

BAB III

METODE PENELITIAN

A. Jenis Penelitian

Jenis penelitian yang dilakukan adalah penelitian kuantitatif. Penelitian ini menggunakan data jamaah umrah di El Bakkah Travel, yang kemudian dianalisis dengan metode statistik yang sesuai. Pendekatan yang digunakan adalah *Naive Bayes Classifier*.

B. Sumber Data

Data yang digunakan adalah data sekunder. Sumber data dalam penelitian ini berasal dari *database internal* milik El Bakkah Travel, yang mengelola dan menyimpan informasi terkait jamaah umrah. Data tersebut merupakan data bulanan jamaah umrah pada bulan September dan Oktober 2024.

C. Variabel Penelitian

Variabel dalam penelitian ini terdiri dari variabel prediktor (X) dan variabel respon (Y). Variabel prediktor meliputi jenis kelamin (X_1), tingkat pendidikan (X_2), pekerjaan (X_3), penghasilan (X_4) dan status pernikahan (X_5) sedangkan variabel respon terdiri dari promo ($Y = 0$) dan reguler ($Y = 1$).

Adapun variabel yang digunakan pada penelitian ini disajikan pada Tabel 3.1.

Tabel 3.1. Variabel yang Digunakan pada Penelitian

No	Atribut	Kategori
1	Paket Umrah (Y)	0 = Promo
		1 = Reguler
2	Jenis Kelamin (X_1)	0 = Laki-laki
		1 = Perempuan
3	Tingkat Pendidikan (X_2)	0 = Tidak Sekolah
		1 = SD/MI
		2 = SMP/MTS
		3 = SMA/MA
		4 = D3

No	Atribut	Kategori
		5 = D4/S1
		6 = S2
4	Pekerjaan (X_3)	0 = Tidak Bekerja
		1 = Petani
		2 = Peg.Swasta
		3 = PNS
		4 = Lainnya
5	Penghasilan (X_4)	0 = Kelompok 1
		1 = Kelompok 2
		2 = Kelompok 3
		3 = Kelompok 4
6	Status Pernikahan (X_5)	0 = Belum Menikah
		1 = Menikah
		2 = Janda
		3 = Duda

Tabel 3.1. memuat variabel prediktor yang digunakan dalam penelitian ini merupakan data bertipe kategori, kecuali variabel penghasilan yang termasuk ke dalam jenis data numerik. Metode *Naive Bayes* secara umum lebih efektif digunakan untuk data kategori, Maka diperlukan proses transformasi data yang dilakukan melalui diskritisasi, yaitu metode yang digunakan untuk mengubah data yang bersifat numerik menjadi data yang bersifat kategori. (Fajriati & Prasetyo, 2023). Salah satu cara yang bisa diterapkan adalah dengan mengelompokkan data ke dalam beberapa interval yang sama, salah satunya menggunakan ukuran nilai letak kuartil. Kuartil merupakan salah satu pengukuran data dalam statistika yang membagi data ke dalam empat bagian sama besar dari data numerik

menjadi data kategori (Maulana dkk., 2022). Pembagian kuartil dalam analisis data terdiri atas kuartil pertama (Q_1) yang memisahkan 25% data terbawah, kuartil kedua (Q_2) atau median yang membagi data menjadi dua bagian yang sama besar, serta kuartil ketiga (Q_3) yang memisahkan 75% data terbawah (Khusaeri, 2019).

D. Teknik Analisis Data

Metode analisis data dalam penelitian ini menggunakan Naive Bayes Classifier dengan bantuan perangkat lunak R-Studio. Metode ini merupakan teknik klasifikasi berbasis probabilitas yang menerapkan Teorema Bayes, dengan asumsi bahwa setiap atribut prediktor bersifat independen satu sama lain ketika nilai variabel respon telah diketahui. Berikut merupakan langkah-langkah yang digunakan untuk melakukan penelitian:

1. Mendeskripsikan data penelitian berdasarkan variabel-variabel yang digunakan.
2. Mengklasifikasikan data penelitian dengan memakai metode Naive Bayes Classifier melalui tahapan berikut:
 - a. Menggunakan data *training* untuk menghitung nilai probabilitas prior dari variabel respon.
 - b. Menghitung nilai probabilitas bersyarat untuk setiap atribut berdasarkan kelas.
 - c. Menghitung nilai probabilitas posterior untuk semua pengamatan.
 - d. Menentukan kelas prediksi dengan menggunakan fungsi arg max.
3. Memilih variabel prediktor menggunakan Informasi Gain untuk setiap kategori data.
4. Menghitung nilai akurasi model klasifikasi yang dilakukan pada data *testing*.