ABSTRAK

Haji merupakan salah satu rukun Islam yang memiliki makna spiritual dan sosial yang mendalam bagi umat muslim di seluruh dunia. Dengan meningkatnya jamaah haji di Indonesia setiap tahunnya, pengelolaan dan pelayanan terhadap calon jamaah haji menjadi tantangan. Faktor demografis seperti usia, pendidikan dan pekerjaan yang mempengaruhi keberangkatan jamaah. Penelitian ini bertujuan untuk mendeteksi faktor keberangkatan jamaah haji menggunakan algoritma *machine learning*, khususnya metode *Naïve Bayes*, *Random Forest* dan *Decision Tree*. Dataset yang dikumpulkan dari Kantor Kementerian Agama Karawang dan diolah menggunakan bahasa pemrograman *Phyton*. Proses penelitian meliputi pengumpulan data, *preprocessing*, split data, implementasi algoritma, dan evaluasi. *Random Forest* mencapai akurasi tertinggi sebesar 99.23%, *Decision Tree* mencatat akurasi 98.75%, dan *Naïve Bayes* memiliki akurasi 76.69%. Hasil evaluasi menunjukkan model mampu memberikan akurasi signifikan dalam mengidentifikasi kategori jamaah haji. Penelitian ini diharapkan dapat memberikan wawasan mendalam mengenai klasifikasi data jamaah haji dan membantu instansi dalam perencanaan serta alokasi sumber daya yang lebih efektif.

Kata Kunci: Jemaah haji, Klasifikasi, Machine Learning

ABSTRACT

Hajj is one of the pillars of Islam that holds profound spiritual and social significance for Muslims worldwide. With the increasing number of Hajj pilgrims in Indonesia each year, the management and services for prospective pilgrims have become a challenge. Demographic factors such as age, education, and occupation influence the departure of pilgrims. This study aims to detect the factors affecting Hajj pilgrim departures using machine learning algorithms, specifically the Naïve Bayes, Random Forest, and Decision Tree methods. The dataset was collected from the Ministry of Religious Affairs Office in Karawang and processed using the Python programming language. The research process includes data collection, preprocessing, data splitting, algorithm implementation, and evaluation. Random Forest achieved the highest accuracy of 99.23%, Decision Tree recorded an accuracy of 98.75%, and Naïve Bayes had an accuracy of 76.69%. The evaluation results demonstrate that the models can provide significant accuracy in identifying Hajj pilgrim categories. This research is expected to offer in-depth insights into the classification of Hajj pilgrim data and assist institutions in planning and allocating resources more effectively.

***Keyword**: Classification, Hajj Pilgrims, Machine Learning*