

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian klasifikasi sentimen terhadap opini Generasi Z mengenai dunia kerja menggunakan tiga pendekatan model Random Forest, Support Vector Machine, dan BERT, dapat disimpulkan sebagai berikut:

1. Penelitian ini telah berhasil mengimplementasikan tiga algoritma, yaitu SVM, Random Forest, dan BERT, untuk mengklasifikasikan sentimen tweet terkait dunia kerja Generasi Z ke dalam tiga label: positif, netral, dan negatif. Proses klasifikasi diawali dengan tahapan pra-pemrosesan data, pembobotan fitur (TF-IDF untuk SVM dan RF, serta tokenisasi kontekstual untuk BERT), dan penyeimbangan kelas menggunakan Random Over Sampling (ROS). Setiap model kemudian dievaluasi menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil evaluasi menunjukkan bahwa ketiga algoritma mampu diterapkan secara efektif untuk tugas klasifikasi sentimen, dengan performa terbaik ditunjukkan oleh model SVM.
2. Model SVM memiliki performa lebih unggul dengan akurasi 90,22%. Pada kelas negatif, SVM berhasil dengan precision 95%, recall 82%, dan F1-score 88%, menunjukkan keakuratan tinggi dalam mendeteksi opini negatif meski beberapa opini negatif terlewat. Pada kelas netral, model mencatat precision 87%, recall sangat tinggi 95%, dan F1-score 90%, mengindikasikan kemampuan yang sangat baik dalam mengenali opini netral. Untuk kelas positif, SVM mencatat precision 93%, recall 87%, dan F1-score 90%, menandakan kemampuan yang akurat dan konsisten dalam mengidentifikasi sentimen positif.
3. Model Random Forest menunjukkan performa dengan akurasi sebesar 80,82%. Untuk kelas negatif, model ini mencatat precision 83%, recall 68%, dan F1-score 75%, yang berarti sebagian besar prediksi negatif tepat namun ada beberapa opini negatif yang belum terdeteksi. Pada kelas

netral, Random Forest mencapai precision 79%, recall 87%, dan F1-score 83%, menunjukkan kemampuan yang baik dalam mengenali opini netral. Sedangkan pada kelas positif, model memperoleh precision 83%, recall 78%, dan F1-score 80%, menunjukkan performa yang cukup seimbang dalam mengklasifikasi sentimen positif.

4. Model BERT (indobenchmark/indobert-base-p2) menunjukkan performa yang sangat baik dengan akurasi sebesar 88,38%. Untuk kelas negatif, BERT menghasilkan precision, recall dan juga F1-score di angka 84%, menandakan bahwa meskipun prediksi opini negatif sangat tepat, sebagian opini dengan nada negatif masih gagal dikenali. Pada kelas netral, model mencapai precision 91%, recall 88%, dan F1-score 89%, yang menunjukkan bahwa model cukup stabil dalam menangkap opini mayoritas. Sedangkan untuk kelas positif, diperoleh precision sebesar 87%, recall 93%, dan F1-score 90%, yang menandakan bahwa BERT mampu memahami opini positif dengan baik, bahkan dalam bentuk ekspresi yang bervariasi dan informal.
5. Secara keseluruhan, SVM menjadi model dengan performa terbaik dalam klasifikasi sentimen opini Gen Z terhadap dunia kerja, dengan akurasi tertinggi sebesar 90,22% dan hasil yang seimbang di semua kelas. BERT menyusul dengan akurasi 88,38%, menunjukkan kemampuan kuat dalam menangani bahasa informal, terutama dalam mengenali opini positif dan netral, meskipun masih ada kekeliruan dalam kelas negatif. Sementara itu, Random Forest memiliki performa paling rendah dengan akurasi 80,82%, khususnya kurang efektif dalam mendeteksi opini negatif. Dengan demikian, SVM dinilai paling optimal, diikuti oleh BERT dan Random Forest.

5.2 Saran

Berdasarkan temuan dari penelitian ini, beberapa saran yang dapat diberikan untuk penelitian serta dapat diimplementasikan dengan penerapan lebih lanjut adalah sebagai berikut :

1. Optimalisasi Model BERT

Penelitian ini menemukan adanya indikasi overfitting pada model BERT, yang ditandai dengan selisih performa antara data pelatihan dan validasi. Untuk mengatasi hal ini, disarankan agar penelitian selanjutnya menggunakan dataset yang lebih besar dan bervariasi agar model dapat belajar secara lebih umum dan tidak hanya menyesuaikan diri dengan data pelatihan. Selain itu, penggunaan teknik *data augmentation* seperti *synonym replacement* atau *back translation* dapat menambah keragaman data tanpa perlu menambah data secara manual. Strategi regularisasi tambahan seperti *weight decay* dan *layer freezing* juga dapat membantu mengurangi kompleksitas model dan meningkatkan stabilitas selama pelatihan. Penerapan langkah-langkah tersebut diharapkan mampu meningkatkan generalisasi model dan mengurangi risiko overfitting dalam analisis sentimen berbasis media sosial.

2. Pengembangan Teknik Pra-Pemrosesan Teks yang Lebih Efektif

Penelitian selanjutnya disarankan untuk mengeksplorasi lebih dalam teknik preprocessing yang dapat menangani teks informal dengan lebih baik. Hal ini mencakup normalisasi kata tidak baku, ekspansi singkatan, penanganan emotikon dan emoji, serta *spelling correction*. Upaya ini diharapkan dapat membantu meningkatkan akurasi model, terutama dalam mendeteksi opini negatif yang kerap ditulis dengan ekspresi tidak langsung.

3. Pelabelan Data dengan Anotator Manusia

Untuk penelitian selanjutnya, disarankan agar proses pelabelan data dilakukan dengan melibatkan anotator manusia guna meningkatkan akurasi label sentimen, khususnya pada data dengan ekspresi yang ambigu dan sesuai dengan perkembangan bahasa Gen Z di masa sekarang. Alternatif lain adalah dengan mengembangkan metode pelabelan otomatis berbasis model deep learning yang dilatih pada data kontekstual, sehingga mampu memahami nuansa bahasa secara lebih menyeluruh. Pendekatan

ini diharapkan dapat menghasilkan data latih yang lebih representatif dan meningkatkan performa model klasifikasi sentimen secara keseluruhan.

