

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang sudah dilakukan, beberapa Kesimpulan dapat disimpulkan sebagai berikut:

1. Model LSTM dapat memprediksi kondisi mental pengguna berdasarkan teks dengan akurasi sebesar 92,07 persen, presisi 89,39 persen, recall 95,27 persen, dan F1-score 92,24 persen. Model ini memiliki kemampuan yang baik dalam mengingat informasi, sehingga cocok digunakan dalam skenario yang membutuhkan deteksi kasus positif. Pendekatan penelitian ini fokus pada kemampuan mendeteksi kasus positif.
2. Model BiLSTM menunjukkan hasil yang paling baik dibandingkan keempat jenis arsitektur lainnya dengan tingkat akurasi sebesar 92,98%, presisi 91,30%, recall 94,83%, dan F1-score 93,04%. Keunggulan ini terjadi karena kemampuannya memproses teks dalam dua arah, yaitu maju dan mundur, secara bersamaan pada setiap langkah waktu, sehingga mampu memahami konteks kalimat dengan lebih lengkap. BiLSTM juga memiliki jumlah false positive yang rendah, yaitu 250, yang menunjukkan keseimbangan yang baik antara presisi dan recall.
3. Model LSTM dengan mekanisme *Attention* mencapai akurasi sebesar 91,83%, presisi 89,37%, recall 94,76%, dan skor F1 sebesar 91,99%. Penambahan mekanisme perhatian membantu model memberi bobot lebih pada kata-kata yang penting, yang terlihat dari perbedaan kecil antara hasil pelatihan dan validasi sebagai tanda bahwa model memiliki kemampuan generalisasi yang baik. Namun performanya masih di bawah BiLSTM. Hal ini terjadi karena meletakkan perhatian setelah LSTM satu arah, sehingga mekanisme perhatian hanya berfungsi pada informasi konteks yang sudah ada sebelumnya, dan tidak bisa mengambil data dari arah yang berlawanan (masa depan). Dengan rata-rata panjang teks sebanyak 72 kata, manfaat dari perhatian belum maksimal karena teks masih tergolong singkat hingga sedang.
4. Model LSTM+MultiHead mencapai tingkat akurasi 91,10%, presisi 88,66%,

recall 94,04%, dan skor F1 sebesar 91,27%. Meskipun memiliki arsitektur yang sangat canggih dengan mekanisme perhatian multi kepala yang dapat menangkap berbagai hubungan antar kata secara bersamaan, penerapan mekanisme perhatian multi kepala setelah LSTM satu arah menyebabkan informasi dari konteks masa depan tetap tidak dapat diakses. Selain itu, desain bangunan yang rumit membuat jumlah parameter meningkat secara besar, sehingga model ini membutuhkan lebih banyak putaran atau data agar bisa mencapai hasil yang stabil. Keterbatasan komputasi (hanya 50 epoch) menyebabkan kinerjanya masih kurang baik dibandingkan model-model lainnya. Artinya, penataan struktur yang tidak tepat dan teks yang tidak terlalu panjang menyebabkan manfaat teoritis dari multi-head attention tidak bisa digunakan secara optimal.

5. Arsitektur BiLSTM dipilih sebagai model terbaik dalam tugas memprediksi kondisi mental dari teks, karena mencapai skor F1 sebesar 93,04% dan nilai AUC-ROC sebesar 98,15%. Model ini menawarkan keseimbangan yang baik antara tingkat keakuratan dan kemampuan untuk beradaptasi dengan berbagai jenis data, termasuk data yang baru diperoleh. Keunggulan BiLSTM dibandingkan LSTM+Attention dan LSTM+MultiHead menunjukkan bahwa untuk penelitian teks dengan panjang sedang seperti dalam penelitian ini, desain dua arah yang terintegrasi lebih efektif daripada menambahkan mekanisme attention setelah pemrosesan satu arah.

5.2. Saran

Berdasarkan pembatasan dan hasil yang ditemukan selama penelitian, berikut adalah saran untuk penelitian lebih lanjut:

1. Penambahan jumlah epoch. Model yang lebih rumit seperti LSTM+MultiHead mungkin bisa berkinerja lebih baik jika dilatih selama lebih banyak epoch, misalnya 100 epoch. Penelitian berikutnya bisa menggali hal ini dengan memperhatikan kecukupan sumber daya komputasi yang ada.
2. Hyperparameter tuning secara sistematis. Menggunakan metode grid search atau random search untuk menemukan kombinasi hyperparameter terbaik,

seperti jumlah unit LSTM, tingkat dropout, dan laju pembelajaran, dapat meningkatkan hasil kerja semua model.

3. Eksplorasi arsitektur hybrid. Menggabungkan BiLSTM dengan mekanisme perhatian multi-head (BiLSTM+MultiHead) bisa menjadi pilihan yang menjanjikan untuk memanfaatkan secara bersamaan kekuatan konteks dua arah dan kemampuan perhatian secara paralel. Selain itu, bisa juga dicoba memperhatikan bagian dalam atau sebelum LSTM (bukan setelahnya) untuk melihat bagaimana pengaturan posisi mempengaruhi hasilnya.
4. Pengujian pada teks yang lebih panjang. Karena rata-rata panjang teks sekitar 72 kata, disarankan untuk menguji model pada dataset yang berisi teks lebih panjang, seperti artikel atau dokumen klinis, agar manfaat dari attention dan multi-head attention dapat dieksplorasi dengan lebih maksimal.

