

BAB III

METODOLOGI PENELITIAN

3.1 Jenis Penelitian

Penelitian ini merupakan penelitian kuantitatif, yaitu jenis penelitian yang menekankan pada pengolahan data numerik dan analisis statistik untuk menjelaskan fenomena tertentu secara objektif (Sugiyono, 2017). Pendekatan yang digunakan dalam penelitian ini adalah klasifikasi data menggunakan algoritma *Naïve Bayes*, dengan tujuan untuk mengelompokkan tingkat *brand awareness* pengikut (followers) terhadap akun *Instagram* Kampus IIB Darmajaya.

Klasifikasi dilakukan berdasarkan persepsi pengguna terhadap elemen visual konten, seperti warna, desain, konsistensi tema, dan elemen grafis lainnya yang disajikan dalam unggahan *Instagram*. Persepsi ini diperoleh melalui data kuesioner yang diolah menjadi atribut-atribut klasifikasi.

Algoritma *Naïve Bayes* dipilih karena memiliki keunggulan dalam menangani dataset bertipe nominal, sederhana dalam penerapan, dan efektif dalam memprediksi kelas dari data baru berdasarkan probabilitas statistik (Septilia Arfida et al., 2022). Dengan pendekatan ini, setiap data responden diklasifikasikan ke dalam tingkat *brand awareness* tertentu, seperti tinggi, sedang, atau rendah, berdasarkan kombinasi atribut yang mereka berikan.

Melalui pendekatan kuantitatif berbasis klasifikasi ini, diharapkan penelitian dapat memberikan gambaran yang akurat mengenai pengaruh konten *Instagram* terhadap *brand awareness* secara terukur dan sistematis.

3.2 Sumber Data

Sumber data dalam penelitian ini terdiri dari data primer dan data sekunder.

- Data primer diperoleh melalui penyebaran kuesioner secara daring (online) kepada pengikut akun *Instagram* resmi Kampus IIB Darmajaya @darmajayathebest. Kuesioner ini bertujuan untuk mengukur persepsi responden terhadap elemen-elemen visual konten *Instagram*, seperti kualitas gambar, estetika warna, dan kejelasan informasi.

- Data sekunder diperoleh dari dokumentasi konten visual yang telah diunggah pada akun *Instagram* resmi Kampus IIB Darmajaya @darmajayathebest, yang digunakan sebagai referensi untuk analisis visual serta pemodelan klasifikasi.

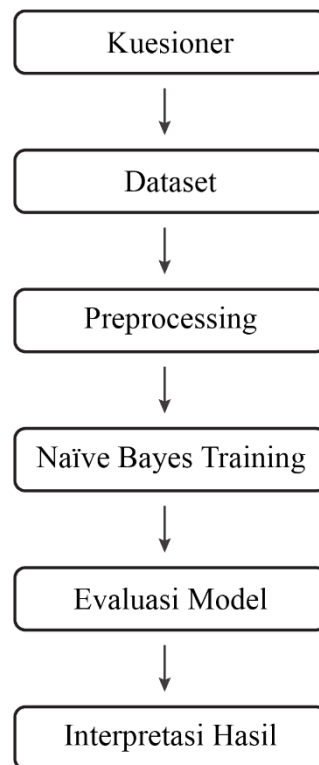
3.3 Metode Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan melalui penyebaran kuesioner daring (*online*) menggunakan *platform Google Form*. Kuesioner ini disebarkan kepada pengikut aktif akun *Instagram* resmi Kampus IIB Darmajaya @darmajayathebest, yang terdiri dari mahasiswa, calon mahasiswa, serta masyarakat umum. Responden dipilih secara *purposive*, yaitu mereka yang secara aktif mengikuti dan berinteraksi dengan konten yang dipublikasikan di akun tersebut.

Instrumen kuesioner terdiri atas pertanyaan-pertanyaan terstruktur yang dirancang untuk mengevaluasi persepsi responden terhadap elemen-elemen visual dalam konten *Instagram*. Aspek yang dinilai meliputi kualitas gambar, estetika warna, konsistensi tema desain, dan kejelasan informasi yang disampaikan. Hasil dari kuesioner ini digunakan sebagai data primer yang kemudian dianalisis menggunakan algoritma *Naïve Bayes* untuk mengelompokkan tingkat *brand awareness* berdasarkan persepsi visual responden.

3.4 Tahapan Klasifikasi Data

Tahapan klasifikasi data dimulai dengan pengumpulan respons kuesioner online yang mengukur persepsi elemen visual *Instagram*. Semua jawaban kemudian dikompilasi menjadi dataset terstruktur. Selanjutnya, dilakukan *Pre-processing* data untuk membersihkan, meng-encode, dan menormalisasi data. Data hasil *Pre-processing* digunakan untuk melatih model *Naïve Bayes* dalam menghitung probabilitas tiap kategori *brand awareness* (tinggi, sedang, rendah). Kinerja model dievaluasi menggunakan *confusion matrix* serta metrik akurasi, presisi, recall, dan *f1-score*. Akhirnya, hasil klasifikasi diinterpretasikan untuk mengetahui pengaruh masing-masing elemen visual terhadap tingkat *brand awareness*.



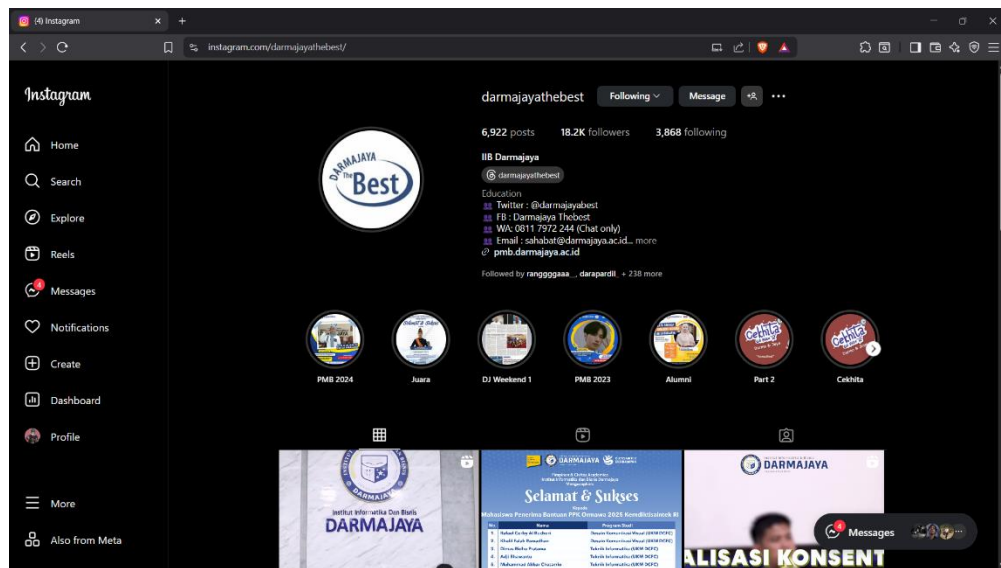
Gambar 3. 1 Alur Penelitian

Gambar 3.1 di atas menjelaskan alur penelitian yang dimulai dengan pengumpulan data melalui Kuesioner yang menghasilkan Dataset terstruktur berisi respons tentang elemen visual Instagram. Dataset ini kemudian melalui tahap Preprocessing untuk dibersihkan, di-encode, dan dinormalisasi. Data yang telah diproses selanjutnya digunakan untuk melatih model *Naïve Bayes (Training)* dalam menghitung probabilitas kategori brand awareness (tinggi, sedang, atau rendah). Setelah pelatihan, Evaluasi Model dilakukan menggunakan confusion matrix serta metrik akurasi, presisi, recall, dan *f1-score* untuk menilai kinerja model, yang puncaknya adalah Interpretasi Hasil guna mengetahui pengaruh elemen visual terhadap tingkat *brand awareness*.

3.5 Populasi dan Sampel

1. **Populasi:** Populasi dalam penelitian ini adalah seluruh pengikut akun *Instagram* resmi milik Kampus IIB Darmajaya, yaitu @darmajayathebest. Berdasarkan data yang diperoleh pada bulan Juli 2025, jumlah pengikut aktif akun tersebut adalah sebanyak 18.200 (18.2k) followers. Populasi ini terdiri

dari mahasiswa aktif, calon mahasiswa, alumni, dan masyarakat umum yang mengikuti akun tersebut serta menjadi target utama dari konten visual yang dipublikasikan oleh pihak kampus. Dengan populasi ini, penelitian diarahkan untuk mengevaluasi persepsi visual followers terhadap konten yang diunggah oleh kampus melalui *platform Instagram*. Jumlah pengikut aktif tersebut dapat dilihat pada Gambar 3.2 berikut.



Gambar 3. 2 Jumlah Populasi Followers Instagram

2. **Sampel:** Penentuan jumlah sampel dilakukan menggunakan **rumus Slovin**, yang digunakan untuk menghitung ukuran sampel dari populasi yang diketahui secara pasti. Adapun rumus Slovin adalah sebagai berikut:

Dengan:

$$n = \frac{N}{1 + N \cdot e^2}$$

$N = 18.200$ (jumlah populasi)

$e = 0,05$ atau 5% (tingkat kesalahan)

Maka:

$$n = \frac{18.200}{1 + 18.200 \cdot (0,05)^2} = \frac{18.200}{1 + 45,5} = \frac{18.200}{46,5} = 391,40$$

Berdasarkan perhitungan tersebut, jumlah minimum responden yang dibutuhkan adalah **391 orang**. Teknik pengambilan sampel yang digunakan adalah **purposive sampling**, yaitu dengan memilih responden yang memenuhi kriteria

tertentu: merupakan pengikut aktif akun @darmajayathebest, pernah melihat atau berinteraksi dengan konten *Feed*, *Story*, atau *Reels*, serta bersedia mengisi kuesioner secara *online*.

3.6 Variabel Penelitian

1. Variabel Bebas (X): Persepsi terhadap elemen visual konten *Instagram*, yang meliputi:
 - Kualitas gambar
 - Estetika warna
 - Konsistensi tema
 - Kejelasan informasi
 - Daya tarik konten Reels
2. Variabel Terikat (Y): Tingkat *Brand awareness*, yang dikategorikan menjadi:
 - Tinggi
 - Sedang
 - Rendah

3.7 Definisi Operasional Variabel

1. **Konten Visual:** Persepsi audiens terhadap aspek estetika, warna, konsistensi tema, dan kejelasan informasi yang terdapat dalam konten *Instagram*.
2. ***Brand awareness*:** Tingkat kesadaran audiens terhadap merek IIB Darmajaya, yang diukur berdasarkan kategori tinggi, sedang, atau rendah, berdasarkan respons terhadap kuesioner.
3. ***Reels*:** Persepsi responden terhadap daya tarik, pesan visual, dan gaya penyampaian konten berbasis video pendek di *Instagram*.

3.8 Metode Analisis Data

Proses analisis data dalam penelitian ini dilakukan melalui beberapa tahapan berikut:

1. ***Pre-processing Data*:** Tahapan ini meliputi normalisasi nilai dan encoding variabel-variabel kategorikal yang terdapat dalam data kuesioner.

Tujuannya adalah untuk memastikan bahwa data siap digunakan dalam proses pelatihan model klasifikasi.

2. **Modeling:** Data yang telah diproses kemudian digunakan untuk melatih model klasifikasi menggunakan algoritma *Naïve Bayes*. Model ini bertugas mengklasifikasikan tingkat *brand awareness* responden berdasarkan persepsi mereka terhadap elemen visual konten *Instagram*.
3. **Evaluasi Model:** Kinerja model dievaluasi menggunakan confusion matrix serta metrik evaluasi berupa akurasi, presisi, *recall*, dan *F1-score*. Evaluasi ini dilakukan untuk mengukur seberapa baik model dalam memprediksi kategori *brand awareness* secara tepat.
4. **Label tingkat *brand awareness*:** Label tingkat *brand awareness* ditentukan berdasarkan rata-rata skor dari lima pertanyaan skala *Likert* (skala 1–5) pada kuesioner. Kategori yang digunakan adalah sebagai berikut:
 - Skor rata-rata 4.0–5.0 dikategorikan sebagai Tinggi
 - Skor rata-rata 2.6–3.9 dikategorikan sebagai Sedang
 - Skor rata-rata 1.0–2.5 dikategorikan sebagai Rendah

Implementasi seluruh proses analisis dilakukan menggunakan bahasa pemrograman *Python*, dengan memanfaatkan *platform Google Colab* sebagai media komputasi berbasis cloud. *Google Colab* digunakan untuk memproses data, melatih, serta menguji model klasifikasi menggunakan algoritma *Naïve Bayes* terhadap data hasil kuesioner yang dikumpulkan melalui Google Form.