

CHAPTER 6

CONCLUSION

The conclusion of this study is that the Naïve Bayes algorithm proposed as a classification method for breast cancer prediction is able to get a good score. The average percentage of data that is classified correctly reaches 96.48% and the average percentage of data that is classified incorrectly is only 3.52%. While the level of effectiveness where the average value of precision and recall respectively is 95.60% and 96.85%. For the highest precision and recall values, when the test data uses a 10% split percentage for testing data with values of 0.985 and 0.985.

The highest value obtained in the test is 98.505 by using a 10% split percentage. From several experiments, it can be seen that the predicted value of True Positive is smaller than the value of True Negative. This appropriate because the dataset used has a higher probability value for the Benign class of 65%. In this study, the probability value of the data distribution also effects the prediction results obtained. It can be seen in the confusion matrix value which tends to tend to a benign positive value, so that a higher True Negative is produced.

Suggestions for further research regarding breast cancer prediction is to use more than one algorithm in order to produce a more optimal value and the results can be compared. So that the advantages of each method can be analysed. Using other existing datasets. Using techniques for imbalanced data, so as to get more accurate results.