# BAB I PENDAHULUAN

## Latar Belakang

Penyakit virus *corona* (COVID-19) menjadi penyakit pandemi hampir merebak di seluruh dunia, bahkan sudah masuk di Indonesia. Berdasarkan Pada tanggal 2 Maret 2020, Indonesia melaporkan pertama kali adanya dua warga Depok, Jawa Barat yang terjangkit penyakit virus *corona* (Ihsanuddin, 2020). Virus *corona* semakin merebak dan menimbulkan masalah dalam bidang Kesehatan, juga menimbulkan dampak di bidang Ekonomi, Pendidikan, Wisata dan lain sebagainya. Indonesia sedang mengalami pandemi dan sedang berjuang menurunkan angka positif COVID-19 dengan melakukan vaksinasi.

Vaksin tidak hanya melindungi mereka yang divaksinasi, tetapi juga masyarakat luas ikut terlindungi dengan memperlambat laju penyebaran virus dalam populasi. Pemerintah Indonesia saat ini juga turut aktif dalam melakukan vaksinasi yang diberikan kepada masyarakat. Menurut data vaksinasi COVID-19 Nasional kementrian Kesehatan per tanggal 08 Oktober 2021, menunjukkan bahwa jumlah vaksinasi dosis 1 masih mencapai 47% dari 208.265.720 penduduk sasaran vaksinasi. Dari total keseluruhan sasaran vaksinasi tersebut terdapat 40% Masyarakat Rentan dan Umum yang sudah melakukan vaksinasi dosis 1. Hal ini menunjukkan bahwa masih banyak masyarakat Indonesia yang tidak mau melakukan vaksinasi.

Berdasarkan Penelitian mengenai vaksin dan algoritma naïve bayes yang pernah dilakukan oleh (Fathonah & Herliana, 2021) dengan judul “Penerapan Text Mining Analisis Sentimen Mengenai Vaksin Covid-19 Menggunakan Metode Naïve Bayes”. Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap Vaksin COVID-19 dengan menggunakan metode Naïve Bayes. Penelitian ini menggunakan metode Naive Bayes. Naïve Bayes dinilai berpotensi baik dalam mengklasifikasi dokumen dibandingkan metode pengklasifikasian yang lain dalam hal akurasi dan efisiensi Berdasarkan hasil pengujian terhadap 100 data latih yang kemudian dipilih kembali dengan menggunakan teknik data crawling menjadi 34 data, didapatkan bahwa analisis sentimen dari pengguna twitter untuk Vaksin COVID-19 ini didapatkan persentase accuracy sebesar 100%.

Penelitian yang kedua dilakukan oleh (Yulita et al., 2021) dengan judul “Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier” ini memiliki tujuan untuk menganalisis pendapat tentang vaksinasi COVID-19 di Indonesia. Analisis dilakukan terhadap data 3780 tweet yang berkaitan vaksinasi dengan menggunakan algoritma Naïve Bayes Classifier. Berdasarkan analisis, dapat diamati bahwa sebagian besar tweet memiliki sikap positif (60,3 %), sementara jumlah tweet yang netral (34,4 %) melebihi jumlah tweet yang menentang (5,4 %). Nilai akurasi yang dihasilkan sebesar 0,93 (93 %).

Penggunaan vaksin COVID-19 pastinya memunculkan respon dan opini dari masyarakat. Kepercayaan masyarakat terhadap vaksin COVID-19 bervariasi karena keterbatasan jenis vaksin dan profil keamanannya. Persepsi masyarakat terhadap vaksin sangat beragam dan saat ini cenderung negatif, karena masyarakat Indonesia saat ini masih banyak terpengahuh oleh informasi *hoax* dan tidak sesuai fakta yang beredar di media sosial. Beredar postingan di Twitter yang memperlihatkan seseorang bernama Dr. Johan Denis mengklaim vaksin COVID-19 tidak terbukti efektif bahkan aman. Baru-baru ini juga sebuah video telah beredar di media sosial yang menunjukan koin menempel pada suntikan vaksin COVID-19 dengan asumsi vaksin tersebut mengandung magnet (Rokom, 2021). Setelah dilakukan penelusuran fakta, semua informasi tersebut adalah *hoax*. Informasi *hoax* tersebut mengakibatkan tingkat kepercayaan masyarakat terhadap vaksin menjadi berkurang.

Menurut survei Kementerian Kesehatan RI, persentase responden yang ingin divaksinasi adalah 64,8%. 27,6% responden menyatakan tidak tahu, dan 7,6% responden menyatakan sangat menentang vaksinasi terhadap COVID-19. Survei dilakukan pada November 2020 dengan lebih dari 112.000 responden (Rosa, 2021). Penolakan terhadap vaksin COVID-19 sempat menjadi trending topik di Twitter. Perspektif masyarakat yang cenderung negatif terhadap vaksin COVID-19 jika tidak segera diatasi dan dibiarkan berlarut-larut maka akan menyebabkan masyarakat semakin enggan melakukan vaksinasi. Banyak media yang mempengaruhi persepsi publik tentang vaksin COVID-19. Media sosial merupakan salah satu media yang paling banyak digunakan untuk menyampaikan pendapat tentang apapun. Twitter adalah salah satu jejaring sosial yang paling banyak digunakan oleh masyarakat Indonesia untuk mengekspresikan pendapat mereka secara bebas.

Ada beberapa metode analisa yang bisa digunakan untuk menganalisis opini publik, seperti di jejaring sosial Twitter. Salah satunya adalah metode sentimen analisis. Metode sentimen analisis adalah metode analisis data yang diperoleh dari internet untuk dapat mengetahui polaritas data yang ada (Laurensz & Eko Sediyono, 2021). Contoh *tweet* yang berkaitan dengan vaksin COVID-19 yang diunggah oleh salah satu pengguna Twitter yang bernama Budijono “Bila setelah vaksin sinovac di suntikan dan ada yang malah kena covid, bahkan bisa koit siapa yang akan tanggung jawab? Bila nakes berjatuhan korban jiwa siapa yang akan merawat para pasien?”.

Analisis persepsi didapatkan dengan mengambil cuitan *tweet* yang sesuai dengan kata-kata yang berhubungan dengan vaksin Sinovac, Astrazeneca, Moderna dan Pfizer. Cuitan Twitter tersebut nantinya akan dilakukan klasifikasi dengan metode Naïve Bayes Multinomial Text. Penulis menggunakan metode Naïve Bayes Multinomial Text karena metode ini menggunakan pendekatan probabilistik yang artinya cocok digunakan untuk melakukan klasifikasi sebuah kelas yang tidak dapat diprediksi secara pasti. *Tweet* tersebut terkadang tidak memiliki atribut yang saling berkorelasi saat menyusun *tweet* dengan kelas sentimen (Yulianto et al., 2019). Selain itu metode Naïve Bayes Multinomial Text bisa menghasilkan akurasi cukup tinggi untuk klasifikasi sentimen teks dengan data *tweet* (Yulianto et al., 2019).

Hasil dari klasifikasi cuitan Twitter nantinya akan menghasilkan persepsi masyarakat apakah sentimen tersebut positif atau negatif. Dari hasil analisis sentimen tersebut diharapkan dapat memberikan informasi kepada pihak pengembang vaksin tentang bagaimana respon masyarakat terhadap vaksin COVID-19 di Indonesia. Berdasarkan latar belakang yang terjadi, maka dilakukanlah penelitian terkait analisis persepsi masyarakat terhadap vaksin COVID-19 di Indonesia.

## Rumusan Masalah

Rumusan masalah dalam penelitian ini yang dapat ditarik dari latar belakang adalah bagaimana penerapan metode Naïve Bayes Multinomial Text untuk menganalisis persepsi masyarakat terhadap vaksin di Indonesia berdasarkan sentimen *twitter*?

## Tujuan

Tujuan dari penelitian ini adalah menganalisis persepsi masyarakat terhadap penggunaan vaksin Sinovac, Vaksin Astrazeneca, Vaksin Pfizer dan Vaksin Moderna di Indonesia dan mengklasifikasi ke dalam kelas positif dan negatif. Kemudian hasil akurasi dari perhitungan sistem akan dibandingkan dengan menggunakan perhitungan WEKA.

## Manfaat

Adapun manfaat yang diharapkan dari penelitian ini adalah sebagai berikut :

1. Mengetahui persepsi cuitan *tweet* terkait vaksin COVID-19apakah positif atau negatif*.*
2. Memberikan informasi kepada pihak pengembang vaksin tentang bagaimana respon masyarakat terhadap vaksin COVID-19.

## Batasan Masalah

1. Fitur yang digunakan untuk mengambil cuitan *tweet* yang berhubungan dengan vaksin Sinovac, Vaksin Astrazeneca, Vaksin Pfizer dan Vaksin Moderna.
2. Pengambilan data menggunakan Twitter API.
3. Jumlah *tweet* yang akan digunakan adalah 500 data *tweet* setiap untuk kategori vaksin Sinovac, 150 data *tweet* vaksin Astrazeneca, 140 data *tweet* vaksin Pfizer dan 120 data *tweet* vaksin Moderna.
4. Data cuitan *tweet* yang diambil menggunakan Bahasa Indonesia.
5. Analisis sentimen dilakukan pada 2 kelas yaitu positif dan negatif
6. Sistem aplikasi berbasis web.
7. Sistem ini sampai ke tahap perhitungan akurasi.

## Metodologi Penelitian

### Tempat dan Waktu Penelitian

Tempat : Lokasi yang peneliti lakukan dalam penelitian ini adalah di kota Malang dengan menggunakan cuitan *tweet* sebagai objek penelitian.

Waktu : Agustus 2021 – Juli 2022

Tabel 1. 1 Waktu Penelitian

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Tahapan |  | | | | | | | | | | | |
| I | II | III | IV | V | VI | VII | VIII | IX | X | XI | XII |
| Perencanaan |  |  |  |  |  |  |  |  |  |  |  |  |
| Analisis Masalah dan Perancangan Sistem |  |  |  |  |  |  |  |  |  |  |  |  |
| Kontruksi Sistem |  |  |  |  |  |  |  |  |  |  |  |  |
| Pengujian Sistem |  |  |  |  |  |  |  |  |  |  |  |  |
| Pengumpulan Data |  |  |  |  |  |  |  |  |  |  |  |  |
| Analisa Data |  |  |  |  |  |  |  |  |  |  |  |  |
| Klasifikasi Naïve Bayes |  |  |  |  |  |  |  |  |  |  |  |  |
| Hasil, Pembahasan dan Kesimpulan |  |  |  |  |  |  |  |  |  |  |  |  |

### Bahan dan Alat Penelitian

Bahan dan alat penelitian yang digunakan oleh peneliti dalam melakukan penelitian ini adalah :

1. *Hardware*

Perangkat keras yang digunakan peneliti dalam melakukan penelitian ini adalah :

1. *Processor intel core* i3 7th Gen
2. *Memory* 4 GB RAM
3. *SSD 500 Gb*
4. *Software*

Perangkat lunak yang digunakan peneliti dalam melakukan penelitian ini adalah:

1. Sistem operasi windows 10
2. Xampp
3. Browser Microsoft Edge Versi 93.0.961.52 (Pembuatan Resmi) (64 bit)
4. Balsamiq
5. Lucid.app
6. Weka version 3.8.6
7. Bahan

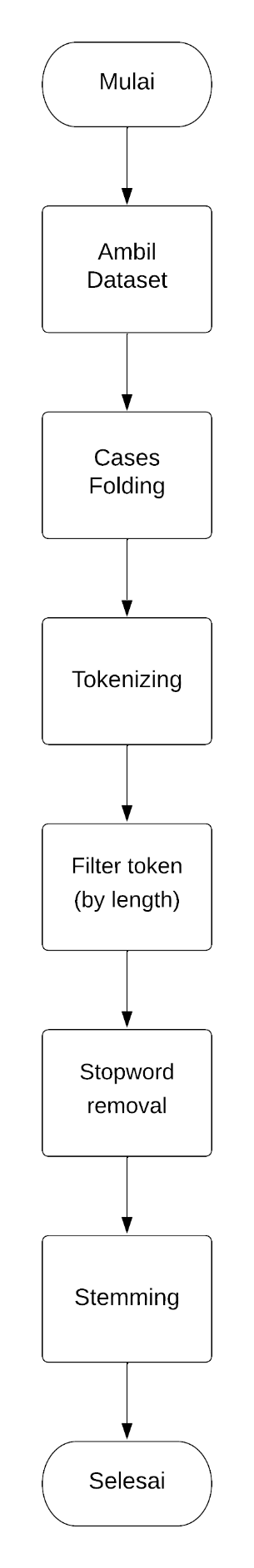
Bahan yang digunakan peneliti dalam melakukan penelitian ini adalah data *tweet* yang diambil dari unggahan twitter yang berhubungan dengan objek penelitian.

### Pengumpulan Data

Pengumpulan data dilakukan oleh peneliti menggunakan twitter API dan menggunakan teknik *web scrapping*. *Web* *scrapping* akan melakukan pengambilan data di Twitter berdasarkan *keyword* yang berhubungan dengan Sinovac, Astrazeneca, Pfizer dan Moderna.

### Analisa Data

Analisa data dilakukan melalui proses pengelompokan data dan memberikan label. Tahapan *preprocessing* meliputi *case folding* dengan merubah teks dalam bentuk baku atau dirubah menjadi huruf kecil (*lowercase*). Kemudian dilanjutkan ke proses *Tokenizing* yaitu menghilangkan karakter khusus, tanda baca, atau karakter non-huruf. Selanjutnya proses *Filter Token* (*By Length)* kata-kata yang lebih pendek dari 4 atau lebih panjang dari 25 karakter akan dihapus, misalnya ‘tdk’, ‘yg’, dan ‘gan’. Kemudian proses *stopword removal* Kata yang tidak relevan, misalnya tetapi, dengan, untuk, yang, dan kata sambung lainnya, akan dihapus. Dan yang terakhir adalah proses *Stemming yaitu* mengubah kata yang berimbuhan menjadi kata dasar.

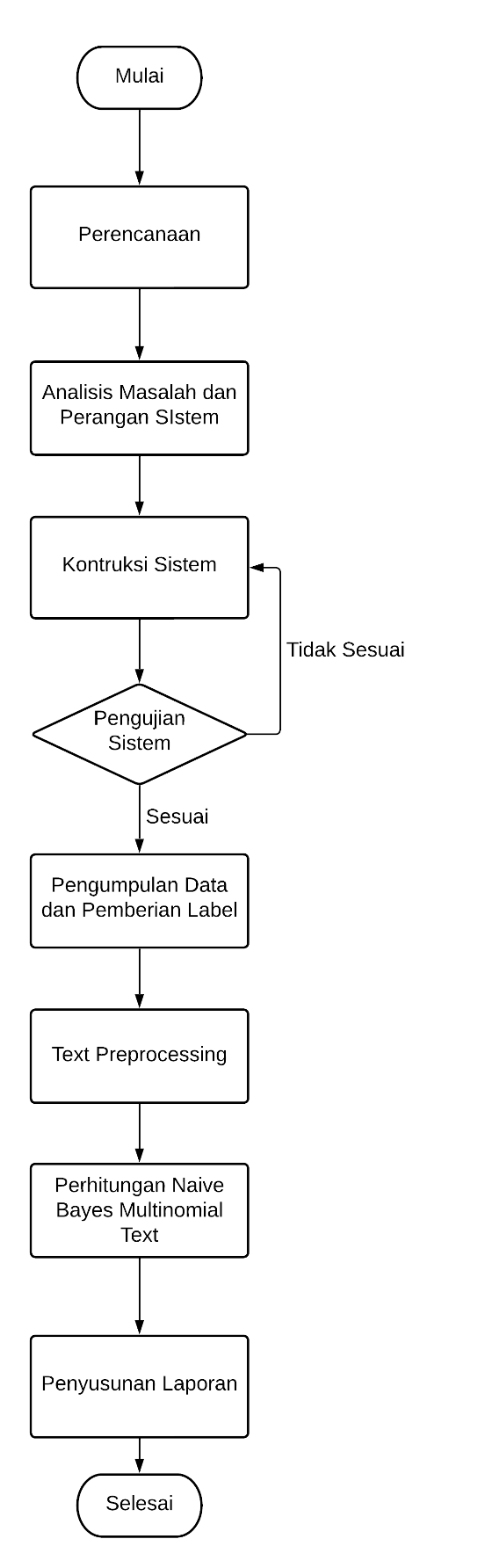


Gambar 1. 1 *Flowchart Preprocessing*

Tabel 1. 2 Contoh Hasil *Preprocessing*

|  |  |
| --- | --- |
| **Tahapan *Preprocessing*** | **Hasil** |
| Kalimat asli | H+1 pasca Vaksin Sinovac, badan pegel2 kek abis workout dan Lengan nyeri banget di sekitaran bekas jarum suntik |
| *Cases folding* | h pasca vaksin sinovac badan pegel kek abis workout dan lengan nyeri banget di sekitaran bekas jarum suntik |
| *Tokenizing* | pasca vaksin sinovac badan pegel kek abis workout dan lengan nyeri banget di sekitaran bekas jarum suntik |
| *Filter token (by length)* | pasca vaksin sinovac badan pegel abis workout lengan nyeri banget sekitaran bekas jarum suntik |
| *Stopword removal* | pasca vaksin sinovac badan pegel abis workout lengan nyeri banget sekitaran bekas jarum suntik |
| *Stemming* | pasca vaksin sinovac badan pegel abis workout lengan nyeri banget bekas jarum suntik |

### Prosedur Penelitian



Gambar 1. 2 Prosedur Penelitian

## Sistematika Penulisan

**BAB I PENDAHULUAN**

Menjelaskan mengenai latar belakang masalah, rumusan masalah, tujuan, manfaat, batasan masalah, metodologi penelitian, dan sistematika penulisan.

**BAB II TINJAUAN PUSTAKA**

Menjelaskan tentang penelitian terkait dan teori-teori yang digunakan sebagai dasar dalam penulisan penelitian.

**BAB III ANALISA DAN PERANCANGAN**

Berisi tentang analisis yaitu identifikasi masalah, pemecahan masalah, perancangan sistem, perancangan data, perancangan user interface dan rancangan pengujian

**BAB IV IMPLEMENTASI DAN PEMBAHASAN**

Menjelaskan tentang pengimplementasian kebutuhan perangkat lunak dan perangkat keras dan gambaran sistem yang digunakan.

**PENUTUP**

Membahas tentang kesimpulan dan saran yang diperoleh dari pembuatan penelitian tugas akhir