

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Kesimpulan Penelitian Komparasi Algoritma *Machine learning* Menggunakan Data klaim asuransi mobil Berdasarkan penelitian ini, dapat diambil kesimpulan sebagai berikut:

1. Model klasifikasi yang berhasil dibangun menggunakan RapidMiner dengan algoritma *Decision Tree*, *K-Nearest Neighbor*, *Naïve Bayes*, *Random Forest*, dan *Logistic Regression*. Model-model ini bertujuan untuk memprediksi klaim asuransi dengan menggunakan data klaim asuransi yang tersedia. Dengan menggunakan RapidMiner, proses pembangunan model klasifikasi menjadi lebih mudah dan dapat menghasilkan model yang dapat memprediksi klaim asuransi.
2. Performa algoritma *Decision Tree*, *K-Nearest Neighbor*, *Naïve Bayes*, *Random Forest*, dan *Logistic Regression* telah dibandingkan dalam memprediksi klaim asuransi. Evaluasi performa dilakukan menggunakan matrik evaluasi seperti akurasi, AUC, *recall*, dan *precision*. Berdasarkan hasil evaluasi, dapat diidentifikasi algoritma yang memberikan performa baik dalam memprediksi klaim asuransi adalah algoritma *K-Nearest Neighbor* dan algoritma *Random Forest* dengan hasil evaluasi akurasi algoritma *K-Nearest Neighbor* (90.46%), *recall Yes* (99.03%), *recall No* (81.88%), *precision Yes* (84.53), *Precision No* (98.83%) dan AUC (0.984) kemudian algoritma *Random Forest* (90.26%), *recall Yes* (97.25%), *recall No* (83.28%), *precision Yes* (85.33%), *Precision No* (96.80%) dan AUC (0.980), yang dapat digunakan untuk pengambilan keputusan yang lebih baik dalam industri asuransi kendaraan.
3. Berdasarkan visualisasi perbandingan performa yang dilakukan, dapat disimpulkan bahwa dalam memprediksi klaim asuransi kendaraan, algoritma *Random Forest* (RF) dan *K-Nearest Neighbors* (K-NN) menunjukkan hasil yang paling unggul dibandingkan dengan algoritma-algoritma lain yang dievaluasi. Dua algoritma ini memperoleh skor yang lebih tinggi dan performa

yang lebih baik dalam membedakan antara klaim yang sah dan klaim yang curang. Oleh karena itu, *Random Forest* (RF) dan *K-Nearest Neighbors* (K-NN) dapat dianggap sebagai pilihan algoritma yang optimal dan efektif dalam melakukan prediksi terkait klaim asuransi kendaraan.

5.2 Saran

Setelah melakukan penelitian ini, berikut adalah beberapa saran untuk penelitian selanjutnya:

1. Lima algoritma *machine learning* telah dibandingkan. Namun, ada banyak algoritma lain yang dapat dieksplorasi. Penelitian selanjutnya dapat memperluas jangkauan algoritma yang dibandingkan. Pastikan dalam memilih algoritma dapat mengatasi *noise* atau data yang tidak konsisten dan tidak terpengaruh oleh fitur atau atribut yang tidak memiliki hubungan seperti *Support Vector Machines*, *Adaboost* dan *Bagging*.
2. Penelitian ini mungkin telah menggunakan fitur-fitur yang relevan untuk memprediksi klaim asuransi kendaraan. Penelitian selanjutnya dapat memberikan wawasan tambahan atau memperbaiki performa prediksi dengan menggunakan *feature selection* seperti *correlation*, *information* dan *information gain*.

