

BAB I

PENDAHULUAN

1.1 Latar Belakang

Program Indonesia Pintar atau yang disebut dengan PIP merupakan program bantuan sosial yang dicanangkan oleh pemerintah dalam bidang pendidikan. Dalam PIP terdapat Kartu Indonesia Pintar atau yang selanjutnya disebut dengan KIP yakni berupa kartu yang nantinya diberikan kepada Peserta Didik dalam satuan pendidikan formal atau nonformal sebagai identitas untuk memperoleh PIP. Dalam lingkup pendidikan tinggi KIP disebut dengan KIP-Kuliah (KIP-K). Pemerintah Indonesia melalui Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek) meluncurkan program KIP-K ini pada tahun 2020 yang merupakan pengembangan dari program bidikmisi yang sudah ada sejak tahun 2011. KIP-K merupakan respons pemerintah terhadap kesenjangan akses terhadap pendidikan tinggi antara mereka yang memiliki latar belakang kemampuan finansial yang baik dengan mereka yang memiliki latar belakang kemampuan ekonomi lemah. Program ini dibuat dengan tujuan memperluas kesempatan bagi anak-anak dari lapisan masyarakat ekonomi lemah agar mendapat akses untuk menempuh pendidikan tinggi dengan lebih mudah. Mengingat peran penting KIPK dalam mencapai keadilan sosial dan meningkatkan kualitas sumber daya manusia. Evaluasi yang efektif dan berkelanjutan terhadap penerimaan program ini dari masyarakat sangatlah penting untuk menjamin kebijakan berjalan tepat sasaran.

Dalam era digital saat ini, platform media sosial telah menjadi sarana utama bagi individu untuk berbagi pengalaman, pandangan, kritik, dan saran tentang berbagai hal termasuk program pemerintah. X (sebelumnya Twitter), merupakan salah satu media sosial yang banyak digunakan yang berisi postingan dan komentar yang dipublikasikan oleh pengguna, postingan dan komentar yang dipublikasikan oleh pengguna mengenai KIPK mencerminkan respon masyarakat secara langsung mulai dari pujian, apresiasi, kritik dan saran terkait program KIPK. Variasi komentar ini menunjukkan adanya respon positif maupun negatif dari pengguna X/twitter terhadap program KIPK. Oleh karena itu, data komentar X/twitter

menjadi sumber yang penting untuk mendapatkan informasi tentang KIPK yang dapat digunakan untuk analisis sentimen.

Untuk mengolah data komentar yang masif dan tidak terstruktur dari X/twitter secara efektif dan objektif, diperlukan pendekatan analisis sentimen. Analisis sentimen adalah metode otomatis untuk mengesktrak, memahami, dan mengklasifikasikan data tekstual ke dalam kategori sentimen (positif dan negatif). Dalam penelitian ini analisis sentimen bermanfaat untuk membantu peneliti mengetahui kecenderungan sentimen masyarakat terhadap penyelenggaraan program KIP-K serta mengidentifikasi isu-isu yang banyak di bahas oleh pengguna platform X/twitter. Hal ini memberikan keunggulan dibandingkan analisis manual yang tidak dapat menangani data komentar X/twitter yang masif dan tidak terstruktur.

Meskipun analisis sentimen menawarkan solusi yang kuat, implementasinya pada klasifikasi data X/twitter sering kali menghadapi masalah, yaitu masalah ketidakseimbangan data (*data imbalance*). Kondisi ini terjadi ketika distribusi jumlah sampel antar kelas sentimen (misalnya, jumlah komentar positif jauh lebih banyak daripada komentar negatif) tidak seimbang. Ketidakseimbangan jumlah data antar kelas dapat memengaruhi kinerja model klasifikasi, sehingga model cenderung lebih baik dalam mengenali kelas yang memiliki jumlah data lebih banyak dibandingkan kelas dengan jumlah data yang lebih sedikit. Kondisi ini berpotensi menghasilkan evaluasi yang kurang representatif terhadap seluruh kelas. Oleh karena itu, diperlukan teknik penyeimbangan data, seperti *Random Undersampling*, untuk menyamakan distribusi antar kelas sehingga proses klasifikasi dapat dilakukan secara optimal. Dalam Penelitian analisis sentimen, terdapat berbagai metode yang dapat digunakan, termasuk salah satunya adalah metode *Multinomial Naïve Bayes* (MNB). Analisis sentimen ini menggunakan ekstraksi fitur TF-IDF.

Multinomial Naïve Bayes (MNB) merupakan salah satu model algoritma analisis dan klasifikasi yang di kembangkan dari algoritma *naïve bayes*. Pemilihan *Multinomial Naïve Bayes* didasarkan pada keunggulannya dalam menangani data diskrit, khususnya pada kemunculan kata dalam sebuah dokumen teks (Ernayanti et al., 2023). Performa akurasi yang tinggi pada klasifikasi teks, serta kecocokannya

dengan karakteristik komentar X/twitter yang bersifat ringkas namun memiliki kosakata yang bervariasi, menjadi alasan pemilihan metode ini.

Berdasarkan latar belakang tersebut, peneliti akan melakukan evaluasi sentimen terhadap KIP-K dengan memanfaatkan data yang diambil dari komentar X/twitter menggunakan model algoritma *Multinomial Naïve Bayes* dan menggunakan Teknik *balancing* yaitu *Random Undersampling*, dengan judul “Evaluasi Sentimen Masyarakat di Platform Media Sosial Terhadap Kartu Indonesia Pintar Kuliah (KIP-K) Berbasis *Multinomial Naïve Bayes*”.

1.2 Rumusan masalah

Berdasarkan latar belakang, maka permasalahan dapat dirumuskan sebagai berikut :

- a. Bagaimana tingkat akurasi, presisi, dan *recall* yang dihasilkan oleh algoritma *Multinomial Naïve Bayes* dengan menerapkan teknik *balancing Random Undersampling* pada evaluasi sentimen masyarakat di platform media sosial terhadap kartu Indonesia pintar kuliah (KIP-K)?.
- b. Berapa jumlah sentimen positif dan negatif pada evaluasi sentimen masyarakat di platform media sosial terhadap kartu Indonesia pintar kuliah (KIP-K) berbasis *Multinomial Naïve Bayes*?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut :

- a. Mengetahui tingkat akurasi, presisi, dan *recall* yang dihasilkan oleh algoritma *Multinomial Naïve Bayes* dengan menerapkan teknik *balancing Random Undersampling* pada evaluasi sentimen masyarakat terhadap kartu Indonesia pintar kuliah (KIP-K) di platform media sosial.
- b. Mengetahui jumlah sentimen positif dan negatif pada evaluasi sentimen masyarakat di platform media sosial terhadap kartu Indonesia pintar kuliah (KIP-K) berbasis *Multinomial Naïve Bayes*.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut :

- a. Mengetahui performa algoritma *Multinomial Naïve Bayes* dengan teknik *balancing Random Undersampling* dalam melakukan evaluasi sentimen masyarakat di platform media sosial terhadap kartu Indonesia pintar kuliah (KIP-K).
- b. Mengetahui jumlah sentimen positif dan negatif pada evaluasi sentimen masyarakat di platform media sosial terhadap kartu Indonesia pintar kuliah (KIP-K) sehingga menjadi dataset yang dapat diproses menggunakan metode *Multinomial Naïve Bayes*.
- c. Hasil penelitian ini dapat digunakan sebagai referensi oleh penelitian selanjutnya dalam melakukan evaluasi sentimen dengan menggunakan metode *Multinomial Naïve Bayes* dan Teknik *balancing Random Undersampling*.

1.5 Batasan Masalah

1. Data yang digunakan berasal dari media sosial X/twitter.
2. Data diambil pada tanggal 1 Juli 2025 – 31 Oktober 2025 dari media sosial X/twitter.
3. Klasifikasi akhir dibagi menjadi 2 kelas yaitu positif dan negatif.
4. Mendapatkan 1000 data dengan 620 data positif dan 380 data negatif.
5. Pembobotan kata menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF).
6. Pembagian data menggunakan *K-Fold Cross Validation* dengan skenario 2-fold, 5-fold, dan 10-fold.
7. Menggunakan algoritma *Multinomial Naïve Bayes*.
8. Menggunakan bahasa pemrograman *Python*.