

ABSTRAK

Penyakit kulit merupakan salah satu jenis penyakit yang umum ditemukan di Indonesia, terutama pada wilayah beriklim tropis. Rumah Sakit Umum Cut Meutia mencatat banyak kasus penyakit kulit yang terdokumentasi dalam rekam medis digital, sehingga memiliki potensi besar untuk dianalisis menggunakan metode klasifikasi berbasis kecerdasan buatan. Penelitian ini bertujuan untuk membandingkan kinerja algoritma *Logistic Regression* dan *Random Forest* dalam mengklasifikasikan delapan jenis penyakit kulit, yaitu Abses, Dermatitis, Selulitis, Psoriasis, Ulkus Dekubitus, Pemfigus Vulgaris, Urtikaria, dan Impetigo berdasarkan 10 gejala utama pasien. Dataset yang digunakan terdiri dari 271 data pasien dan diolah melalui proses pra-pemrosesan, pelatihan, dan pengujian menggunakan Python dan Scikit-learn. Evaluasi model dilakukan berdasarkan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma *Logistic Regression* memperoleh akurasi sebesar 94,55%, sedangkan *Random Forest* memperoleh akurasi sebesar 92,73%. Dengan demikian, *Logistic Regression* terbukti lebih efektif dan efisien dalam klasifikasi penyakit kulit berdasarkan dataset yang digunakan. Temuan ini diharapkan dapat berkontribusi pada pengembangan sistem bantu diagnosis berbasis kecerdasan buatan di bidang medis.

Kata Kunci: Penyakit Kulit, Logistic Regression, Random Forest, Klasifikasi, Data Mining, Evaluasi Model

ABSTRACT

Skin disease is one of the most common diseases in Indonesia, especially in tropical climates. Cut Meutia General Hospital records many cases of skin diseases documented in digital medical records, so it has great potential to be analyzed using artificial intelligence-based classification methods. This study aims to compare the performance of Logistic Regression and Random Forest algorithms in classifying eight types of skin diseases, namely Abscess, Dermatitis, Cellulitis, Psoriasis, Decubitus Ulcer, Pemphigus Vulgaris, Urticaria, and Impetigo based on 10 main patient symptoms. The dataset used consisted of 271 patient data and was processed through pre-processing, training, and testing using Python and Scikit-learn. Model evaluation was performed based on accuracy, precision, recall, and F1-score metrics. The results showed that the Logistic Regression algorithm obtained an accuracy of 94.55%, while Random Forest obtained an accuracy of 92.73%. Thus, Logistic Regression proved to be more effective and efficient in the classification of skin diseases based on the dataset used. These findings are expected to contribute to the development of artificial intelligence-based diagnosis assistance systems in the medical field.

Keywords: *Skin Disease, Logistic Regression, Random Forest, Classification, Data Mining, Model Evaluation*