

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil dari penelitian mengenai penerimaan peserta didik baru, dapat disimpulkan bahwa klasifikasi penerimaan peserta didik baru dengan memanfaatkan teknik *data mining* dan menganalisis kinerja algoritma *C4.5* dan *Naïve Bayes* mendapatkan hasil yang signifikan. Melalui percobaan perhitungan menggunakan *Google Colab* dan bahasa pemrograman *Python*, dengan data sebanyak 187 dan dibagi menjadi *training* 80% (149 data) dan *testing* 20% (38 data). Dengan perolehan algoritma *C4.5* mencapai tingkat akurasi sebesar 97,3%, sementara algoritma *Naïve Bayes* mencapai 92,1%. Kedua algoritma ini memberikan hasil akurasi yang baik, namun algoritma *C4.5* menunjukkan kinerja yang lebih unggul dibandingkan dengan algoritma *Naïve Bayes*.

5.2 Saran

Penulis menyadari banyak kekurangan dan keterbatasan selama proses penelitian ini. Penulis menyarankan beberapa metode untuk pengembangan tulisan ini yang lebih baik:

- a) Penelitian ini melakukan pembagian data secara manual dengan menggunakan data *training* sebesar 80% dimulai dari data teratas dan data *testing* 20% dari data terbawah. Menurut penelitian yang dilakukan oleh Wildan Attariq Firmansyach, Umi Hayati, dan Yudhistira Arie Wijaya implementasi metode *K-Fold Cross Validation* dapat membagi data secara acak dan memperbaiki masalah *Overfitting* dan *Underfitting* sehingga dalam proses perhitungan memperoleh model yang lebih akurat[26].
- b) Penelitian ini menggunakan 187 *dataset* dan memberikan nilai akurasi algoritma *C4.5* 97,3% dan *Naïve Bayes* 92,1%. Namun menurut penelitian yang dilakukan oleh Nurfazriah Attami, Agung Triayudi, dan Rima Tamara Aldisa dengan penambahan jumlah *dataset* dapat menghasilkan akurasi yang lebih baik. Pada penelitian tersebut menggunakan *dataset* 1.043 memperoleh nilai akurasi untuk algoritma *C4.5* 99,04% dan *Naïve Bayes* 92,9%[27].