

1. Pendahuluan

Transportasi memiliki peran yang sangat penting dalam menunjang pertumbuhan ekonomi masyarakat dan merupakan urat nadi dalam pembangunan ekonomi suatu negara. Keberhasilan pembangunan di bidang ekonomi harus ditunjang dengan pengembangan sistem transportasi yang baik, sesuai dengan kebutuhan dan perkembangan zaman. Melalui transportasi kita bisa berpergian, mengirim barang, meningkatkan nilai ekonomi suatu wilayah, hingga menjaga stabilitas harga barang. Hadirnya aplikasi GoJek, besutan anak negeri yang menjadikan jasa transportasi sebagai aktivitas utama dari perusahaannya juga membuat sepeda motor sebagai salah satu moda transportasinya semakin jauh melangkah dan membuat transportasi berperan semakin penting dalam menunjang pertumbuhan ekonomi masyarakat.

Sepeda motor sebagai salah satu moda transportasi dapat menjadi sumber mata pencaharian bagi sebagian orang, dan ongkos untuk mengirim barang kini lebih murah karena perusahaan atau bisnis baru tidak perlu membeli sepeda motor dan menambah asset untuk bidang logistik. Pelaku bisnis dapat dengan mudah mengantarkan barangnya dan sampai di hari yang sama. Berpergian pun menjadi lebih mudah dan tidak perlu khawatir lagi harus menunggu lama. Meski semua sudah serba cepat dan instan, sebuah mesin tetap perlu perawatan. Mesin sepeda motor yang rusak karena kurangnya perawatan dapat mengganggu roda perekonomian. Perawatan yang sudah dilakukan pun kadang masih belum cukup untuk meminimalisasi kemungkinan mesin rusak saat di perjalanan. Terbukti dengan data yang didapat dari KOMINFO, bahwa setiap jam rata-rata 3 orang meninggal akibat kecelakaan jalan di Indonesia, dengan 9% kecelakaan terjadi karena faktor kendaraan^[1], dan dari awal tahun hingga Februari 2020 jumlah kecelakaan lalu lintas pada sepeda motor mencapai angka 144 kasus dan menjadi kasus terbanyak dibanding dengan kendaraan lain^[2]. Solusi yang tersedia saat ini pun seperti bengkel, belum mampu untuk mengatasi rusaknya mesin atau kendala lain yang dialami oleh sepeda dan pengendara motor, karena ada kemungkinan jarak yang jauh dari tempat kendala tersebut terjadi.

Solusi yang cepat dan tepat sangat dibutuhkan, namun tidak semua pengendara motor mengerti kondisi sepeda motornya dan mampu mengatasi masalah tersebut seorang diri. Berangkat dari kasus ini, maka diperlukan sistem klasifikasi keluhan mengenai mesin dan non mesin sepeda motor sebagai alternatif dari solusi yang ada. Dengan klasifikasi pertanyaan mengenai sepeda motor ini, banyak hal yang mampu dikembangkan, seperti chatbot atau bot pintar. Sistem klasifikasi tersebut dapat dibangun menggunakan klasifikasi teks.

Terkait studi klasifikasi teks, Pada tahun 2018 Baygin^[3] juga melakukan penelitian mengenai klasifikasi dokumen teks berdasarkan Naïve Bayes menggunakan fitur N-Gram, dimana dalam penelitiannya klasifikasi teks dokumen menggunakan Naïve Bayes mendapat tingkat akurasi sebesar 92%. Pada tahun yang sama dengan Baygin, Venkatesh^[4] dan koleganya melakukan skema klasifikasi dan optimasi untuk data teks menggunakan *machine learning Naïve Bayes classifier* dan mendapatkan penggunaan naïve bayes dalam klasifikasi teks lebih baik dibandingkan klasifikasi teks menggunakan K-Nearest Neighbor. Mengikuti kesuksesan Baygin menggunakan Naïve bayes sebagai pengklasifikasi teks, Dalibor^[5] menggunakan Naïve Bayes sebagai klasifikasi lirik lagu, dari 207 lagu didapatkan hasil prediksi yang menjanjikan dengan presisi sebesar 93% dan F-Measure sebesar 94%. Penelitian serupa juga dilakukan oleh Rahman^[6] dan membandingkan Decision Tree, K-NN dan Multinomial Naïve Bayes untuk klasifikasi topik dan dari ketiga pengklasifikasian tersebut, didapatkan hasil akurasi terbaik sebesar 91.8% untuk Naïve Bayes.

Pada tahun 2018 Ardy Wibowo^[7] dan koleganya melakukan penelitian untuk mengetahui pengaruh normalisasi kata dan pemilihan fitur chi-squared pada klasifikasi teks SVM dan mendapatkan 95.05% pada akurasi ketika performansi stemming ditingkatkan oleh chi-squared. Patrick Adolf^[8] dan koleganya juga melakukan penelitian untuk membandingkan beberapa metode klasifikasi pada machine learning dalam kasus berbasis teks di Twitter dan mendapatkan akurasi dari metode klasifikasi SVM sebagai metode terbaik dibandingkan 4 metode lainnya dengan akurasi sebesar 91.7%. Andreas Chandra^[9] melakukan penelitian untuk membandingkan seleksi fitur untuk dataset teks yang tidak seimbang dan didapatkan hasil akurasi SVM sebesar 89.07% pada saat fitur diseleksi dengan menggunakan chi-squared. Selanjutnya, Zun Hlaing^[10] dan koleganya pada tahun 2018 juga membandingkan Naïve bayes dan SVM pada klasifikasi dokumen dan mendapatkan SVM lebih baik dibandingkan dengan Naïve Bayes pada saat klasifikasi dokumen.

Pada penelitian ini, model prediksi dibangun untuk memprediksi pertanyaan dan keluhan mengenai sepeda motor pada kasus ini, keluhan pengguna sepeda motor dibatasi menjadi dua kelas, yaitu keluhan terkait mesin dan non mesin. Model prediksi dibangun dengan menggunakan metode Naïve bayes dan Support Vector Machine. Dataset yang digunakan pada penelitian ini adalah data pertanyaan dan keluhan dari pengendara motor yang mengalami masalah pada motornya.