Klasifikasi Penyakit Kanker Payudara Berdasarkan Berkas Rekam Medis Pasien Rawat Inap Menggunakan Algoritma C4.5 Di Rumah Sakit Baladhika Husada Jember. Classification of Breast Cancer Disease Based on Inpatient Medical Record Files Using the C4.5 Algorithm at Baladhika Husada Hospital Jember. Bakthtiyar Hadi Prakoso, S.Kom., M.Kom. (Pembimbing 1).

> Wa Ode Rezky Amalia Program Studi Manajemen Informasi Kesehatan Jurusan Kesehatan

ABSTRACT

Cancer is a disease that occurs due to abnormal growth of body tissue cells that turn into cancer cells during their development. These cancer cells can spread to other parts of the body and can cause death. Based on data on the top ten inpatient diseases at Baladhika Husada Jember Hospital, breast cancer ranked first with a total of 8,194 cases during the period 2021 to 2023. In addition, the mortality rate due to this disease shows an increasing trend, namely 15 deaths in 2021, 18 deaths in 2022, and reaching 20 deaths in 2023. This study aims to classify breast cancer based on risk factors and symptoms using the C4.5 algorithm, a data mining method. The modeling was conducted using RapidMiner software with a crossvalidation approach and model performance evaluated using a confusion matrix. The dataset consists of 302 records with 11 attributes. The results of the testing show that the C4.5 model achieved an accuracy of 78.15%, precision of 78.91%, and recall of 76.82%. The decision tree formed places the attribute of personal history of breast cancer as the root of the tree with the highest gain ratio of 0.1840 and produces 34 classification rules.

Keywords: C4.5, Breast Cancer, Cross Validation