

ABSTRAK

Kesehatan ibu hamil merupakan aspek penting dalam menjamin keselamatan ibu dan janin. Deteksi dini terhadap risiko kesehatan pada masa kehamilan sangat diperlukan untuk mencegah komplikasi yang dapat berujung pada kematian ibu atau bayi. Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja algoritma Decision Tree dan Random Forest dalam klasifikasi risiko kesehatan ibu hamil menggunakan dataset Maternal Health Risk dari UCI Machine Learning Repository. Proses penelitian meliputi tahapan preprocessing data seperti penanganan missing values, normalisasi, dan deteksi outlier. Selain itu, dilakukan juga penanganan data tidak seimbang menggunakan metode SMOTE untuk memastikan distribusi kelas yang seimbang. Model dievaluasi menggunakan metrik akurasi, presisi, recall, F1-score, dan validasi silang (k-fold cross-validation). Hasil eksperimen menunjukkan bahwa algoritma Random Forest memiliki kinerja yang lebih unggul dibandingkan Decision Tree, baik pada data seimbang maupun tidak seimbang. Random Forest mampu memberikan akurasi yang lebih tinggi (86,64%) dan stabil serta lebih tangguh terhadap variasi data. Penelitian ini menyimpulkan bahwa Random Forest lebih efektif digunakan dalam klasifikasi risiko kesehatan ibu hamil dan dapat menjadi alternatif metode yang andal dalam sistem pendukung keputusan medis. Penelitian ini menyimpulkan bahwa Random Forest lebih efektif digunakan dalam klasifikasi risiko kesehatan ibu hamil dan dapat menjadi alternatif metode yang andal dalam sistem pendukung keputusan medis.

Keywords : Kesehatan ibu hamil, klasifikasi, Decision Tree, Random Forest, SMOTE, machine learning.

ABSTRAK

The health of pregnant women is an important aspect in ensuring the safety of the mother and fetus. Early detection of health risks during pregnancy is needed to prevent complications that can lead to the death of the mother or baby. This study aims to analyze and compare the performance of the Decision Tree and Random Forest algorithms in the classification of health risks of pregnant women using the Maternal Health Risk dataset from the UCI Machine Learning Repository. The research process includes data preprocessing stages such as handling missing values, normalization, and outlier detection. In addition, unbalanced data was handled using the SMOTE method to ensure balanced class distribution. The model was evaluated using accuracy, precision, recall, F1-score, and k-fold cross-validation metrics. Experimental results show that the Random Forest algorithm has superior performance compared to Decision Tree, both on balanced and unbalanced data. Random Forest is able to provide higher accuracy (86.64%) and is stable and more resilient to data variations. This study concludes that Random Forest is more effectively used in the classification of health risks of pregnant women and can be a reliable alternative method in medical decision support systems. This study concludes that Random Forest is more effectively used in the classification of health risks of pregnant women and can be an alternative method that is a reliable method in medical decision support systems.

Keywords : Health of pregnant women, classification, Decision Tree, Random Forest, SMOTE, machine learning.