

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil yang diperoleh dari kedua metode tersebut dapat disimpulkan bahwa

1. Model LSTM:

- Model LSTM memiliki tingkat akurasi sebesar 97.73%, yang menunjukkan kemampuan model dalam mengklasifikasikan data dengan benar.
- Presisi model sebesar 95.80% menunjukkan bahwa model cenderung memberikan sedikit jumlah false positive (klasifikasi yang salah sebagai kelas positif).
- Recall model sebesar 59.69% menunjukkan bahwa model memiliki kinerja yang lebih rendah dalam mengidentifikasi data yang sebenarnya sebagai kelas positif.
- F1-Score model sebesar 73.55% mengindikasikan keseimbangan antara presisi dan recall.

2. Model Bi-directional LSTM:

- Model Bi-directional LSTM memiliki tingkat akurasi sebesar 97.24%, yang menunjukkan kemampuan model dalam mengklasifikasikan data dengan benar.
- Presisi model sebesar 78.26% menunjukkan bahwa model cenderung memberikan sejumlah false positive (klasifikasi yang salah sebagai kelas positif) yang lebih tinggi dibandingkan dengan model LSTM.
- Recall model sebesar 65.97% menunjukkan peningkatan kinerja model dalam mengidentifikasi data yang sebenarnya sebagai kelas positif dibandingkan dengan model LSTM.

- F1-Score model sebesar 71.59% mengindikasikan keseimbangan antara presisi dan recall.

Dalam kedua model, kinerja umumnya baik dengan tingkat akurasi yang tinggi. Namun, terdapat perbedaan dalam presisi dan recall antara kedua model tersebut. Model LSTM memiliki presisi yang lebih tinggi, yang berarti lebih sedikit klasifikasi yang salah sebagai kelas positif. Di sisi lain, model Bi-directional LSTM memiliki recall yang lebih tinggi, yang menunjukkan kemampuan yang lebih baik dalam mengenali data sebenarnya sebagai kelas positif.

Dalam konteks aplikasi yang spesifik, pilihan model yang lebih sesuai tergantung pada preferensi dan kebutuhan Anda. Jika penting untuk menghindari false positive, model LSTM mungkin lebih baik. Namun, jika mengenali lebih banyak data yang sebenarnya sebagai kelas positif lebih penting, model Bi-directional LSTM mungkin lebih sesuai.

Kedua model juga dapat dianalisis lebih lanjut melalui laporan (classification report) yang memberikan insight tentang kinerja model dalam mengklasifikasikan setiap kelas secara lebih rinci. Hal ini dapat membantu dalam pemahaman lebih mendalam tentang kekuatan dan kelemahan masing-masing model dalam konteks kasus yang diberikan.

5.2 Saran

Berikut beberapa saran berdasarkan hasil evaluasi metode LSTM dan BiLSTM dimana beberapa saran yang dapat dipertimbangkan berdasarkan hasil evaluasi model:

1. Eksplorasi teknik pemrosesan teks tambahan: Terlepas dari model yang digunakan, ada beberapa teknik pemrosesan teks tambahan yang dapat meningkatkan performa model. Misalnya, pembersihan teks yang lebih lanjut, normalisasi kata, penghapusan stopword, atau penggunaan stemming atau lemmatization. Eksplorasi teknik-teknik ini dapat

membantu mengurangi noise dan meningkatkan kualitas data teks yang diolah.

2. Penyesuaian parameter model: menyesuaikan parameter model seperti ukuran vektor, jumlah layer, jumlah unit, dan fungsi aktivasi. Melakukan eksperimen dengan berbagai konfigurasi model dapat membantu mengoptimalkan performa model dan meningkatkan akurasi serta keseimbangan antara presisi dan recall.
3. Penggunaan ensemble learning: dapat mencoba menggunakan ensemble learning dengan menggabungkan beberapa model dan membuat keputusan berdasarkan mayoritas atau bobot. Ensemble learning dapat meningkatkan keakuratan dan stabilitas prediksi dengan memanfaatkan kekuatan berbagai model.
4. Perluas dataset: Jika memungkinkan, perluas dataset dengan mengumpulkan lebih banyak data atau menggunakan teknik augmentasi data seperti oversampling atau data synthesis. Dataset yang lebih besar dan beragam dapat membantu model mempelajari pola yang lebih baik dan menghasilkan prediksi yang lebih akurat.
5. Evaluasi metrik lainnya: Selain metrik-metrik standar seperti akurasi, presisi, dan recall, pertimbangkan untuk mengevaluasi metrik lain seperti area di bawah kurva ROC (AUC-ROC) atau mean average precision (MAP). Metrik-metrik ini dapat memberikan wawasan yang lebih lengkap tentang kinerja model dalam membedakan antara kelas positif dan negatif.
6. Validasi lintas waktu: Jika dataset mencakup data yang dikumpulkan dari waktu ke waktu, pertimbangkan untuk melakukan validasi lintas waktu. Metode ini membagi data menjadi bagian pelatihan dan pengujian berdasarkan waktu, sehingga menguji kinerja model pada data yang tidak dilihat sebelumnya. Hal ini dapat memberikan gambaran yang lebih realistis tentang bagaimana model akan berkinerja dalam penggunaan praktis.

7. Evaluasi lebih banyak arsitektur model: Selain LSTM dan Bi-directional LSTM, dapat menjajaki arsitektur model lainnya seperti CNN (Convolutional Neural Network) atau Transformer. Setiap arsitektur memiliki kekuatan dan kelemahan sendiri dalam pemrosesan teks, dan eksplorasi berbagai arsitektur dapat membantu menemukan pendekatan yang lebih efektif untuk tugas NLP tertentu.

