

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Pada bulan Maret tahun 2020, masyarakat dihebohkan oleh kasus pertama virus yang berasal dari China yaitu COVID-19 (*Corona Virus Disease 2019*) atau virus Corona, COVID-19 diperkirakan akan berlanjut untuk waktu yang belum dapat dipastikan (Azizah, 2020). Pemerintah melakukan berbagai opsi untuk memutuskan penyebaran virus ini, hingga saat ini COVID-19 telah merenggut banyak korban jiwa dari berbagai kalangan dan usia (Merdeka, 2020). Pada ayat Al-Qur'an surah Al-Anbiya ayat 83-84 yang mengisahkan tentang ujian yang didapatkan oleh Nabi Ayyub AS dari Allah SWT berupa penyakit. Yang berbunyi :

﴿وَأَيُّوبَ إِذْ نَادَى رَبَّهُ أَيْ مَسْنِي الضُّرِّ وَأَنْتَ أَرْحَمُ الرَّحِيمِينَ ۝٨٣ فَاسْتَجَبْنَا لَهُ فَكَشَفْنَا مَا بِهِ مِنْ ضُرِّهِ وَأَتَيْنَاهُ أَهْلَهُ وَمِثْلَهُمْ مَعَهُمْ رَحْمَةً مِّنْ عِنْدِنَا وَذَكَرَى لِلْعَبِيدِينَ ۝٨٤﴾ (الانبیاء/21:83-84)

Artinya: “(Ingatlah) Ayyub ketika dia berdoa kepada Tuhannya, “(Ya Tuhanku,) sesungguhnya aku telah ditimpa penyakit, padahal Engkau Tuhan Yang Maha Penyayang dari semua yang penyayang. Maka, Kami mengabulkan (doa)-nya, lalu Kami lenyapkan penyakit yang ada padanya, Kami mengembalikan keluarganya kepadanya, dan (Kami melipatgandakan jumlah mereka) sebagai suatu rahmat dari Kami dan pengingat bagi semua yang menyembah (Kami).” (Al-Anbiya'/21:83-84)

Dalam kisah yang dijelaskan pada ayat tersebut merupakan suatu bentuk peringatan kepada umat manusia, bahwa siapapun dapat mengalami ujian berupa penyakit. Oleh karena itu, sebagai manusia harus selalu berdoa dan berikhtiar, yang mana salah satu bentuk ikhtiar dengan menjaga kesehatan dan sebagai tenaga medis terus berikhtiar melalui penelitian klinis untuk mendapatkan pengobatan yang tepat terhadap penyakit terkait, seperti halnya terhadap virus COVID-19 ini.

Pada bulan Januari tahun 2021, program vaksinasi COVID-19 dilakukan oleh pemerintah. Meski telah beberapa bulan program vaksinasi berjalan, tetapi program vaksinasi tersebut tidak secara merata diterima oleh masyarakat. Sehingga, program tersebut sempat menuai beragam tanggapan di kalangan masyarakat (Syarifuddin, 2020).

Hal itu membuat masyarakat mengungkapkan pendapat dan opini mereka di berbagai jejaring sosial, salah satunya adalah Twitter. Twitter merupakan sosial media

yang hanya mengizinkan pengguna untuk mengunggah 140 huruf yang biasa disebut *twit* atau kicauan (Suryono, 2018). Tetapi tidak semua *tweet* yang dikirim oleh pengguna memiliki makna yang sama. Oleh karena itu, analisis perlu dilakukan terhadap opini yang dilontarkan oleh *netizen* di Twitter tentang isu vaksinasi dengan mengklasifikasi positif dan negatif, untuk melakukan klasifikasi opini dan respons pengguna Twitter dapat menggunakan analisis sentimen. Analisis sentimen merupakan kegiatan yang mengacu pada komputasi linguistik, pengolahan bahasa, serta *text mining* yang ditujukan untuk menganalisis opini dan evaluasi. Analisis sentimen berfungsi untuk mengelompokkan suatu teks pada sebuah kalimat yang kemudian menentukan apakah teks atau opini tersebut bersifat positif, negatif atau netral (Sri Lestari, 2021). Terdapat beberapa algoritma dalam klasifikasi untuk melakukan klasifikasi *text* yaitu *Naive Bayes*, *Support Vector Machine*, dan *K-Nearest Neighbor*. Pada dasarnya *K-Nearest Neighbor* berfungsi untuk mencari sejumlah K pola terdekat, *Support Vector Machine* berfungsi untuk mencari *hyperlane* terbaik untuk membagi kelas dan *Naive Bayes* berfungsi untuk mencari nilai probabilitas terbesar dari setiap kelas (Fithriasari, 2020).

Beberapa penelitian terkait dengan analisis sentimen terhadap isu vaksin di aplikasi Twitter yang telah dilakukan oleh (Syarifuddin, 2020) dengan hasil akurasi yang diperoleh dari penelitian tersebut sebesar 63.21% untuk metode Naive Bayes dan KNN sebesar 58.10%. dan penelitian yang dilakukan oleh (Muhammad Dani Mulyawan, 2021) dengan hasil yang diperoleh dari penelitian menunjukkan bahwa metode SVM cukup baik untuk mengklasifikasi data dengan akurasi sebesar = 82.738%.

Berdasarkan penjelasan permasalahan di atas, maka penulis mengusulkan penelitian yang berjudul Analisis Sentimen Terhadap *Tweet* Pada Aplikasi Twitter Terhadap Vaksinasi COVID-19. Penelitian ini menggunakan empat metode klasifikasi yaitu *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* dan *Logistic Regression* untuk menentukan algoritma sentimen dari *tweet* yang membahas isu vaksin. Dengan hasil persentase dan respons yang didapatkan dari penelitian ini bisa digunakan sebagai evaluasi bagi pemerintah tentang pengadaan vaksin, selain itu hasil persentase dari penelitian ini juga memberikan manfaat sebagai edukasi untuk masyarakat agar yang belum vaksin segera melaksanakan vaksin dan tidak terpengaruh oleh opini atau pendapat negatif.

Penulis menggunakan aplikasi Twitter untuk menganalisis *tweet* berbahasa Indonesia tentang vaksinasi menggunakan Jupyter Notebook dan Google Colab.

Banyaknya opini dan respons melalui tweet inilah yang akan dimanfaatkan untuk melakukan analisis sentimen. Untuk mendapatkan data yang dibutuhkan penulis menggunakan data *open source* dari *website* kaggle.com, setelah mendapatkan data yang dibutuhkan data tersebut akan memasuki tahap *preprocessing*. Pada tahap ini data yang didapat akan dianalisis menggunakan algoritma *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* dan *Logistic Regression*.

1.2 Perumusan Masalah

Sesuai dengan latar belakang maka perumusan masalah dapat disimpulkan sebagai berikut:

1. Bagaimana cara mengimplementasikan identifikasi data dari Twitter menggunakan metode *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* dan *Logistic Regression* terkait Vaksinasi COVID-19?
2. Bagaimana hasil akurasi data dengan menggunakan metode *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* dan *Logistic Regression* mengenai tweet tentang Vaksinasi?
3. Bagaimana Analisis Sentimen Terhadap *Tweet* Pada Aplikasi Twitter Terhadap Vaksinasi COVID-19 menurut tinjauan Islam?

1.3 Tujuan Penelitian

Penelitian ini memiliki tujuan sebagai berikut:

1. Mengimplementasikan metode algoritma *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* dan *Logistic Regression* pada penambahan data tentang *tweet* mengenai isu vaksinasi COVID-19.
2. Mengetahui kinerja algoritma *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* dan *Logistic Regression* dalam mengidentifikasi tweet positif, negatif atau netral.
3. Memahami Analisis Sentimen Terhadap *Tweet* Pada Aplikasi Twitter Terhadap Vaksinasi COVID-19 menurut tinjauan Islam.

1.4 Research Gap

Pada penelitian yang berjudul Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode Naïve Bayes dan KNN (Syarifuddin, 2020). Peneliti Menggunakan *dataset* yang diambil melalui *website* Twitter sebanyak 1098

tweets. Lalu *dataset* tersebut di uji melalui perbandingan AUC dengan algoritma KNN dan Naive Bayes. Hasil yang didapatkan 63.21% untuk Naive Bayes dan KNN sebesar 58.10%.

Pada penelitian yang berjudul Analisis Sentimen Terkait Vaksin COVID-19 Pada Data Twitter Menggunakan *Support Vector Machine* (Muhammad Dani Mulyawan, 2021). *Dataset* yang digunakan oleh peneliti merupakan data sekunder melalui *Application Programming Interface*(API) yang telah disediakan oleh Twitter. *Dataset* kemudian diolah dan diuji menggunakan algoritma *Support Vector Machine*. Hasil akurasi yang didapatkan sebesar 82.73%.

Pada penelitian yang berjudul Sentimen Analisis Informasi Covid-19 Menggunakan *Support Vector Machine* dan *Naive Bayes* (Ratino, 2020). Peneliti mendapatkan *dataset* dengan cara mengambil komentar pengguna Instagram menggunakan *Instagram Scraper*. Hasil akurasi yang didapatkan adalah 78,02% dengan *score* AUC sebesar 0,714 sedangkan algoritma *Support Vector Machine* mendapatkan akurasi sebesar 80,23% dengan AUC 0,904.

1.5 Metode yang Diajukan

Algoritma yang digunakan pada penelitian skripsi ini adalah *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* dan *Logistic Regression*. Untuk mendapatkan data penulis menggunakan data dari sebuah web *open source* yaitu Kaggle dan *dataset* tersebut diolah dengan cara *preprocessing*.

Setelah melakukan *preprocessing* maka *dataset* akan diuji menggunakan algoritma yang dipilih untuk mendapatkan dan membandingkan hasil akurasi yang didapat dari kedua algoritma.

1.6 Batasan Masalah

1. Dalam penelitian ini menggunakan bahasa pemrograman Python. Untuk melihat akurasi data yang dihasilkan oleh metode terkait *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* dan *Logistic Regression*.
2. Hanya *tweet* berbahasa Indonesia yang digunakan dan terkait dengan masalah vaksinasi.

3. *Dataset* yang digunakan merupakan hasil dari *data open source* dan menggunakan query “#Vaksinasi”, “Covid19” dan “#Vaksin”. menggunakan Twitter API.
4. Algoritma yang digunakan *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* dan *Logistic Regression*.

1.7 Manfaat Penelitian

1. Hasil persentase penelitian ini dapat digunakan untuk menunjukkan respons masyarakat terhadap vaksin.
2. Untuk mengetahui kinerja algoritma *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* dan *Logistic Regression* dalam mengidentifikasi tweet positif, negatif atau netral.
3. Hasil dari penelitian ini dapat digunakan sebagai bahan acuan evaluasi untuk pemerintah tentang pengadaan vaksin.
4. Penelitian ini dapat digunakan sebagai edukasi untuk masyarakat agar melakukan vaksinasi dan tidak terpengaruh oleh opini atau pendapat negatif.
5. Hasil dari penelitian ini dapat digunakan sebagai pembandingan dalam penelitian sejenis.