

BAB I PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi informasi dan komunikasi yang berlangsung dengan cepat memberikan pengaruh terhadap munculnya bervariasi macam aplikasi media sosial, *twitter* ialah salah satunya. Selain menjadi media informasi dan komunikasi, *twitter* juga sering dipakai untuk wadah untuk menyampaikan opini oleh penggunanya atau biasa disebut dengan *Tweet*. Para pengguna *twitter* biasanya akan memberikan opini pada suatu informasi terbaru atau terhangat mengenai suatu hal yang menjadi *trending* topik atau topik utama (Prabowo dkk., 2023). Melalui berbagai *Tweet* pada suatu topik yang sedang *trending*, menjadikan *twitter* sebagai media yang memiliki penyebaran informasi paling cepat. Pengguna *twitter* dapat mengirimkan *Tweet* secara bebas mengenai opininya dan tanpa adanya batasan. Berdasarkan hal tersebut, *twitter* dapat menjadi sarana terbaik untuk melakukan analisis sentimen.

Topik yang menjadi perbincangan pada *twitter* juga beragam, contohnya politik, ekonomi, sosial, olahraga, dan lain sebagainya. Olahraga yang sering menjadi topik perbincangan pada *twitter* adalah sepakbola, karena sepakbola adalah jenis olahraga yang memiliki banyak penggemar diseluruh dunia. Pada olahraga sepakbola terdapat ajang kejuaraan tingkat internasional yang disebut sebagai piala dunia. Piala dunia adalah kompetisi yang diadakan tiap 4 tahun sekali oleh FIFA dan FIFA selalu memilih suatu negara sebagai tuan rumah. (Maulana, 2023). Salah satu negara yang ditunjuk sebagai tuan rumah piala dunia U-20 adalah Indonesia, tetapi hal tersebut tidak jadi dilaksanakan karena disebabkan oleh banyak faktor. Gagalnya piala dunia U-20 di Indonesia menjadi suatu masalah menarik yang menjadi topik perbincangan bagi publik, permasalahan ini mengundang berbagai *Tweet* yang berisi opini atau komentar.

Berdasarkan *Tweet* mengenai gagalnya piala dunia U-20 di Indonesia yang berisi berbagai opini sentimen yang bersifat positif, negatif, dan netral, sehingga memungkinkan untuk menggunakan analisis sentimen. Analisis sentimen merupakan prosedur secara otomatis dalam membuat sebuah data teks guna memperoleh informasi sentimen yang dimasukkan ke dalam kalimat (Hasan & Wahyudi, 2018). Naive Bayes adalah salah satu algoritma yang dapat digunakan untuk melakukan analisis sentimen,

dan sistem akan melakukan pengelompokan akan kemungkinan dari data yang dikumpulkan (Setiawan *et al.*, 2021). Analisis sentimen pada penelitian dilaksanakan untuk melihat berdasarkan opini terhadap suatu masalah yang berisi sentimen positif, negatif, serta netral dengan algoritma *Multinomial Naïve Bayes*.

MultinomialNaïve Bayes merupakan salah satu teknik bayes yang digunakan untuk menghitung keseringan dari tiap kemunculan data yang berupa kata pada dokumen (Sriyano & Setiawan, 2021). Metode *MultinomialNaïve Bayes* memiliki banyak kelebihan, diantaranya tingkat akurasi tinggi, kecepatan komputasi rendah, mudah untuk diimplementasikan, serta minimum *error rate*. Hal ini menjadi alasan dipilihnya teknik *MultinomialNaïve Bayes* dalam temuan tersebut.

Banyak penelitian sebelumnya mengenai analisis sentimen dengan metode *Multinomial*, seperti temuan oleh Winahyu & Suharjo (2021) berjudul “Aplikasi Web Analisis Sentimen Dengan Algoritma *MultinomialNaïve Bayes*” data digunakan diperoleh dari *twitter*, menunjukkan hasil bahwa metode *MultinomialNaïve Bayes* terbukti paling akurat dalam mengklasifikasikan sentimen *Tweet* dengan nilai akurasi mencapai 69%. Temuan lainnya dilakukan oleh Hanafi & Solichin (2023) berjudul “Analisis Sentimen Terhadap PSSI Atas Tragedi Kanjuruhan Menggunakan *MultinomialNaïve Bayes*”. Data yang digunakan diperoleh dari *twitter* pada *Tweet* mengenai tanggapan masyarakat kepada PSSI baik sebelum atau sesudah tragedi kanjuruhan, hasil dari pengujian dengan metode *MultinomialNaïve Bayes* memiliki kemampuan tertinggi untuk mengklasifikasikan sentimen dari *Tweet* sebelum tragedi kanjuruhan sebanyak 73%% dan setelah tragedi sebesar 68%.

Temuan sebelumnya oleh Imam & Santoso (2023) berjudul “Analisis Sentimen Pada *Twitter* Terhadap Gagalnya Pelaksanaan Piala Dunia Di Indonesia Menggunakan Metode *Naïve Bayes*” memperoleh hasil bahwa, jumlah data *Tweet* sebanyak 501, dengan rincian 434 *Tweet* positif dan 67 *Tweet* negatif mengenai gagalnya piala dunia U-20 di Indonesia, *Naïve Bayes* sendiri memiliki akurasi tinggi yaitu 76.05% mengklasifikasikan popularitas sentimen pada *Tweet*. Pembeda dari temuan tersebut dengan temuan ini yaitu metode yang digunakan, dalam temuan sebelumnya, metode *Naïve Bayes* digunakan, kali ini, metode *MultinomialNaïve Bayes* digunakan.

Berdasarkan uraian diatas, adanya urgensi mengenai gagalnya pelaksanaan pildun U-20 di Indonesia, mengundang banyak opini positif ataupun negatif. Dari uraian di atas, peneliti mengambil judul penelitian **“Analisis sentimen pada *twitter* pasca dibatalkannya pagelaran piala dunia U-20 yang diselenggarakan di Indonesia menggunakan metode *MultinomialNaïve Bayes*”**, dengan tujuan untuk mengetahui pendapat atau opini masyarakat pada gagalnya pildun U-20 di Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana hasil klasifikasi sentimen masyarakat terhadap pembatalan penyelenggaraan Piala Dunia U-20 di Indonesia menggunakan metode *MultinomialNaïve Bayes*?
2. Berapa tingkat akurasi, presisi, dan *recall* dari metode *MultinomialNaïve Bayes* dalam mengklasifikasikan sentimen tersebut?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Menganalisis sentimen masyarakat terhadap pembatalan penyelenggaraan Piala Dunia U-20 di Indonesia menggunakan metode *MultinomialNaïve Bayes*.
2. Mengetahui tingkat akurasi, presisi, dan *recall* hasil implementasi metode *Multinomial Naïve Bayes* dalam mengklasifikasikan sentimen masyarakat terhadap pembatalan penyelenggaraan Piala Dunia U-20 di Indonesi.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Bagi Penulis

Penelitian ini dapat menambah pengetahuan dan pengalaman dalam menerapkan teknik *text mining*, analisis sentimen, metode *Multinomial Naïve Bayes*, teknik *Synthetic Minority Oversampling Technique* (SMOTE), serta evaluasi model menggunakan *K-Fold Cross Validation* pada data media sosial *Twitter*.

2. Bagi Peneliti Lain

Penelitian ini diharapkan dapat menjadi referensi dalam pengembangan penelitian di bidang analisis sentimen, khususnya yang menggunakan metode *Multinomial Naïve Bayes* pada data media sosial.

3. Bagi Pengembangan Ilmu Pengetahuan

Penelitian ini diharapkan dapat memberikan kontribusi sebagai bahan perbandingan dan acuan bagi penelitian selanjutnya yang berkaitan dengan analisis sentimen maupun pengembangan metode klasifikasi teks.

1.5 Batasan Penelitian

Adapun batasan penelitian dalam penelitian ini adalah sebagai berikut:

1. Sumber data yang digunakan berasal dari media sosial *Twitter* pada periode 11 April 2023 sampai dengan 24 Juni 2023 yang diperoleh melalui teknik *crawling* menggunakan Google Colaboratory.
2. Data yang digunakan dibagi ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral.
3. Dataset yang digunakan berupa unggahan *Tweet* berdasarkan *trending topic* terkait pembatalan penyelenggaraan Piala Dunia U-20 pada periode 11 April 2023 sampai dengan 24 Juni 2023. Data *Tweet* yang digunakan tidak mencakup gambar, video, maupun tautan (*link*).
4. Metode klasifikasi yang digunakan adalah *Multinomial Naïve Bayes* tanpa melakukan perbandingan dengan metode klasifikasi lainnya.
5. Dataset yang digunakan sebanyak **921** komentar *Twitter* hasil proses *crawling*. (Sesuaikan menjadi 931 jika memang itu yang benar di datasetmu, tetapi harus konsisten dengan abstrak, BAB IV, dan hasil pengujian).
6. Pengujian model dilakukan menggunakan tiga skenario *K-Fold Cross Validation*, yaitu $K = 3$, $K = 6$, dan $K = 9$.
7. Penelitian ini menerapkan teknik *Synthetic Minority Oversampling Technique* (SMOTE) untuk menangani ketidakseimbangan kelas pada dataset.