

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Dari hasil eksplorasi data terhadap dataset Titanic dapat diketahui bahwa dari 891 penumpang, penumpang yang selamat sebanyak 342 orang dan memiliki peluang selamat sekitar 38%. Dari penumpang yang selamat, jumlah penumpang perempuan yang selamat cukup tinggi yaitu 233 orang dibandingkan dengan laki-laki sebanyak 109 orang. Peluang selamat perempuan lebih tinggi daripada laki-laki yaitu 74.20%. Peluang selamat tertinggi terdapat pada penumpang kelas 1, perempuan, penumpang yang berpergian dengan keluarga, dan berusia antara 0 – 11 tahun. Peluang selamat tertinggi terjadi pada penumpang perempuan dan berasal dari kelas 1.

Faktor eksternal (di luar dataset Titanic) juga mempengaruhi kelangsungan hidup. Material kapal yang digunakan adalah baja dengan kandungan oksigen dan sulfur yang tinggi sehingga menyebabkan peningkatan kerapuhan dan membentuk magnesium sulfida yang memfasilitasi perambatan retak. Suhu air yang rendah dan beban benturan yang tinggi membuat paku keling pada pelat lambung tidak berfungsi dengan baik dan berkontribusi pada kegagalan baja lambung kapal dan tenggelamnya Titanic. Jumlah sekoci yang hanya menampung 1.178 dari total penumpang 2.240 serta tidak memadainya prosedur darurat menyebabkan banyak orang yang tenggelam. Badai geomagnetik yang terjadi di Atlantik Utara mungkin berperan dalam tragedi Titanic baik secara langsung maupun tidak langsung yang menyebabkan kesalahan perhitungan posisi SOS Titanic.

Berdasarkan hasil evaluasi dan perbandingan terhadap empat model yaitu Logistic Regression, Decision Tree, Random Forest, dan Naive Bayes, model Decision Tree dipilih sebagai model terbaik untuk memprediksi kelangsungan hidup penumpang Titanic. Pemilihan tersebut berdasarkan tingkat akurasi yang tinggi, presisi yang baik dalam mengidentifikasi

penumpang yang benar-benar selamat, serta nilai konsistensi performa yang tinggi. Dengan demikian, model Decision Tree tidak hanya memberikan prediksi yang tepat secara signifikan, tetapi juga terbukti konsisten dan dapat diandalkan dalam berbagai skenario data.

5.2. Saran

Berdasarkan penelitian yang dilakukan, adapun beberapa saran yang diberikan untuk penelitian selanjutnya adalah sebagai berikut:

1. Menggunakan model *machine learning* yang kompleks seperti *Neural Network*, *XGBoost*, atau *Gradient Boosting*. Model tersebut dapat menangani kompleksitas data yang dapat memberikan hasil yang akurat.
2. Menggunakan teknik *feature engineering* yang lebih baik seperti pemilihan fitur dengan metode RFE (*Recursive Feature Elimination*).
3. Melakukan evaluasi dengan metrik tambahan untuk mendapatkan gambaran yang lebih baik tentang kinerja model
4. Menerapkan teknik penanganan data yang tidak seimbang untuk memperbaiki distribusi kelas yang tidak seimbang dan meningkatkan kinerja model dalam mendeteksi kelas minoritas