ABSTRACT

Multimorbidity is used to describe a condition in which a person suffers from several diseases or medical disorders, both chronic and acute, without any one disease being the primary focus. Diagnosing diseases in patients with multimorbidity is very complex because these patients have more than one disease at the same time. Machine learning can help predict disease diagnoses more quickly and efficiently. Multi-label classification is the appropriate approach because it allows for classification results of more than one label in the same object. This study aims to apply multi-label classification using the Binary Relevance method with the Naive Bayes algorithm to patients based on multimorbidity conditions and evaluate its performance. Binary Relevance works by applying a binary classifier to each label, then combining the outputs to form a set of predicted labels. The Naive Bayes algorithm is used in this study because it can determine parameters well in the classification process even when using relatively little training data. The dataset used in this study consisted of patient medical records based on multimorbidity conditions in 2024-2025 at the ABC Main Clinic, comprising 274 patients. This study uses six labels consisting of stroke, heart failure, ischemic heart disease, kidney failure, urinary tract infection, and typhoid fever. The feature variables in this study are continuous numeric and categorical, so Gaussian Naive Bayes and Categorical Naive Bayes are used. This study applied three types of training and test data divisions, namely with a proportion of 80%:20%, 70%:30%, and 60%:40%. The results show that the 80%:20% ratio provides the best overall performance with a hamming loss of 8,48%, micro precision of 84,95%, micro recall of 84,95%, and micro F1-score of 84,95%.

Keywords: Multimorbidity, Multi-Label Classification, Binary Relevance, Naive Bayes