

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan terhadap dataset yang diperoleh dari *Coimbra UCI Machine Learning Repository kaggle* sebanyak 116 data dan 9 variabel seperti Usia, BMI, Glukosa, Insulin, HOMA, Leptin, Adiponektin, resistin, *MCP-1* yang dilakukan preprocessing untuk pemeriksaan missing value, dan duplikat data. Setelah dilakukan preprocessing maka tidak terdapat missing value atau duplikat pada data tersebut. Hasil pengujian yang dilakukan dengan metode *naive bayes* pada *dataset* kanker payudara menggunakan confusion matrix untuk data *training* 80%, data data *testing* 20% mendapatkan nilai akurasi sebesar 80%, Precision 100% dan Recall 67% menunjukkan bahwa model memiliki performa yang cukup baik dalam mengklasifikasi suatu data.

5.2 Saran

Dari kesimpulan diatas maka penulis memberikan saran yang diharapkan dapat memberikan manfaat bagi peneliti selanjutnya.

1. Pada penelitian ini masih kurangnya data yang digunakan, sehingga pada penelitian selanjutnya bisa menggunakan data baru untuk melakukan klasifikasi pada kanker payudara.
2. Menggunakan evaluasi model lain seperti K-Fol Cross- Validation yang

lebih stabil dan akurat, terutama pada dataset kecil.

3. Menggunakan metode *machine learning* lain sebagai perbandingan dengan metode naive bayes misalnya: metode Knn(K-Nearest Neighbor), Svm(Support Vector Machine), Clustering atau metode lain.