

BAB 1

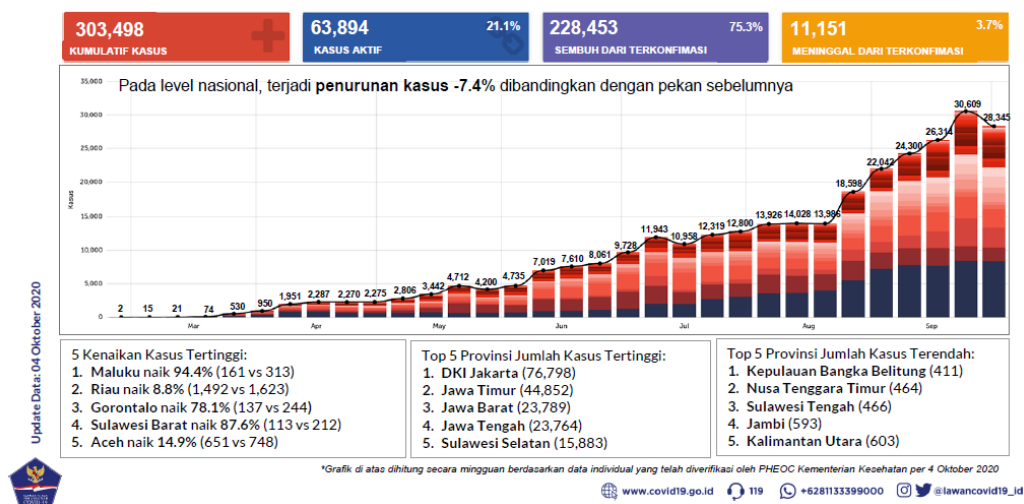
PENDAHULUAN

Pada bab ini akan dijelaskan mengenai latar belakang, rumusan masalah, batasan masalah tujuan, manfaat, dan metodologi penulisan.

1.1 Latar Belakang

Dunia telah digemparkan oleh virus corona (2019-nCov) yang lebih kita kenal Covid-19, Wabah ini ditetapkan sebagai pandemi global oleh World Health Organization (WHO) pada tanggal 11 Maret 2020[1]. Indonesia mengumumkan kasus pertama pada tanggal 2 Maret 2020 yang menjangkit dua orang asal Depok Jawa Barat [2]. Pada update terbaru kali ini jumlah kasus terkonfirmasi virus corona Covid-19 bertambah 10.047 kasus, dan tercatat sebanyak 846.765 pasien positif virus, 695.807 pasien sembuh, dan 24.645 pasien meninggal [3].

Perkembangan Kasus Positif COVID-19 Mingguan



Gambar 1.1 Perkembangan Covid di Indonesia

Penyakit akibat virus memang pada umumnya merupakan ‘self-limiting disease’ yang mengandalkan kekuatan pertahanan imune tubuh. Karena itu banyak disuarakan untuk meningkatkan daya tahan tubuh untuk mencegah tertularnya infeksi virus, dan walaupun tertular, tubuh akan kuat melawannya. Pertahanan tubuh memiliki 5 faktor yang mempengaruhinya sehingga lebih kuat dalam menangkal virus corona ini, berikut kelima faktornya:

1. Lingkungan lebih dominan dibandingkan genetik

Fungsi tubuh secara umum dapat dipengaruhi oleh 2 faktor, yaitu informasi genetik yang dibawa dari keturunan orang tua, dan kedua faktor lingkungan, faktor lingkungan sangat berperan penting dalam perkembangan komponen sistem imun karena dapat membuat imune di tubuh kita bertahan lebih lama.

2. Makanan

Faktor makanan juga merupakan komponen paling penting untuk membentuk diri kita, tidak terkecuali sistem imune tubuh. Makanan yang kaya akan lemak dan gula menyebabkan stress pada sistem imune kita, sehingga mudah terserang penyakit.

3. Usia

Pada umumnya sel-sel imune akan berada pada tingkat puncaknya saat manusia sudah dewasa dan dapat menurunkan aktivitas pertahannya.

4. Kondisi kesehatan

Penyakit kronis (yang diderita lama) dapat meningkatkan kerentanan terhadap infeksi virus ini, kondisi pikiran juga cenderung meningkatkan manusia terinfeksi.

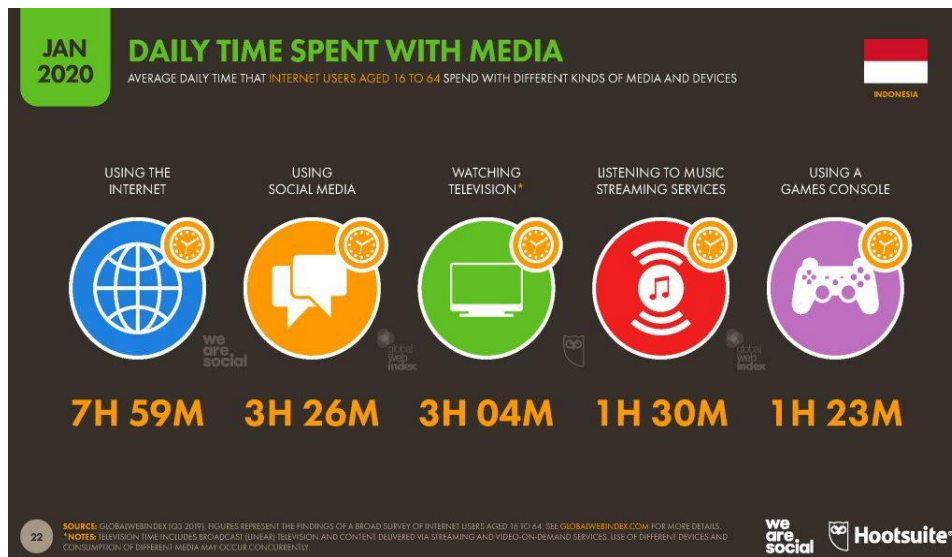
5. Konsumsi obat-obatan

Terdapat beberapa golongan obat-obatan yang dapat menurunkan kemampuan sistem imun kita dalam jangka waktu yang lama dan sangat mungkin menurunkan kekebalan tubuh kita [5].

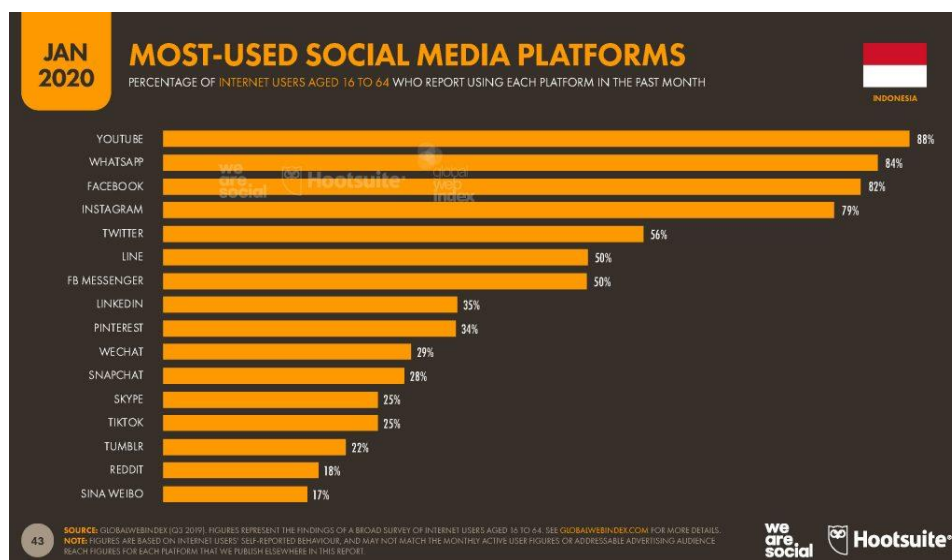
Manusia adalah makhluk yang diciptakan berbeda-beda atau disebut unik, oleh karena itu diperlukannya pertolongan ke dua dan sangat mungkin untuk mencegah penyebaran virus ini, yaitu dengan mengembangkan vaksin.

Dalam menyikapi hal ini Pemerintah Republik Indonesia telah mengeluarkan himbauan vaksinasi untuk masyarakat Indonesia di tahun 2021 ini[6], Rencana kegiatan vaksinasi tersebut, mulai dari sasaran pelaksana, distribusi vaksin, peralatan pendukung, logistik, pelaksanaan pelayanan, pemantauan dan penanggulangan kejadian, strategi komunikasi, pencatatan dan pelaporan, pendanaan, pembinaan, dan terakhir pengawasan. Diharap semua berjalan dengan baik dan terhindar dari hal-hal yang justru akan merugikan. Rencana kegiatan vaksinasi tersebut juga haruslah mempertimbangkan berbagai masukan, di antaranya adalah dengan melihat bagaimana tanggapan penggunaan masyarakat terhadap wacana vaksinasi tersebut.

Masyarakat banyak memberikan pendapatnya di berbagai media, Penelitian kali ini menggunakan media social twitter sebagai sumber dokumen, karena riset pada Januari 2020 twitter menempati nomor lima sebagai platform media social terbanyak digunakan yaitu sebanyak 79 persen atau sekitar 126 juta penduduk yang menggunakan twitter dan waktu yang digunakan dalam menggunakan social media yaitu 3 jam 26 menit.



Gambar 1.2 Waktu penggunaan media sosial



Gambar 1.3 Urutan platform sosial media di Indonesia

Media sosial menjadi media untuk menyuarakan suara masyarakat. Dalam hal ini media sosial sebagai platform untuk merespon program vaksinasi yang terjadi di Indonesia. Media sosial telah mengubah protes yang terjadi dengan konvensional menjadi lebih digital menggunakan media sosial. Media sosial mengubah society menjadi berani naik suara dengan menggunakan social media sebagai platform [8].

Ada beragam teknik klasifikasi text, pertama *machine learning* diantaranya *Support Vector Machine*, *Neural Network*, *Naïve bayes*, *Bayesian Network*, *Maximum Entropy*, dan *K Nearest Neighbor*, Dan yang kedua *Lexicon-Based Approach* diantaranya *Statistical* dan *Semantic*. Metode yang saya gunakan adalah *K-Nearest neighbour* (KNN), dimana algoritma ini dapat membantu mengambil keputusan dengan melihat pilihan mayoritas terbanyak.

Salah satu contoh penggunaan algoritma *K-Nearest Neighbour* (KNN). Anda mengambil keputusan (kelas) antara dokumen atau tidak dokumen ke sebuah acara. Untuk mendukung pengambilan keputusan tersebut, anda melihat kebanyakan dari keputusan teman-teman anda (instance lainnya). Teman-teman anda pilih berdasarkan kedekatannya dengan anda. Ukuran kedekatan pertemanan ini bermacam-macam: tetangga rumah, satu hobi yang sama, satu kelas, atau hal-hal lainnya. Ukuran-ukuran tersebut dapat digunakan bersamaan, misalnya si A itu tetangga, satu hobi, dan satu kelas; sedangkan si B hanya satu kelas saja [9].

Alasan penulis menggunakan Algoritma *K-NN*, karena mudah dipahami dan diimplementasikan, efektif dengan data skala kecil, konsistensi pengujian kuat yang menjamin tingkat *error* tidak lebih dari dua kali, memiliki tingkat akurasi paling tinggi [10].

Berikut perbandingan algoritma yang lain dengan *K-NN* [11],

Tabel 1.1 Perbandingan Algoritma

Parameter	KNN	Naïve bayes	Decision Tree
Determinan/tidak determinan	Tidak determinan	Tidak determinan	Determinan
Efektifitas	Data skala kecil	Data skala global	Data skala besar
Kecepatan	Lama diproses jika data data besar	Lebih cepat dari KNN	Cepat
Data sampel	Tidak bisa jika data masih berantakan	Dapat diproses meskipun data berantakan	Dapat diproses meskipun data berantakan
Akurasi	Tingkat akurasi tertinggi	Hanya bagus untuk data skala sangat besar	Akurasi tinggi

Berikut tabel sampel pengujian perbandingan algoritma.

Tabel 1.2 Sampel Perbandingan Algoritma

Dataset	Size of Dataset	KNN	Naïve bayes	Decision Tree
Weather nominal	Small (14 instance)	100%	92,857%	100%
Segment Challenge	Medium (1500 instace)	100%	81,667%	99%
Supermarket	Large (4627 instace)	89,842%	63,713%	63,713%

Dengan memanfaatkan dokumen dari media sosial *twitter*, penelitian ini bertujuan untuk menganalisis respon masyarakat terhadap penggunaan vaksin dengan cara mengklasifikasikan respon tersebut ke dalam respon positif dan negatif yang nantinya sebagai langkah awal menyelesaikan masalah kesehatan dikarenakan corona virus di Indonesia.

1.2 Rumusan Masalah

1.2.1 Permasalahan

Pemerintah Indonesia perlu mengantisipasi keraguan masyarakat tentang program vaksinasi yang disuarakan di sosial media dengan cara mengukur respon positif maupun respon negatif.

1.2.2 Solusi Permasalahan

Pengukuran ini didapat dari proses perhitungan menggunakan metode *K-NN* dengan ekstraksi *TF-IDF* dan menghasilkan *akurasi*, *recall*, *precision*, *specificity* yang nantinya sebagai acuan kepercayaan terhadap program vaksinasi.

1.2.3 Pertanyaan Penelitian

Agaimana akurasi, recall, specificity, precision terhadap Analisa sentimen penggunaan vaksin di twitter menggunakan metode *K-Nearest Neighbor* dan berapa K terbaik nya ?

1.3 Batasan Masalah

Agar pembuatan skripsi ini lebih terarah dan tidak melebar, maka perlu adanya batasan masalah. Batasan permasalahan yang dibahas sebagai berikut:

1. Dokumen yang diambil merupakan *tweet* dari twitter dalam rentang Bulan September 2020 hingga Juli 2021.
2. Dokumen *tweet* yang di ambil dari twitter berdasarkan kata kunci “Vaksin covid”.
3. *Tweet* yang diambil adalah dokumen berbahasa Indonesia.
4. Dokumen yang digunakan 300 dokumenset.
5. Hasil klasifikasi adalah 2 kelas, yaitu positif dan negatif.
6. Untuk pelabelan dokumenset masih secara manual dengan tiga orang sebagai penentu sentimen positif atau negatif.
7. Untuk menentukan tingkat tetangga terdekat hanya dibatasi $K = 1$ sampai $K=10$.
8. Split Dataset hanya ada satu dimana 70% sebagai data latih dan 30% sebagai data uji.

1.4 Tujuan

Mengetahui respon masyarakat terhadap penggunaan vaksin menggunakan *TF-IDF* sebagai pengekstraksi dengan metode *K-Nearest Neighbor* sehingga dapat mengklasifikasi ulasan positif dan negatif, sehingga respon tersebut dapat diukur baik buruknya.

1.5 Manfaat

Adapun manfaat yang ingin dicapai dan di peroleh dari penelitian ini adalah:

1. Penelitian ini memiliki manfaat untuk mengetahui ulasan di Twitter baik itu berupa ulasan positif maupun negatif untuk dijadikan evaluasi bagi pemerintah sebagai acuan meningkatkan kualitas pelayanan sesuai yang diinginkan masyarakat.

1.6 Metedologi Penulisan

Metedologi penulisan yang menjadi langkah-langkah dalam penyusunan skripsi ini berfungsi untuk membantu mencapai tujuan agar tidak menyimpang dari batasan masalah yang ditentukan. Skripsi ini terdiri dari 5 bab yaitu:

BAB I : PENDAHULUAN

Pada bab ini berisi mengenai latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, dan metedologi penulisan.

BAB II : TINJAUAN PUSTAKA

Pada bab ini akan dijelaskan tentang penelitian yang telah dilakukan sebelumnya, sentimen analisis, vaksin, Covid-19, *K-Nearest Neighbor*.

BAB III : PERANCANGAN SISTEM

Bab ini membahas tentang metode penelitian yang akan dibuat analisis sistem.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini berisi perhitungan akurasi sistem menggunakan metode *K – Nearest neighbour (KNN)* dan dievaluasi menggunakan *Confusion Matrix*.

BAB V : PENUTUP

Pada bab ini diuraikan mengenai kesimpulan yang diambil dari keseluruhan isi dari laporan serta saran yang disampaikan untuk penelitian selanjutnya.