

BAB I

PENDAHULUAN

1.1 Latar Belakang

Mengenai kebijakan tahun 2023 yang diluncurkan oleh pemerintah Indonesia, mengumumkan bahwa rencana penyebaran nyamuk *Wolbachia* sebagai upaya menekan kasus Demam Berdarah *Dengue* (DBD). Namun, kebijakan ini menimbulkan perdebatan di masyarakat. Sebagian mendukung karena melihat potensi keberhasilannya dalam mengurangi penyebaran DBD, sementara sebagian lainnya menolak karena kekhawatiran akan dampak jangka panjang serta kurangnya pemahaman terhadap teknologi ini. Dari permasalahan ini, muncul gagasan penelitian untuk mengetahui nilai probabilitas dari sentimen masyarakat terhadap nyamuk *Wolbachia* melalui analisis komentar YouTube dengan penerapan model *Naïve Bayes*.

Masalah nyamuk *Wolbachia* kini menjadi topik utama dalam pembahasan publik di Indonesia, terutama terkait dengan kebijakan yang dikeluarkan oleh pemerintah. Kementerian Kesehatan secara resmi telah mengadopsi inovasi teknologi berbasis *Wolbachia* ini dengan harapan dapat secara signifikan mengurangi penyebaran demam berdarah (DB), penyakit yang disebabkan oleh virus DENV (*Dengue Virus*) dan ditularkan oleh vektor utamanya, nyamuk *Aedes aegypti*. Kasus pertama infeksi DENV tercatat di Jakarta dan Surabaya pada tahun 1968 yang kemudian dengan cepat menyebar menjadi masalah kesehatan masyarakat nasional yang serius [1]. Menurut data terbaru dari Kementerian Kesehatan, beban kasus DHF di Indonesia tetap tinggi. Tercatat 53.131 kasus dengan 404 kematian mulai Maret 2024 dan dalam seminggu, angka tersebut melonjak drastis menjadi 60.296 kasus dengan total 455 kematian [2].

Teknologi *Wolbachia* telah diterapkan di berbagai negara seperti Brasil, Australia, Vietnam, Fiji, dan beberapa negara lainnya dengan hasil yang menjanjikan dalam mengendalikan penyakit yang ditularkan oleh nyamuk, termasuk DBD [3]. Studi yang dilakukan di Fiji, Vanuatu, dan Kiribati menunjukkan bahwa pelepasan nyamuk yang telah diinfeksi *strain wMel Wolbachia* berhasil menyebabkan infeksi ini menyebar dalam populasi nyamuk lokal dan bertahan selama beberapa generasi [4]. Efektivitas metode ini terbukti di Yogyakarta, di mana sejak penyebaran nyamuk *Wolbachia* pada

tahun 2017, kasus DBD berhasil berkurang sebesar 77% dan angka rawat inap akibat DBD menurun hingga 86% [5]. Namun, masih ada spekulasi di masyarakat mengenai potensi dampak negatif di masa depan yang menyebabkan sebagian masyarakat menolak teknologi ini.

Kebijakan pemerintah terkait nyamuk *Wolbachia* telah memicu polarisasi pandangan di kalangan masyarakat dengan satu pihak menyatakan dukungan karena efektivitasnya yang diharapkan, sementara pihak lain menolaknya karena kekhawatiran terhadap implikasi jangka panjang dan kurangnya pemahaman publik terhadap aspek teknologi yang terlibat. Oleh sebab permasalahan tersebut, menjadikan penelitian ini penting karena membahas isu yang sedang hangat dan belum banyak diteliti. Dengan menganalisis dari penelitian ini diharapkan dapat membantu pemerintah membuat strategi komunikasi yang lebih baik dengan melihat perasaan publik, sehingga dapat mengurangi potensi konflik sosial dan terus meningkatkan penerimaan publik terhadap inovasi kesehatan ini.

Platform *online*, terutama media sosial memainkan peran sentral dengan memberikan akses kepada pengguna untuk berpartisipasi, berbagi, dan memproduksi berbagai jenis konten digital seperti jejaring sosial, blog, dan sebagainya sehingga khalayak luas dapat mengakses sumber informasi [6]. Di antara semua saluran yang tersedia, YouTube menempati posisi penting sebagai platform populer yang banyak digunakan oleh individu dan kelompok untuk mengekspresikan pendapat mereka. Indonesia merupakan negara dengan jumlah pengguna YouTube terbesar keempat di dunia dengan total 139 juta pengguna, menurut penelitian yang dilakukan oleh *We Are Social* per Oktober 2023 [7]. Maka dari itu, variasi komentar yang tercatat di YouTube merupakan sumber data primer yang esensial dan representatif untuk menganalisis dan memahami sentimen publik terhadap suatu isu.

Teori Bayes yang dikembangkan oleh Thomas Bayes, seorang matematikawan Inggris pada abad ke-18 adalah asal dari algoritma klasifikasi *Naïve Bayes* yang didasarkan pada gagasan probabilitas. Algoritma ini berfungsi untuk melakukan klasifikasi dan prediksi dengan menentukan nilai probabilitas tertinggi dari suatu peristiwa. Berkat kemudahan implementasinya, *Naïve Bayes* telah diakui sebagai salah satu algoritma paling populer di bidang data *mining* [8].

Dalam pencarian nilai probabilitas, digunakan analisis sentimen untuk mengidentifikasi berbagai kategori emosional seperti positif, negatif, dan netral terkait suatu isu. Analisis sentimen memanfaatkan

pemrosesan bahasa alami, analisis teks, serta linguistik komputasional untuk mengekstrak informasi subjektif dalam suatu sumber. Saat ini, analisis sentimen telah mendapatkan pengakuan yang luas. Tidak hanya di kalangan peneliti, melainkan juga di kalangan perusahaan, pemerintah, dan berbagai organisasi [9]. Salah satu keuntungannya ialah memberikan wawasan rasional kepada pemangku kepentingan mengenai suatu isu dan manfaat yang dapat diperoleh dari analisis tersebut.

Studi literatur sebelumnya telah menerapkan berbagai metode klasifikasi dalam analisis sentimen termasuk *Support Vector Machine* (SVM), Regresi Logistik, dan *Naïve Bayes* yang diterapkan pada data dari banyak situs jejaring sosial seperti YouTube, Instagram, dan sebagainya. Misalnya, penelitian oleh Nurul Hijriani dan Ermatita yang membandingkan *Naïve Bayes Classifier* (NBC) dan SVM pada komentar dari situs web Apify mencatat akurasi 82,16% untuk NBC dan 83,26% untuk SVM [10]. Selain itu, Styawati dkk. mencapai akurasi tinggi sebesar 98,67% dengan kernel linier dan 98,34% dengan kernel RBF menggunakan metode SVM pada 2.000 titik data dari Twitter [11]. Sementara itu, Sri Mulyani dan Rice Novita menunjukkan akurasi 91% melalui kombinasi *Naïve Bayes* dan *K-Fold Cross Validation* pada 4.783 komentar YouTube [12]. Selain itu, Ritika Singh dan Ayushka Tiwari menggunakan beberapa model dengan dataset 1.500 titik data dalam studi mereka dan memperoleh hasil akurasi yang bervariasi, yaitu 87% untuk *Naïve Bayes*, 88% untuk SVM dan Regresi Logistik, 87% untuk *Decision Tree* dan *K-Nearest Neighbor*, serta 88% untuk *Random Forest* [13]. Dalam studi lain, Nurul Rezki, Sri Astuti T., dan Siswanto menggunakan algoritma SVM pada 10.000 tweet yang dikumpulkan melalui *crawling* di Twitter dan memperoleh akurasi 89,87% [14]. Akhirnya, studi yang melibatkan 100 titik data oleh Asrikam B., Elfira U., dan Paulus M.A. yang menerapkan algoritma *Naïve Bayes* menghasilkan akurasi sebesar 82% [15].

Namun, belum ada penelitian yang secara spesifik menganalisis sentimen masyarakat Indonesia terhadap isu penyebaran nyamuk *Wolbachia* melalui komentar di YouTube. Sebagian besar penelitian sebelumnya, seperti penelitian yang dilakukan oleh Nurul Hijriani dan Ermatita menitikberatkan pada analisis sentimen dengan memanfaatkan data dari Instagram. Oleh karena itu, penelitian ini memiliki dua kebaruan: (1) Pemanfaatan data Youtube untuk menganalisis sentimen terkait isu nyamuk *Wolbachia*. (2) Uji coba strategi optimasi model *Naïve Bayes Multinomial* dibandingkan dengan penelitian sebelumnya. Meskipun penelitian yang dilakukan oleh Nurul Hijriani dan Ermatita memanfaatkan SMOTE untuk mengatasi ketidakseimbangan, namun penelitian tersebut belum

melibatkan tahap optimasi *hyperparameter* secara mendalam. Oleh karena itu, penelitian ini melanjutkan dengan menguji tiga pendekatan yaitu *baseline* (NBM), *Naïve Bayes Multinomial* + SMOTE, dan kombinasi *Naïve Bayes Multinomial* + SMOTE + *GridSearchCV*. Strategi ini bertujuan untuk mencapai keseimbangan optimal antara *precision* dan *recall* pada kelas minoritas.

Melalui analisis sentimen publik, diharapkan bahwa penelitian ini akan memberikan informasi yang bermanfaat dan relevan dalam memahami secara komprehensif persepsi masyarakat Indonesia untuk pemerintah dan semua pemangku kepentingan terhadap kebijakan pelepasan nyamuk *Wolbachia*. Penelitian ini diberi judul berdasarkan urgensi dan tujuan yaitu “Analisis Sentimen Terkait Isu Penyebaran Nyamuk *Wolbachia* di Indonesia Pada Platform YouTube Menggunakan *Naïve Bayes*”.

1.2 Rumusan Masalah

Masalah utama dalam studi ini dapat dirumuskan berdasarkan latar belakang yang telah dijelaskan di atas menjadi dua pertanyaan fundamental:

1. Sejauh mana metode *Multinomial Naïve Bayes* (MNB) efektif dan secara teknis layak digunakan dalam menganalisis sentimen yang diungkapkan dalam komentar YouTube terkait penyebaran nyamuk *Wolbachia* di Indonesia?
2. Bagaimana kinerja metode *Multinomial Naïve Bayes* saat digunakan untuk mengklasifikasikan sentimen dalam komentar YouTube pada isu yang sama?

1.3 Tujuan Penelitian

Penelitian ini mengacu pada rumusan masalah yang telah ditetapkan menjadi tiga tujuan utama:

1. Melakukan analisis sentimen publik terhadap kebijakan distribusi nyamuk *Wolbachia* di Indonesia melalui penerapan metode *Multinomial Naïve Bayes*.
2. Mengklasifikasikan sentimen dalam data komentar YouTube dan membaginya menjadi tiga kategori: negatif, positif, dan netral masing-masing sebagai komponen penting dari analisis.
3. Mengevaluasi secara komprehensif kinerja dalam pengolahan dan interpretasi data teks melalui metode *Multinomial Naïve Bayes* didasarkan pada metrik evaluasi standar seperti *recall*, *F1-score*, akurasi, dan presisi.

1.4 Batasan Masalah

Penelitian ini menetapkan dua batasan utama untuk memastikan bahwa cakupan penelitian tetap terfokus dan spesifik:

1. Data yang digunakan terbatas pada komentar yang diambil dari platform YouTube yang secara eksplisit membahas isu distribusi nyamuk *Wolbachia* di Indonesia.
2. Dalam analisis sentimen, metode *Multinomial Naïve Bayes* (MNB) diterapkan sebagai teknik klasifikasi tunggal tanpa melibatkan perbandingan dengan algoritma klasifikasi lainnya.

