

***Classification of Pneumonia Disease Using the Random Forest Algorithm  
(Case Study: Jember Lung Hospital)***

Bakhtiyar Hadi Prakoso, S.Kom., M.Kom (*Supervisor*)

**Dyah Nurfazdila**

*Health Information Management Study Program  
Department of Health*

***ABSTRACT***

*Pneumonia is an acute infection of the lung parenchyma caused by pathogens such as bacteria, viruses, fungi, and parasites, excluding Mycobacterium tuberculosis. Based on the report of the ten most common inpatient diseases at Jember Lung Hospital in 2023, pneumonia ranked first with a total of 2,426 cases. This study aimed to classify pneumonia using the Random Forest algorithm at Jember Lung Hospital. This research employed a quantitative approach using data mining techniques and was analyzed using RapidMiner software. The sample consisted of 372 inpatient medical records from 2024 with 13 variables: cough, sputum production, shortness of breath, chest pain, nausea and vomiting, fatigue, loss of appetite, fever, tachypnea, rhonchi, wheezing, signs of infection, and infiltrates. The results of parameter tuning with a 90:10 split ratio indicated that the optimal configuration of the Random Forest algorithm was achieved with  $n\_estimators = 9$  and  $max\_depth = 5$ . Model performance evaluation yielded an accuracy of 86.84%, precision of 85.00%, and recall of 89.47%. The modeling results showed that the infiltrate variable had the highest attribute weight (0.395). These findings indicate that pulmonary infiltrates are the primary indicator in pneumonia classification.*

***Keywords*** : *Confusion Matrix, Classification, Pneumonia, Random Forest*