

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai penerapan algoritma Naïve Bayes untuk mengklasifikasikan potensi keterlambatan pengiriman pada ekspedisi JNE Sariharjo, Sleman, Yogyakarta, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Langkah penerapan algoritma Naïve Bayes dalam penelitian ini dimulai dari proses pengumpulan data historis pengiriman dari database JNE Sariharjo pada periode 1–24 September 2024. Data kemudian diproses melalui tahapan text preprocessing yang meliputi cleaning data, penanganan missing values, pelabelan kelas (“Tepat Waktu” dan “Terlambat”), serta categorical encoding. Setelah data siap, dilakukan proses training dan testing model Naïve Bayes menggunakan bahasa pemrograman PHP. Proses prediksi dilakukan dengan menghitung probabilitas posterior dari tiap kelas berdasarkan fitur input seperti tujuan, layanan, barang, berat, dan jumlah barang, lalu hasil prediksi ditampilkan dalam sistem berbasis web.
2. Faktor-faktor yang memengaruhi akurasi prediksi keterlambatan pengiriman di antaranya adalah:
 - a) Jenis layanan (servis): Berpengaruh signifikan terhadap estimasi waktu pengiriman. Layanan YES cenderung lebih cepat dibandingkan REG, OKE, atau JTR.
 - b) Berat dan jumlah barang: Fitur numerik ini dipertimbangkan dalam model dengan pendekatan probabilistik berdasarkan deviasi dari nilai rata-rata. Misalnya, berat mendekati rata-rata kelas tertentu akan menghasilkan probabilitas lebih tinggi terhadap kelas tersebut.
 - c) Tujuan dan jenis barang: Kategori tujuan pengiriman dan jenis barang juga memberikan pengaruh terhadap kecenderungan pengiriman tepat waktu atau terlambat. Frekuensi kemunculan fitur dalam kelas “terlambat” maupun “tepat waktu” mempengaruhi nilai likelihood dalam perhitungan probabilitas.

3. Keakuratan model prediksi berbasis Naïve Bayes menunjukkan performa yang cukup baik. Berdasarkan hasil evaluasi dari confusion matrix terhadap 827 data pengiriman, diperoleh hasil:
 - a) Akurasi: 91,66%
 - b) Precision: 85,81%
 - c) Recall: 73,89%
 - d) F1-Score: 79,40%

Nilai-nilai ini membuktikan bahwa model Naïve Bayes yang diterapkan mampu mengklasifikasikan status pengiriman dengan akurasi tinggi, meskipun terdapat sedikit penurunan recall karena ketidakseimbangan distribusi kelas “terlambat” yang jauh lebih sedikit dibandingkan “tepat waktu”.

5.2 Saran

Untuk pengembangan lebih lanjut, berikut beberapa saran yang dapat dijadikan pertimbangan:

1. Penambahan variabel prediktor, seperti kondisi cuaca, kepadatan lalu lintas, serta jarak pengiriman, diharapkan dapat meningkatkan akurasi dan sensitivitas model dalam memprediksi keterlambatan.
2. Penerapan teknik balancing data, seperti *SMOTE* atau *undersampling*, direkomendasikan untuk mengatasi ketidakseimbangan data antara kelas “Tepat Waktu” dan “Terlambat”, agar model lebih adil dalam memberikan klasifikasi.
3. Penggunaan data yang lebih banyak dan beragam dapat membantu model mengenali pola keterlambatan secara lebih luas serta mengurangi dominasi kelas mayoritas.
4. Validasi model menggunakan metode k-fold cross validation atau split test yang terpisah dari data pelatihan awal dapat memberikan gambaran performa model yang lebih obyektif dan menghindari overfitting.