bahkan di Amerika terdapat sekitar 20-30 perusahaan yang memfokuskan pada layanan analisis sentimen [11].

2.2 Publik

Istilah publik dalam pengertian sehari-hari di Indonesia sering dipahami sebagai negara atau umum, hal ini biasa dijumpai dalam pola Bahasa Indonesia yang menterjemahkan publik seperti pada istilah public administration yang diterjemahkan sebagai administrasi negara. Kata publik sebenarnya sudah dite-

rima menjadi Bahasa Indonesia baku menjadi publik yang berarti umum, atau orang banyak [12].

Istilah publik mengalami perubahan seiring dengan bentuk-bentuk perkembangan masyarakat itu sendiri, dari masyarakat feodal hingga masyarakat liberal-demokratik. Dalam masyarakat Yunani, publik diartikan secara sederhana sebagai suatu kehidupan bersama. Pembagian struktur masyarakat dari Aristoteles (384-322SM) – kelas bawah (budak), prajurit, dan kelas terdidik (bangsawan) – misalnya, tidak serta merta membebaskan mereka untuk dapat saling hidup bersama dalam suatu hak yang sama di dalam ruang publik. Pembagian kelas sosial di Athena juga membedakan dengan tegas antara warga negara dengan warga asing (*metic*). Warga asing memiliki hak yang sama dalam berbagai sisi kehidupan dengan warga negara asli Athena, namun mereka tidak diberi hak politik walaupun sudah menetap dalam kurun waktu yang lama. Begitu pula warga negara tidak menyiratkan seluruh warga yang lahir di Athena memiliki hak politik. Wanita, pedagang, maupun petani tidak dianggap sebagai warga negara yang memiliki hak politik dan urusan publik [13].

2.3 Sosial Media Tiktok

Tik tok merupakan sebuah aplikasi yang dapat memberikan efek spesial yang unik dan menarik yang bisa digunakan oleh para pengguna aplikasi ini dengan mudah untuk membuat video pendek yang keren dan dapat menarik perhatian banyak orang yang melihatnya. Aplikasi tik tok memiliki banyak konten - konten video yang sangat mudah untuk membuatnya. Dengan melihat, menghafal, menirukan dapat membuat video bebas, dapat memberikan contoh-

contoh video pendek yang menarik serta menjadi contoh pengguna lainnya untuk menyalurkan ide-ide kreatifnya agar memiliki viewers yang tinggi.

Aplikasi tik tok memiliki beberapa fitur menarik seperti penambahan efek video, suara, stiker, berbagi video, dan tentunya mampu untuk bertukar pesan kepada pengguna lainnya. Aplikasi tik tok berisikan berbagai macam jenis konten yang bervariasi, dimulai dari konten edukasi yang mendidik, memasak, menggambar, menjahit dan banyak hal-hal positif di dalamnya. Selain banyak hal-hal baik di dalamnya, tik tok juga menyajikan cukup banyak konten yang sebaiknya tidak dilihat oleh anak-anak dibawah umur. Penggunaan aplikasi yang minim pengawasan orang tua, bisa menyebabkan anak-anak tersebut dapat mencontoh perbuatan - perbuatan yang mungkin belum tepat dilakukan untuk seumurannya mereka [14].

Aplikasi tik tok adalah media sosial yang banyak digunakan pada anak-anak saat ini. Media sosial tiktok adalah sebuah jaringan sosial dan platform video musik yang berdurasi dengan jangka waktu pendek. Salah satu contoh adalah akun tiktok @biawak97 dan @khanayaaazzahra1, yang mana Sakha Ibadil Qiram selaku pemilik akun tiktok @biawak97 yang merupakan informan utama dan informan pendukung ialah Khanaya Azzahra Salsabilla dan Naura. Informan utama selaku konten kreator tiktok yang menggemari aplikasi tiktok sangat rajin mengunggah video di akun tiktok yang ia miliki. Namun, Khanaya Azzahra Salsabila dan Naura merupakan pengguna aktif aplikasi tiktok serta konten kreator (newbie), dimana newbie yang berarti orang baru yang menggemari hal-hal yang ia pelajari [15].

2.4 Badan Penyelenggara Jaminan Sosial (BPJS)

BPJS kesehatan yaitu asuransi jaminan kesehatan yang menjamin segala biaya pengobatan, pelayanan dan penggunaan fasilitas kesehatan untuk seluruh masyarakat Indonesia [16].

BPJS Kesehatan (Badan Penyelenggara Jaminan Sosial) adalah organisasi yang dibiayai oleh pemerintah yang dibentuk oleh pemerintah untuk menyelenggarakan jaminan kesehatan kepada seluruh rakyat Indonesia, terutama pegawai negeri sipil, pensiunan pegawai negeri sipil, militer/politisi, veteran, penggerak diri sendiri dan keluarganya. lembaga pemerintah yang ditunjuk. dan entitas lain atau anggota masyarakat [17].

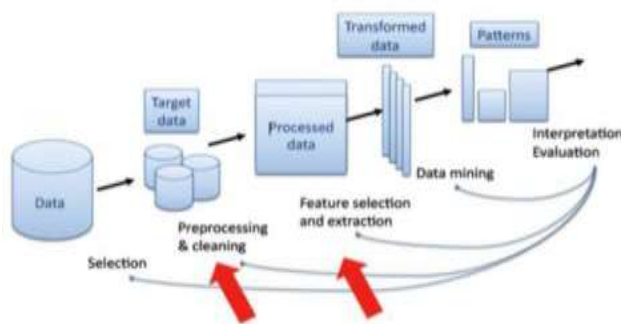
2.5 *Text Mining*

Text mining adalah algoritma komputer yang bertujuan untuk menggali suatu informasi atau mengekstrak informasi baru yang sebelumnya belum diketahui ketika belum dilakukan proses *text mining* yang berasal dari *teks* yang tidak terstruktur dalam jumlah besar. Atau dalam pengertian lain *text mining* adalah kegiatan menambang dari suatu data berupa *text* untuk menemukan sebuah inti dari suatu dokumen atau kata-kata tertentu apa yang bisa mewakili dokumen dan menganalisa keterhubungan antar dokumen yang diperiksa oleh komputer [18].

Text mining adalah satu langkah dari *analisis teks* yang dilakukan secara otomatis oleh komputer untuk menggali informasi yang berkualitas dari suatu rangkaian *teks* yang terangkum dalam sebuah dokumen. Prosedur utama dalam metode ini terkait dengan menemukan kata-kata yang dapat mewakili isi dari

dokumen untuk selanjutnya dilakukan analisis keterhubungan antar dokumen dengan menggunakan metode statistik tertentu seperti *analisis* kelompok, *klasifikasi* dan *asosiasi* [19].

Tahapan dalam *text mining* secara umum dapat dilihat pada gamabar dibawah ini.



Gambar 2.1 Tahapan *Text Mining*

2.6 Data Training

Data *training* adalah data yang digunakan sebagai acuan untuk membangun model *klasifikasi*. Pada skenario pengujian ini data *training* sudah dipersiapkan sebelumnya dan disimpan ke dalam dataset. Hasil data *training* yang didapat akan diuji dengan data baru yang masih belum diketahui *classtype* kategorinya [20].

Data *training* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Istilah data *training* memiliki hakikat sebagai disiplin ilmu yang tujuan utamanya adalah untuk menemukan,

menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki. Data *training*, sering juga disebut sebagai *Knowledge Discovery in Database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [21].

2.7 *Preprocessing*

preprocessing untuk data yang akan di gunakan dalam proses analisa sentimen, dimana data yang kita proses akan kita ambil informasi yang terkandung didalamnya dalam hal sentimen penulisnya yaitu negatif, positif atau netral. Guna memudahkan dalam mengelola data maka data perlu kita berikan analisa sentimen secara manual dengan membaca maksud dari kalimat yang ada dalam sentimen tersebut, sehingga dapat diberikan penilaian bahwa sentimen tersebut merupakan sentimen negatif, positif atau netral [22].

2.8 *Tokenize*

Tokenizing atau tokenisasi adalah proses yang paling awal dalam melakukan *text mining*. Dalam proses ini, input stream yang didapat dari file *text* akan dipecah-pecah menjadi bagian bagian yang lebih kecil. Sebagai contoh pemecahan kalimat menjadi kata-kata (*tokens*) [23].

2.9 *Stop Word*

Langkah yang selanjutnya adalah *Stop Word Removal*. *Stop Word* sendiri didefinisikan sebagai sekumpulan kata yang tidak berhubungan (*irrelevant*) dengan subyek utama yang dimaksud, meskipun kata tersebut sering muncul

didalam data yang digunakan. Kata-kata yang dimaksud biasanya adalah jenis kata sambung, imbuhan, dan lain sebagainya [24].

2.10 Normalisasi

normalisasi adalah memastikan bahwa tidak terjadi Partial Dependency, yaitu kondisi ketika salah satu atribut dalam satu entitas bergantung pada sebagian kunci primer yang ada di entitas tersebut. Dalam penelitian ini, bentuk kedua dari normalisasi ini dapat dilihat pada penjabaran [24].

2.11 Filtering

Filtering adalah tahap mengambil kata penting dari hasil proses token. Bisa menggunakan algoritma *stoplist* atau *wordlist*. *Filtering* dapat juga diartikan sebagai proses mengambil kata – kata penting dari hasil proses token atau penghapusan stopwords [25].

2.12 Stemming

Stemming digunakan untuk memeriksa relevannya beberapa dokumen dibutuhkan proses stemming. Stemming adalah di mana proses pemetaan kata pada kalimat berimbuhan menjadi kata asli (tanpa kata imbuhan awalan,akhiran,sisipan,kombinasi) yang di jalankan algoritma tertentu [26].

2.13 Naive Bayes

Secara umum algoritma *naive bayes* untuk membangun sebuah hasil kesimpulan adalah sebagai berikut “*Naive Bayes Classifier* merupakan salah satu algoritma dalam teknik data mining yang menerapkan teori bayes dalam klasifikasi. Teorema keputusan *bayes* adalah pendekatan statistik yang

fundamental dalam pengenalan pola (*Pattern Recognition*), *Naive bayes* didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output [27].

Naive Bayes merupakan suatu bentuk *klasifikasi* data dengan menggunakan metode probabilitas dan statistik. Algoritma *Naive Bayes* dapat diartikan sebagai sebuah metode yang tidak memiliki aturan, *Naive Bayes* menggunakan cabang matematika yang dikenal dengan teori probabilitas untuk mencari peluang terbesar dari kemungkinan *klasifikasi* dengan cara melihat *frekuensi* tiap *klasifikasi* pada data *training* [28].

Naive bayes merupakan teknik prediksi berbasis probabilitas sederhana yang berdasar pada penerapan aturan *bayes* dengan asumsi ketidaktergantungan yang kuat. Selain itu *naive bayes* juga dapat menganalisis variabel-variabel yang paling mempengaruhinya dalam bentuk peluang [29].

1. Hitung probabilitas *prior* setiap kategori, yang menjadi kategori ada 3 kategori yaitu positif, netral, negatif.

$$P_{\frac{positif}{netral}} = \frac{P_{\frac{positif}{netral}}}{|c|} \dots (4)$$

2. Menghitung probabilitas *likelihood* setiap term dari semua dokumen, Banyaknya term tergantung pada hasil praproses data.

$$P(w|\frac{positif}{netral}/\frac{negatif}{negatif}) = \frac{(nk(\frac{positif}{netral})+1)}{(n.\frac{negatif}{negatif}+|r|)} \dots (5)$$

3. Proses klasifikasi data uji yang termasuk ke dalam data training dengan memilih 2 kata data uji. Nilai yang lebih tinggi merupakan kelas baru dari data tersebut.

$$p\left(\frac{\text{positif}}{\text{netral}} \middle| d\right) = P\left(\frac{\text{positif}}{\text{netral}}\right) \times np(w|\text{positif/netral/negatif}) \dots (6)$$

Berikut ini adalah perhitungan manual penyelesaian menggunakan metode Naïve Bayes.

2.14 *Term Frequency - Inverse Document Frequency (Tf-Idf)*

Algoritma Term Frequency – Inverse Document Frequency (TF-IDF) merupakan salah satu algoritma yang sangat populer untuk melakukan klasifikasi teks yang dapat digunakan untuk menganalisis hubungan antara sebuah kalimat dengan sekumpulan dokumen. Metode TF-IDF merupakan metode untuk menentukan frekuensi relatif dari setiap kata-kata atau token-token yang mana setiap kata tersebut akan diberikan pembobotan berupa nilai berdasarkan penting atau tidaknya suatu kata dalam dokumen berdasarkan jumlah kemunculan kata dalam suatu dokumen dan mengukur kata-kata tersebut terhadap keseluruhan dokumen yang ada. Metode TF-IDF ini merupakan gabungan dari dua konsep untuk penghitungan bobot yaitu Tf (term frequency) dan idf (inverse document frequency) [30].

$$tf - idf(t, d) = tf(t, d) * idf(t)$$

Term Frequency (TF) menyatakan jumlah kata t yang muncul dalam sebuah dokumen d . Pendekatan paling sederhana dari konsep ini adalah

dengan menyatakan bobot suatu kata t sebagai jumlah kemunculannya pada dokumen d .

$$tf(t, d) = \text{jumlah } t \text{ dalam } d / \text{jumlah kata pada } d$$

Document Frequency menyatakan jumlah dokumen yang memuat kata t .

$$df(t) = \text{jumlah dokumen yang memuat term } t$$

Inverse Document Frequency (IDF) untuk menguji seberapa relevan kata tersebut. Tujuannya adalah untuk menemukan proporsi dokumen dalam korpus yang mengandung istilah atau term tersebut. Kata yang unik akan mendapatkan nilai kepentingan yang lebih tinggi daripada kata yang umum.

$$idf(t) = \log(N/df(t))$$

Keterangan:

$idf(t)$ = inverse document frequency

N = jumlah keseluruhan dokumen

2.15 Aplikasi

Mengemukakan bahwa Aplikasi adalah *software* yang dibuat oleh suatu perusahaan komputer untuk mengerjakan tugas-tugas tertentu, misalnya *Ms.World*, *Ms.Excel* [31].

Aplikasi adalah program siap pakai yang dapat digunakan untuk menjalankan perintah-perintah dari pengguna aplikasi tersebut dengan tujuan mendapatkan hasil yang lebih akurat sesuai dengan tujuan pembuatan aplikasi tersebut, aplikasi mempunyai arti yaitu pemecahan masalah yang menggunakan salah satu teknik pemrosesan data aplikasi yang biasanya berpacu pada sebuah

komputansi yang diinginkan atau diharapkan maupun pemrosesan data yang diharapkan. Pengertian aplikasi secara umum adalah alat terapan yang difungsikan secara khusus dan terpadu sesuai kemampuan yang dimilikinya, aplikasi merupakan suatu perangkat komputer yang siap pakai bagi *user* [32].

Aplikasi dapat diartikan sebagai suatu program berbentuk perangkat lunak yang berjalan pada suatu sistem tertentu yang berguna untuk membantu berbagai kegiatan yang dilakukan oleh manusia [33].

2.16 UML(Unifield Modeling Language)

Unifield Modeling Language merupakan salah satu metode pemodelan visual yang digunakan dalam perancangan dan pembuatan sebuah *software* yang berorientasikan pada objek [34].

Pada perkembangan teknik pemograman berorientasi objek, muncullah sebuah standarisasi bahasa pemodelan untuk pembangunan perangkat lunak yang dibangun dengan menggunakan teknik pemograman berorientasi objek yaitu *Unified Modeling Language (UML)* [35].

UML merupakan bahasa visual untuk pemodelan dan komunikasi mengenai sebuah sistem dengan menggunakan diagram dan teks-teks pendukung [36].

2.16.1 Use Case Diagram

Use Case Diagram merupakan diagram yang harus dibuat pertama kali saat pemodelan perangkat lunak berorientasi objek dilakukan [37].

Use case atau *diagram use case* merupakan pemodelan untuk kelakuan (*behavior*) sistem informasi yang akan dibuat [38].

Use Case merupakan langkah pertama dalam memodelkan sebuah sistem. *Use Case* merupakan pemodelan untuk kebutuhan sebuah sistem fungsional, setiap *Use Case* digambarkan sebagai kunci dari suatu skenario yang dilakukan oleh aktor dan diringkas dalam sebuah batas sistem, setiap *Use Case* dihubungkan dengan sebuah garis notasi [38].

2.16.2 Class Diagram

Class Diagram dibuat setelah diagram *use case* dibuat terlebih dahulu. Pada diagram ini harus menjelaskan hubungan apa saja yang terjadi antara suatu objek dengan objek lainnya sehingga terbentuklah suatu sistem aplikasi [37].

Diagram kelas atau *class diagram* menggambarkan struktur sistem dari sini pendefinisian kelas-kelas yang akan dibuat untuk membangun sistem [29].

Class diagram adalah salah satu pemodelan yang cukup penting dalam *UML*, fungsinya adalah untuk membuat sebuah *logical models* dari sebuah sistem [28].

2.16.3 Activity Diagram

Activity Diagram menggambarkan *workflow* (aliran kerja) atau aktifitas dari sebuah sistem atau proses bisnis [37].

Diagram aktivitas atau *activity diagram* menggambarkan *workflow* (aliran kerja) atau aktivitas dari sebuah sistem atau proses bisnis atau menu yang ada pada perangkat lunak [29].

Activity diagram merupakan pemodelan yang menggambarkan sebuah sistem kerja dari sebuah objek atau sebuah sistem, sebuah *activity diagram*

digambarkan dengan sebuah alur secara terstruktur proses kerja dari *use case* yang sedang diproses dari titik awal sampai titik akhir, setiap aktivitas digambarkan dengan notasi-notasi sesuai fungsinya [28].

2.16.4 *Sequence Diagram*

Sequence diagram menggambarkan kelakuan objek pada *use case* dengan mendeskripsikan waktu hidup objek dan *message* yang dikirimkan dan diterima antar objek. Oleh karena itu untuk menggambar diagram sekuen maka harus diketahui objek-objek yang terlibat dalam sebuah *use case* beserta metode-metode yang dimiliki kelas yang diinstansiasi menjadi objek itu”. Membuat diagram sekuen juga dibutuhkan untuk melihat skenario yang ada pada *use case* [26].

Sequence diagram merupakan *UML* yang menggambarkan interaksi antar objek di dalam dan disekitar sistem, termasuk pengguna, *display*, dan sebagainya berupa *message* yang digambarkan terhadap waktu [28].

Sequence diagram merupakan interaksi object dalam sistem, dapat dilihat pada gambar dibawah ini object-object yang ada di sistem [29].

2.17 *PHP (Hypertext Preprocessor)*

Hypertext Markup Language (HTML) adalah bahasa yang digunakan untuk menulis halaman *web*. *HTML* merupakan pengembangan dari standar. Pemformatan dokumen teks yaitu *Standard Generalized Markup Language (SGML)*. *HTML* sebenarnya adalah dokumen *ASCII* atau teks biasa, yang dirancang untuk tidak tergantung pada suatu sistem operasi tertentu [27].

PHP merupakan singkatan dari *Hypertext Preprocessor*. Ia merupakan bahasa berbentuk skrip yang ditempatkan dalam *server* dan diproses di *server*.

Hasilnyalah yang dikirimkan ke klien, tempat pemakai menggunakan *browser* [26].

PHP adalah bahasa pemrograman untuk dijalankan melalui halaman *web*, umumnya digunakan untuk mengolah informasi di *internet* [39].

2.18 *MySQL(My Structured Query Language)*

MySQL adalah sebuah perangkat lunak sistem manajemen basis data *SQL* (*database management system*) atau *DBMS* yang *multithread*, *multi-user*, dengan sekitar 6 juta instalasi di seluruh dunia. *MySQL AB* membuat *MySQL* tersedia sebagai perangkat lunak gratis dibawah lisensi *GNU General Public License (GPL)*, tetapi mereka juga menjual dibawah lisensi komersial untuk kasus-kasus dimana penggunaannya tidak cocok dengan penggunaan *GPL* [40].

MySQL merupakan singkatan dari *My Structured Query Language* adalah sebuah program *database server* yang mampu menerima dan mengirimkan data dengan sangat cepat serta multi *user*, sedangkan *SQL* adalah bahasa perintah dasar yang ada didalamnya [41].

2.19 *Flowchart*

Flowchart merupakan gambaran secara grafik dari langkah-langkah dan urutan-urutan prosedur dari suatu program yang dibuat. *Flowchart* menggambarkan urutan logika dari suatu prosedur pemecahan masalah, sehingga *flowchart* merupakan langkah-langkah penyelesaian masalah yang dituliskan dalam simbol-simbol tertentu [42].

Flowchart adalah cara untuk menjelaskan tahap-tahap pemecahan masalah dengan merepresentasikan simbol-simbol tertentu yang mudah dipahami, mudah digunakan dan standar [43].

Tujuan penggunaan *flowchart* adalah untuk menggambarkan suatu tahapan penyelesaian masalah secara sederhana, terurai dan rapi dengan menggunakan simbol-simbol yang standar yang dapat di mengerti oleh programmer [44].

2.20 Penelitian Terkait

Tabel 2.1 Penelitian Terkait

No	Penulisan dan Tahun	Judul	Metode	Hasil
1.	(Syarifuddin, M. 2020).	Analisis Sentimen Opini Publik Mengenai Covid-19 Pada <i>Tiktok</i> Menggunakan Metode <i>Naive Bayes</i> Dan <i>Knn</i>	<i>Naive Bayes</i>	Berdasarkan hasil penelitian, dapat ditarik beberapa kesimpulan diantaranya adalah COVID19 merupakan penyakit yang tidak sedikit meresahkan masyarakat umum, apalagi ditambah dengan kicauan di media social tentang COVID-19 yang semakin merabak, salah satunya adalah <i>tiktok</i> yang menambah masyarakat ketakutan.
2.	(Samsir, S., Ambiyar,	Analisis Sentimen	<i>Naive Bayes</i>	Penelitian ini menerapkan algoritma <i>Naive Bayes</i> untuk klasifikasi sentimen. Analisis sentimen dalam study ini menggunakan data <i>Tiktok</i>

	A., Verawardina , U.,Edi, F., & Watrianthos, R. 2021).	Pembelajaran Daring Pada <i>Tiktok</i> di Masa Pandemi COVID-19 Menggunakan Metode <i>Naive Bayes</i> .		dengan kata kunci 'pembelajaran daring', 'kuliah', 'belajar', 'online', 'daring', dan tagar #BelajarDariRumah yang difilter dengan kata kunci 'online' dan 'rumah' pada twit- twit dalam bahasa Indonesia pada minggu pertama November 2020. Penelitian pada periode tersebut menunjukkan 30% sentimen positif, 69% sentimen negatif, dan 1% netral.
3.	(Nugroho, A. 2018).	Analisis Sentimen Pada Media Sosial <i>Tiktok</i> Menggunakan <i>Naive Bayes Classifier</i> Dengan Ekstraksi Fitur <i>N-Gram</i> .	<i>Naive Bayes</i>	Berdasarkan dari hasil penelitian terhadap model algoritma <i>naïve bayes</i> yang di kombinasikan dengan ekstraksi fitur n-gram terhadap analisa persebaran konten terhadap isu kenaikan tarif dasar listrik, maka dapat ditarik kesimpulan bahwa algoritma <i>naïve bayes</i> mampu mencapai tingkat akurasi 89,67%, dan pengaruh ekstraksi fitur n-gram yang diterapkan dapat meningkatkan nilai akurasi pada algoritma <i>naïve bayes</i> sekitar 2,33%, yaitu menjadi 92,00% dalam pengklasifikasian data komen.
4.	Mahardhika, Y. S., & Zuliarso, E. (2018).	Analisis Sentimen Terhadap Pemerintahan Joko Widodo	<i>Naive Bayes</i>	Dengan hasil akurasi yang cukup tinggi yaitu 97% maka metode <i>Naive Bayes Classifier</i> dapat digunakan untuk melakukan klasifikasi tweet dengan sentimen negatif dan positif secara otomatis.

		Pada Media Sosial <i>Tiktok</i> Menggunakan Algoritma <i>Naives Bayes Classifier</i>		
5.	Pintoko, B. M., & Lhaksmana, K. M. (2018).	Analisis Sentimen Jasa Transportasi Online pada <i>Tiktok</i> Menggunakan Metode <i>Naive Bayes Classifier</i> .	<i>Naive Bayes</i>	Berdasarkan hasil analisis yang dilakukan terkait sentimen terhadap jasa transportasi online dengan menggunakan metode <i>Naive Bayes</i> untuk menentukan kelas prediksi, didapatkan hasil sentimen positif sebesar 88.60% dan sentimen negatif sebesar 11.40% dengan akurasi sebesar 86.80%. Hasil menunjukkan tingkat sentimen positif lebih besar dibandingkan dengan tingkat sentimen negatif. Dengan hasil penelitian ini bisa dilihat bahwa banyak masyarakat yang merasa puas dengan layanan jasa transportasi online di Indonesia berdasarkan data dari <i>Tiktok</i> .
6.	Sudiantoro, A. V., & Zuliarso, E. (2018).	Analisis Sentimen <i>Tiktok</i> Menggunakan Text Mining Dengan	<i>Naive Bayes</i>	Algoritma <i>Naive Bayes Classifier</i> sangat efektif untuk digunakan sebagai proses klasifikasi tweet yang dibutuhkan dalam sistem analisis sentiment ini dimana

		Algoritma <i>Naive Bayes Classifier</i>		nilai yang di dapatkan dalam pengujian sampai 84%.
7.	Syarifuddin, M. (2020).	Analisis Sentimen Opini Publik Terhadap Efek Psbb Pada Tiktok Dengan Algoritma <i>Decision Tree-Knn-Naive Bayes</i> .	<i>Naive Bayes</i>	Dalam studi ini, upaya dilakukan untuk mengklasifikasikan sentimen analisis dari tweet di tiktok tentang efek PSBB bagi masyarakat di Indonesia. Untuk merangkum pandangan masyarakat ini teknik penambahan teks digunakan, dan penggalian data menggunakan tiga algoritma yang berbeda yaitu Decision Tree, K-NN, dan <i>Naive Bayes</i> . Tiga algoritma memprediksi label pada dataset. Dari hasil pengujian secara keseluruhan algoritma <i>Decision Tree</i> memiliki nilai tertinggi dari nilai <i>accuracy</i> dan <i>recall</i> .
8.	Naraswati, N. P. G., Nooraeni, R., Rosmilda, D. C., Desinta, D., Khairi, F., & Damaiyanti, R. (2021).	Analisis Sentimen Publik dari Tiktok Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan <i>Naive Bayes</i>	<i>Naive Bayes</i>	Berdasarkan hasil pengujian yang telah dilakukan, dapat disimpulkan bahwa masyarakat lebih banyak memberikan sentimen negatif terhadap kebijakan penanganan COVID-19 khususnya PSBB, wajib masker, dan jam malam yaitu sebesar 72,72 persen. Pada sentimen positif, tiga kata dengan frekuensi kemunculan terbanyak yaitu demo, jakarta, dan kerja.

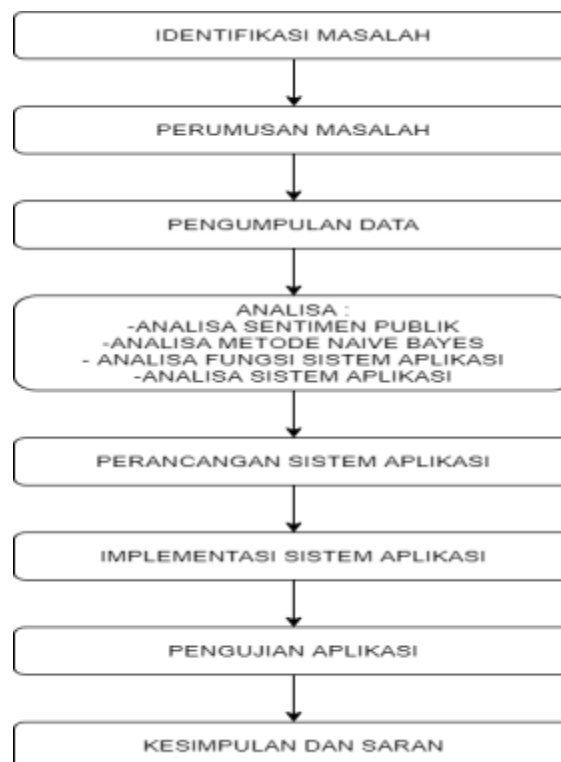
		<i>Classification</i>		
9.	Ruhyana, N. (2019).	Analisis Sentimen Terhadap Penerapan Sistem Plat Nomor Ganjil/Genap Pada <i>Tiktok</i> Dengan Metode Klasifikasi <i>Naive Bayes</i> .	<i>Naive Bayes</i>	Dari hasil penelitian untuk mencari sentiment dalam hal penerapan sistem ganjil genap untuk tanggapan pada pengguna tiktok dan dilakukan analisa sentiment dengan label positif dan negatif dengan metode klasifikasi <i>naive bayes</i> pengujian dilakukan dengan data training dan data testing yang di ambil dari <i>tiktok</i> menggunakan software <i>rapid miner</i> sehingga memudahkan memberikan hasil dari penelitian tersebut dengan hasil <i>accuracy</i> 86,67%, presisi 71,43% dan recall 80,00%.
10.	Ramadhan, D. A., & Setiawan, E. B. (2019).	Analisis Sentimen Program Acara Di Sctv Pada <i>Tiktok</i>	<i>Naive Bayes</i>	Berdasarkan hasil penelitian untuk Analisis Sentimen Program Acara di SCTV pada <i>Tiktok</i> menggunakan Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> yang telah dilakukan mendapatkan Labelling terbaik didapatkan dengan cara <i>Labelling</i> Otomatis sedangkan akurasi terbaik dari

		Menggunakan Metode <i>Naive Bayes</i> Dan <i>Support Vector Machine</i>		Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> adalah Metode <i>Support Vector Machine</i> dengan Seluruh Program Acara didapatkan hasil akurasi 88,57% berikut perkategori yang diperoleh Berita didapatkan hasil akurasi 79,81%, ISSN : 2355-9365 e- <i>Proceeding of Engineering</i> : Vol.6, No.2 Agustus 2019 Page 9742 <i>Entertainment</i> didapatkan hasil akurasi 89,80%, Sinetron didapatkan hasil akurasi 73,68% dan FTV didapatkan hasil akurasi 87,74%.
--	--	---	--	---

BAB 3

METODOLOGI PENELITIAN

Penelitian ini dilakukan dengan melaksanakan tahapan demi tahapan yang berhubungan. Tahapan-tahapan tersebut dijabarkan dalam metode penelitian. Metode penelitian diuraikan kedalam bentuk skema yang jelas, teratur, dan sistematis. Berikut tahapan-tahapan penelitian dapat dilihat pada gambar 3.1



Gambar 3.1 Metodologi Penelitian

Penjelasan dari tahapan-tahapan penelitian pada gambar 3.1 dapat dilihat pada penjelasan dibawah ini:

3.1 Identifikasi Masalah

Identifikasi masalah merupakan tahapan awal yang dilakukan dalam penelitian ini bertujuan untuk mengamati masalah yang terjadi yaitu mengklasifikasikan komentar pengguna *tiktok* terhadap permasalahan BPJS yang sering terjadi. Langkah untuk dapat memahami masalah yang telah ditentukan ruang lingkup masalah atau batasannya. Dengan menganalisa masalah yang telah ditemukan tersebut, maka diharapkan masalahnya dapat dipahami dengan baik. Maka dari itu diperlukan sebuah sistem baru, yaitu dengan membuat suatu aplikasi yang mampu mengklasifikasikan komentar pengguna *tiktok* pada masalah dihapusannya sebagian pengguna BPJS gratis.

3.2 Perumusan Masalah

Berdasarkan hasil dari tahapan identifikasi masalah sebelumnya, maka tahapan selanjutnya adalah perumusan masalah. Pada tahapan perumusan masalah akan dirumuskan masalah yang dianggap sebagai penelitian dalam tugas akhir ini. Permasalahan-permasalahan yang menjadi rumusan masalah dalam penelitian ini didapatkan dari penelitian, terkait data pengamatan pendahuluan sebelumnya. Solusi yang didapatkan pada tahapan perumusan masalah yang akan menjadi judul penelitian tugas akhir ini yaitu “ Aplikasi *analisis sentiment public Tiktok* pada masalah non aktif kartu BPJS gratis menggunakan metode *Naive Bayes*”. Penggunaan metode *Naïve Bayes* dalam Aplikasi ini didasari oleh keunggulan metode *Naïve Bayes* yaitu *Naive Bayes* merupakan suatu bentuk *klasifikasi data* dengan menggunakan metode *probabilitas* dan *statistik*.

3.3 Pengumpulan Data

Pengumpulan data adalah tahapan-tahapan yang bertujuan dalam memperoleh data-data informasi yang berhubungan dengan penelitian tugas akhir ini. Pada tahapan pengumpulan data ini juga berguna untuk mengumpulkan semua kebutuhan data yang akan diproses nantinya menggunakan metode *Naïve Bayes*. Dalam pengumpulan data ini penulis mendapatkannya dari komentar pengguna *tiktok* terhadap dihapusnya BPJS gratis yang sudah banyak membantu masyarakat.

3.4 Analisa

Tahapan selanjutnya adalah melakukan analisa metode sistem dari penelitian Skripsi ini. Adapun tahapan analisa dalam penelitian Skripsi ini adalah sebagai berikut:

3.4.1 Analisa Metode *Naïve Bayes*

Tahapan ini yaitu menganalisa cara kerja dari metode *Naïve Bayes*, dengan menganalisa metode ini, nantinya bisa di terapkan pada aplikasi yang akan di bangun, karena permasalahan yang ditemukan oleh penulis dalam penelitian ini yaitu dapat mengetahui bagaimana pendapat public terhadap dampak dihapusnya BPJS gratis melalui media *social tiktok*

3.4.2 Analisa Fungsi Sistem Aplikasi

Setelah melakukan tahapan analisis terhadap metode *Naïve Bayes*, maka selanjutnya adalah analisa fungsional sistem yang akan dibangun. Adapun tahapan-tahapan analisis fungsional yaitu dalam pembuatan *flowchart*.

3.4.3 Analisa Sistem Aplikasi

Setelah menganalisa sistem lama, maka tahapan selanjutnya dengan menganalisa sistem yang baru. Analisa dalam pembuatan sistem ini menggunakan metode *Naïve Bayes* serta penggunaan *Unified Modelling Language* untuk menganalisa kebutuhan sistem.

3.5 Perancangan Sistem Aplikasi

Setelah tahapan analisa selesai dilakukan, maka tahapan selanjutnya adalah perancangan sistem. Tahapan perancangan sistem terdiri dari :

3.5.1 Perancangan Basis Data

Setelah dilakukannya analisa sistem yang akan dibuat, maka tahap berikutnya ialah analisa dan perancangan basis data yang kita lakukan untuk melengkapi komponen dalam pembuatan sistem.

3.5.2 Perancangan Struktur Menu

Rancangan struktur menu ini kita perlu untuk memberikan gambaran terhadap menu-menu atau fitur pada sistem yang akan dibuat

3.5.3 Perancangan Antar Muka (*Interface*)

Dalam mempermudah komunikasi antara sistem dengan pengguna, maka perlu di rancang antar muka (*interface*). Dalam perancangan *interface* hal terpenting yang harus dilakukan ialah bagaimana menciptakan tampilan yang baik dan mudah dimengerti oleh pengguna.

3.6 Implementasi Sistem Aplikasi

Beberapa komponen pendukung yang memiliki peran yang sangat penting dalam implementasi sistem diantaranya adalah perangkat keras (*hardware*) dan perangkat lunak (*software*). Adapun spesifikasi dari perangkat keras (*hardware*) dan perangkat lunak (*software*) yang di gunakan sebagai berikut :

1. Perangkat keras (*hardware*), antara lain :

Prosesor : Intel Core i5-5300U

Memory (RAM) : 4 GB RAM

System type : 64-bit Operating System

Harddisk : SSD 100 GB

2. Perangkat Lunak (*Software*), antara lain :

Sistem operasi : Windows 10

3.7 Pengujian Aplikasi

Pengujian sistem ini dilakukan dengan cara menggunakan *Black Box*. Dalam Pengujian *Black Box* ini berfokus pada perangkat lunak untuk mendapatkan hasil yang baik. Apabila terjadi *error* atau tidak sesuai dengan tujuan yang diharapkan maka dilakukan penganalisaan sistem kembali hingga tidak ditemukan *error* ,

3.8 Kesimpulan dan Saran

Tahapan terakhir adalah menarik kesimpulan dari hasil penelitian terkait penggunaan metode *Naive Bayes* pada aplikasi ini, Pada tahapan ini juga berisikan saran peneliti bagi pembaca untuk melakukan pengembangan terhadap penelitian ini kedepannya.