

BAB I

PENDAHULUAN

1.1. Latar Belakang

Ibadah haji merupakan salah satu rukun Islam yang memiliki makna spiritual dan sosial yang mendalam bagi umat Muslim di seluruh dunia. Setiap tahun, jutaan jamaah dari berbagai latar belakang negara berkumpul di tanah suci untuk melaksanakan ibadah ini. Ibadah Haji termasuk salah satu rukun Islam yang diwajibkan oleh Allah bagi muslim yang mampu mengerjakannya (Muslim 2020). Hal ini berarti bahwa setiap muslim yang melaksanakan ibadah haji memiliki kemampuan baik secara pengetahuan, keilmuan yang memadai terkait pelaksanaan dan tata cara pelaksanaan ibadah haji. Mampu juga dalam arti finansial selama pelaksanaan ibadah haji, fisik yang sehat, psikis sehat dan keamanan yang memungkinkan dalam melaksanakan ibadah haji (Agustin Selfia 2022).

Menurut Kementerian Agama Republik Indonesia, jumlah Jemaah haji dari Indonesia mencapai lebih dari 200.000 orang setiap tahun dan terus meningkat. Kementerian Agama (Kemenag) Karawang mencatat ada sebanyak 2.055 calon jemaah haji berangkat ke tanah suci pada tahun 2024. Dan juga pada 2025 tersedia sebanyak 2.055 kuota calon jemaah haji. Dengan pertumbuhan jumlah jemaah yang signifikan, terdapat beberapa tantangan yang kompleks terkait pengelolaan data selama pelaksanaan ibadah haji mengenai calon jemaah.

Masalah yang terkait dengan faktor keberangkatan jemaah haji mencakup berbagai aspek, seperti ketidakmerataan distribusi kuota, keterbatasan akses informasi, serta tantangan dalam memenuhi kebutuhan spesifik jemaah berdasarkan usia, jenis kelamin, pendidikan, dan pekerjaan. Oleh karena itu, solusi untuk permasalahan ini yaitu menggunakan pendekatan algoritma klasifikasi pada *machine learning* diantaranya *Naïve Bayes*, *Random Forest*, dan *Decision Tree*.

Penelitian terdahulu yang menerapkan metode klasifikasi menggunakan algoritma *Naïve Bayes* untuk memprediksi kelayakan peminjam untuk pinjaman dalam pengaturan yang kooperatif dengan menggunakan Teknik *k-fold cross validation* yang mendapatkan tingkat akurasi 78,08% (Lestari et al. 2020). Selanjutnya, analisa perbandingan algoritma *Random Forest* dan *Naïve Bayes* untuk klasifikasi curah hujan berdasarkan iklim di Indonesia dengan metode *Naïve Bayes* yang menunjukkan hasil nilai akurasi yang tinggi dengan hasil 36,61% sedangkan *Random Forest* hanya sedikit lebih rendah (Prakoso Indaryono 2024).

Selanjutnya, penelitian yang menggunakan algoritma *Random Forest* memiliki akurasi 97.26% dan kinerja terbaik dalam menentukan penerima bantuan raskin, dibandingkan dengan *Support Vector Machine* (92.15%) dan *Naïve Bayes* (88.39%) (Kurniawan et al. 2023). Selain itu metode *Random Forest* menunjukkan tingkat akurasi 85,5% yang lebih unggul dibandingkan *Decision Tree* dengan akurasi 84,4% dalam memprediksi keberhasilan pengobatan *imunoterapi* untuk penyakit kutil (Yulian Pamuji, et al. 2021).

Berdasarkan latar belakang dan beberapa referensi tersebut, tujuan dari penelitian ini adalah untuk menganalisa dengan pendekatan tiga algoritma *machine learning* yaitu *Naïve Bayes*, *Random Forest*, *Decision Tree*, penulis berminat untuk melakukan penelitian terhadap klasifikasi data dengan jumlah yang mencapai 5000 jemaah haji, serta akan dilihat berdasarkan karakteristik demografis dengan mengimplementasikan dengan menggunakan tiga algoritma ini. Proses implementasi dilakukan dengan bahasa pemrograman *Python*. Dengan memberikan rekomendasi berbasis teknologi menggunakan metode klasifikasi ini, dapat meningkatkan efisiensi untuk pengelolaan data yang sebelumnya masih dilakukan secara manual. Pendekatan ini juga mengisi kesenjangan literatur karena belum banyak dibahas pada penelitian sebelumnya terkait analisis faktor demografis keberangkatan Jemaah haji menggunakan algoritma klasifikasi *machine learning*.

1.2. Rumusan Masalah

1. Bagaimana penerapan algoritma *Machine Learning* dapat membantu mengklasifikasikan jemaah haji berdasarkan faktor demografis?
2. Bagaimana tingkat akurasi model klasifikasi yang dihasilkan oleh algoritma *Machine Learning* dalam mengidentifikasi kategori jemaah haji?
3. Bagaimana pengaruh faktor demografis (usia, jenis kelamin, pendidikan, dan pekerjaan) terhadap kebutuhan dan pengalaman jemaah haji selama pelaksanaan ibadah?

1.3. Tujuan Penelitian

1. Menerapkan algoritma klasifikasi *machine learning* pada calon jemaah haji berdasarkan faktor-faktor yang berpengaruh secara signifikan untuk meningkatkan efisiensi pengelolaan data jemaah haji.
2. Mengetahui tingkat akurasi model klasifikasi yang dihasilkan oleh algoritma *Naïve Bayes*, *Random Forest*, *Decision Tree*.
3. Menganalisis pengaruh faktor demografis (usia, jenis kelamin, pendidikan, dan pekerjaan) terhadap kebutuhan dan pengalaman jemaah haji selama pelaksanaan ibadah.

1.4. Manfaat Penelitian

1. Mengetahui tingkat akurasi dengan model klasifikasi untuk meningkatkan kinerja dalam pengambilan keputusan strategis berbasis data.
2. Instansi dapat membantu dalam perencanaan dan alokasi sumber daya berdasarkan karakteristik demografis jemaah.
3. Peneliti dapat melakukan penerapan algoritma *Machine Learning* untuk klasifikasi data dan memberikan wawasan mendalam tentang karakteristik Jema