

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Berdasarkan hasil analisa dan pengujian yang telah dilakukan dengan menggunakan metode naive bayes, random forest, dan support vector machine dalam memprediksi penyakit jantung koroner, diperoleh beberapa kesimpulan sebagai berikut:

1. Pada proses penelitian menggunakan metode naive bayes dataset dibagi menjadi 3 bagian yaitu :
  - a. Data training 80% dan data testing 20% untuk orang yang tidak terkena penyakit jantung memperoleh hasil nilai presisi 87%, recall 90%, f1-score 88%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 90%, recall 88%, f1-score 89%. Untuk nilai akurasi sebesar 89%.
  - b. Data training 70% dan data testing 30% untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 85%, recall 83%, f1-score 84%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 86%, recall 88%, f1-score 87%. Untuk nilai akurasi sebesar 86%.
  - c. Data training 60% dan data testing 40% untuk orang yang tidak terkena penyakit jantung memperoleh hasil nilai presisi 85%, recall 85%, f1-score 85%. Sedangkan untuk yang terkena penyakit jantung memperoleh hasil nilai presisi 88%, recall 88%, f1-score 88%. Untuk nilai akurasi sebesar 87%.
2. Pada proses penelitian menggunakan metode random forest dataset dibagi menjadi 3 bagian yaitu :
  - a. Data training 80% dan data testing 20% untuk orang yang tidak terkena penyakit jantung memperoleh hasil nilai presisi 86%, recall 83%, f1-score 84%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 85%, recall 88%, f1-score 86%. Untuk nilai akurasi sebesar 85%.
  - b. Data training 70% dan data testing 30% untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 80%, recall 79%, f1-score 80%. Sedangkan untuk

orang yang terkena penyakit jantung memperoleh hasil nilai presisi 82%, recall 83%, f1-score 83%. Untuk nilai akurasi sebesar 81%.

- c. Data training 60% dan data testing 40% untuk orang yang tidak terkena penyakit jantung memperoleh hasil nilai presisi 80%, recall 83%, f1-score 81%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 87%, recall 84%, f1-score 84%. Untuk nilai akurasi sebesar 83
3. Pada proses penelitian menggunakan metode support vector machine dataset dibagi menjadi 3 bagian yaitu :
    - a. Data training 80% dan data testing 20% untuk orang yang tidak terkena penyakit jantung memperoleh hasil nilai presisi 89%, recall 83%, f1-score 86%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 85%, recall 91%, f1-score 88%. Untuk nilai akurasi sebesar 87%.
    - b. Data training 70% dan data testing 30% untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 84%, recall 76%, f1-score 80%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 81%, recall 88%, f1-score 84%. Untuk nilai akurasi sebesar 82%.
    - c. Data training 60% dan data testing 40% untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 84%, recall 81%, f1-score 82%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 86%, recall 88%, f1-score 87%. Untuk nilai akurasi sebesar 85%.
  4. Pada proses penelitian menggunakan metode logistic regression dataset dibagi menjadi 3 bagian yaitu :
    - a. Data training 80% dan data testing 20% untuk orang yang tidak terkena penyakit jantung memperoleh hasil nilai presisi 89%, recall 86%, f1-score 88%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 88%, recall 91%, f1-score 89%. Untuk nilai akurasi sebesar 89%.
    - b. Data training 70% dan data testing 30% untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 87%, recall 79%, f1-score 82%. Sedangkan untuk

orang yang terkena penyakit jantung memperoleh hasil nilai presisi 83%, recall 90%, f1-score 86%. Untuk nilai akurasi sebesar 85%.

- c. Data training 60% dan data testing 40% untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 84%, recall 79%, f1-score 81%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 85%, recall 88%, f1-score 87%. Untuk nilai akurasi sebesar 84%.

5. Pada proses penelitian menggunakan metode decision tree dataset dibagi menjadi 3 bagian yaitu :

- a. Data training 80% dan data testing 20% untuk orang yang tidak terkena penyakit jantung memperoleh hasil nilai presisi 69%, recall 86%, f1-score 77%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 64%, recall 66%, f1-score 74%.. Untuk nilai akurasi sebesar 75%.

- b. Data training 70% dan data testing 30% untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 75%, recall 86%, f1-score 80%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 86%, recall 76%, f1-score 80%. Untuk nilai akurasi sebesar 80%.

- c. Data training 60% dan data testing 40% untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 76%, recall 81%, f1-score 79%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 85%, recall 81%, f1-score 83%. Untuk nilai akurasi sebesar 81%.

6. Pada proses penelitian menggunakan metode gradient boosting dataset dibagi menjadi 3 bagian yaitu :

- a. Data training 80% dan data testing 20% untuk orang yang tidak terkena penyakit jantung memperoleh hasil nilai presisi 77%, recall 83%, f1-score 80%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 83%, recall 78%, f1-score 81%.. Untuk nilai akurasi sebesar 80%.

- b. Data training 70% dan data testing 30% untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 78%, recall 90%, f1-score 84%. Sedangkan untuk

orang yang terkena penyakit jantung memperoleh hasil nilai presisi 90%, recall 78%, f1-score 84%. Untuk nilai akurasi sebesar 84%.

- c. Data training 60% dan data testing 40% untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 75%, recall 85%, f1-score 79%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 87%, recall 78%, f1-score 82%. Untuk nilai akurasi sebesar 81%.

7. Pada proses penelitian menggunakan validasi silang keenam metode mendapatkan hasil sebagai berikut :

- a. Metode naive bayes pada proses menggunakan validasi silang untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 84%, recall 93%, f1-score 89%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 93%, recall 84%, f1-score 89%. Untuk nilai akurasi sebesar 89%.
- b. Metode random forest pada proses menggunakan validasi silang untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 84%, recall 90%, f1-score 87%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 90%, recall 84%, f1-score 87%. Untuk nilai akurasi sebesar 87%.
- c. Metode svm pada proses menggunakan validasi silang untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 92%, recall 83%, f1-score 87%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 85%, recall 94%, f1-score 89%. Untuk nilai akurasi sebesar 88%.
- d. Metode Logistic regression pada proses menggunakan validasi silang untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 89%, recall 83%, f1-score 86%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 85%, recall 90%, f1-score 88%. Untuk nilai akurasi sebesar 87%.
- e. Metode decision tree pada proses menggunakan validasi silang untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 92%, recall 79%, f1-score 85%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh

hasil nilai presisi 83%, recall 94%, f1-score 88%. Untuk nilai akurasi sebesar 87%.

- f. Metode gradient boosting pada proses menggunakan validasi silang untuk orang yang tidak terkena penyakit memperoleh hasil nilai presisi 81%, recall 90%, f1-score 85%. Sedangkan untuk orang yang terkena penyakit jantung memperoleh hasil nilai presisi 90%, recall 81%, f1-score 85%. Untuk nilai akurasi sebesar 85%.
8. Dari pengujian enam metode seperti naive bayes, random forest, support vector machine, logistic regression, decision tree, dan gradient boosting dengan pembagian dataset menjadi 3 kelompok yaitu data training 80% dan data testing 20%, data training 70% dan data testing 30, dan data training 60% dan data testing 40%. Dapat diambil kesimpulan bahwa metode naive bayes dan logistic regression lebih baik dalam pengujian data latih 80% dan data uji 20%. Sedangkan untuk pengujian dalam data latih 70% dan data uji 30% serta data latih 60% dan data uji 40% metode naive bayes lebih baik daripada metode random forest, svm, logistic regression, decision tree, dan gradient boosting. Berdasarkan pengujian menggunakan validasi silang dapat diambil kesimpulan metode naive lebih unggul dari metode svm, random forest, logistic regression, decision tree, dan gradient boosting.

## 5.2 Saran

Adapun saran untuk pengembangan penelitian selanjutnya adalah sebagai berikut

:

1. Diharapkan data yang digunakan dalam penelitian selanjutnya bisa lebih banyak dan lebih beragam untuk dilakukan pengujian.
2. Untuk penelitian selanjutnya diharapkan mampu meningkatkan akurasi yang lebih tinggi dalam memprediksi penyakit jantung koroner.
3. Diharapkan penelitian selanjutnya menggunakan lebih banyak model algoritma untuk dilakukan perbandingan.