

BAB V

KESIMPULAN DAN SARAN

5.1 KESIMPULAN

Berdasarkan hasil penelitian dan implementasi sistem *chatbot* berbasis Natural Language Processing (NLP) untuk pelayanan administrasi MBKM di Institut Informatika dan Bisnis Darmajaya, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Penelitian ini telah berhasil membangun dan mengimplementasikan sebuah sistem *chatbot* berbasis web yang dapat diakses untuk memberikan informasi seputar administrasi MBKM. Sistem ini menggunakan algoritma Multinomial Naive Bayes sebagai metode klasifikasi untuk mengenali maksud (*intent*) dari pertanyaan yang diajukan oleh pengguna.
2. Ada 20 *class intent* yang dibangun. Berdasarkan hasil evaluasi, salah satu temuan krusial adalah adanya indikasi overfitting yang signifikan pada model yang dibangun. Hal ini terbukti dari disparitas yang sangat besar antara akurasi pada data latih (*training accuracy*) yang mencapai 93.62%, dengan akurasi pada data uji (*testing accuracy*) yang jauh lebih rendah, yaitu sebesar 60.56%. Kondisi ini menunjukkan bahwa meskipun model mampu mengenali pola pada data latih dengan sangat baik, kemampuannya untuk melakukan generalisasi terhadap data baru masih sangat terbatas.
3. *Chatbot* yang dibangun telah diintegrasikan dengan Flask API dan diimplementasikan pada antarmuka berbasis web dalam bentuk mockup landing page yang memuat informasi utama MBKM serta dilengkapi fitur *chatbot* interaktif. Sistem ini telah diuji untuk mengetahui *user acceptance* yang melibatkan mahasiswa dan staf administrasi MBKM, dan terbukti dapat berfungsi sesuai tujuan awal, yaitu memberikan layanan informasi MBKM secara cepat, responsif, dan aktif selama 24 jam.

5.2 SARAN

Adapun beberapa saran yang dapat diberikan untuk pengembangan lebih lanjut dari penelitian ini adalah sebagai berikut:

1. Meningkatkan jumlah dan variasi data pelatihan (training data) agar model klasifikasi intent dapat mengenali lebih banyak jenis pertanyaan dengan akurasi yang lebih tinggi. Hal ini dapat dilakukan dengan mengumpulkan data pertanyaan dari interaksi pengguna secara langsung.
2. Ketidakseimbangan dataset antar-intent dapat menyebabkan bias pada model dan mempengaruhi akurasi prediksi. Oleh karena itu, pada penelitian selanjutnya disarankan untuk melakukan teknik oversampling, undersampling, atau class weighting agar distribusi data antar-intent menjadi lebih seimbang.
3. Pemilihan parameter alpha yang kurang optimal pada validation curve dapat memicu terjadinya overfitting atau underfitting. Penelitian berikutnya disarankan untuk mengeksplorasi rentang nilai alpha yang lebih luas serta mengombinasikan pencarian parameter menggunakan metode seperti Grid Search atau Random Search.
4. Akurasi model yang masih terbatas (60,56%) pada data uji menunjukkan perlunya pengayaan data latih, khususnya untuk intent yang jarang muncul, agar kemampuan generalisasi model meningkat. Teknik data augmentation berbasis NLP, seperti parafrase otomatis atau penerjemahan bolak-balik (back translation), dapat digunakan untuk menambah variasi pertanyaan pada dataset. Penggunaan model berbasis deep learning seperti LSTM atau BERT juga dapat dipertimbangkan untuk meningkatkan akurasi prediksi intent.