CHAPTER 1
INTRODUCTION

1.1. Background

Breast cancer is a type of tumor that is found and develops aggressively in breast cells. Breast cancer is the most common type of cancer in women although the potential for male cancer can also occur but the chances are very small. This disease can occur at any age and therefore the risk will increase with age. [1][2]

Tumor cell growth becomes uncontrolled and will continue to change and divide more rapidly. So that normal cells cannot grow and can even die because they are pressed by the growth of tumor cells that continue to grow and are invasive. [3]

Breast cancer ranks initial and is one among the primary contributors to cancer deaths. Globocan 2020 knowledge states that quantity of recent cases of carcinoma reached 68.858 cases (16,6%) out of a complete of 396.914 new cases of cancer in Indonesia. Meanwhile, the quantity of deaths reached quite 22 thousand cases. [4]
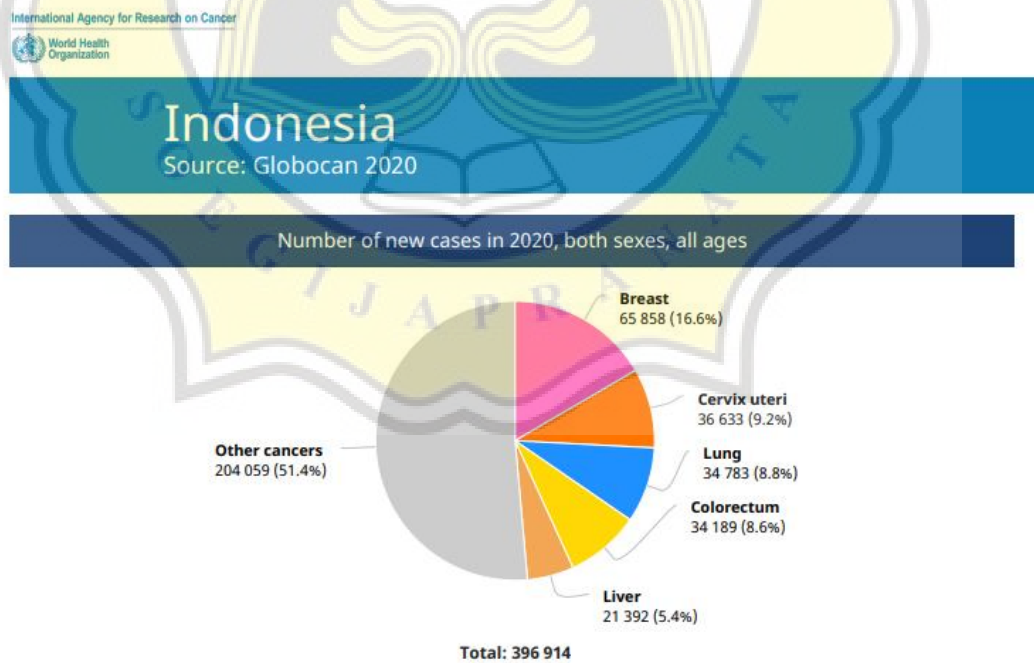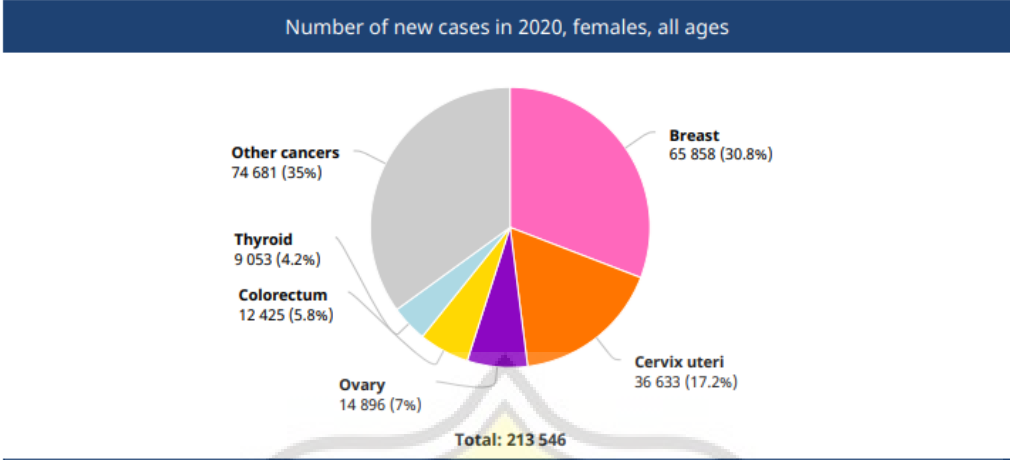

Figure 1.1 New Cases of Cancer

Figure 1.2 New Cases of Breast Cancer

Having a family history of breast cancer can increase your chances of getting this disease. Then factor in age, about 80% of breast cancer cases can occur at the age of over 50 years. Although it does not rule out this occurs in women at the age of the city. An unhealthy lifestyle can increase the risk of breast cancer such as smoking, lack of exercise, and an unhealthy diet that can lead to obesity. Then radiation exposure, the effect of hormone therapy being undertaken.

The application of a healthy lifestyle by doing physical activity, eating healthy food, can be done for prevention. [12]

The doctor will recommend a biopsy if abnormalities are found on the mammogram results. A breast biopsy may be a procedure for taking tissue samples from the breast to detect any abnormalities and to gauge whether the tissue is malignant or benign.

A system that will automatically predict recurring events on the results of carcinoma examinations will greatly assist doctors in making a decision. Naïve Bayes classifier takes advantage of opportunities by considering existing attributes.

The Naïve Bayes Classifier algorithm is one among the classifier which will predict the probability of a class of data that will fall into a specific class. This algorithm is often to research various things, one of which is to determine the type of breast cancer. Previously, there have been also those that had conducted similar research, but the attributes used to classify classes were different. In this study, a performance evaluation of the Naïve Bayes Classifier was built using the

Wisconsin UCI breast cancer datasets to detect and evaluate whether the tissue is malignant or benign.

1.2. Problem Formulation

When going to determine the type of cancer suffered by the patient, the doctors will read the result of the examination. From the results of the examination, it can be seen that there are differences in criteria if it includes tumors or cancer. Formulation of the problem is :

1. Can the application of the *Naïve Bayes* method determined whether the data is malignant or benign ?

1.3. Scope

As for limitations of the problem so as not to deviate from the aims and objectives, the limitations are as follows:

1. The method used is *Naïve Bayes*.

2. The data used is data from the *Wisconsin Breast Cancer Dataset* from *UCI Machine Learning Repository.*

3. The attributes used include *Sample code number, clump thickness, uniformity cell size, uniformity cell shape, marginal adhesion, single epithelial cell size, bare nucleoli, bland chromatin, normal nucleoli, mitoses,* and one attribute as a label, namely the *class* attribute with fields 2 for benign and 4 for malignant.

4. The final result that is expected from this research is to be able to classify the type of cancer based on the value obtained.

1.4. Objective

The purpose of this study was to build a breast cancer classifier on a UCI dataset that can accurately classify as benign or malignant using Naïve Bayes classifier.