

BAB 5

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil implementasi dan uji coba sistem, simpulan dari penelitian ini dapat disusun sebagai berikut.

1. Implementasi XLM-RoBERTa untuk melakukan *multilingual text classification* pada teks berbahasa Inggris dan bahasa Indonesia telah selesai dibuat. XLM-RoBERTa *model* diambil dari dokumentasi resmi www.huggingface.com yang di *train* ulang menggunakan 2 *dataset* utama (bahasa Inggris dan bahasa Indonesia) dari www.kaggle.com.
2. Hasil dari uji coba menunjukkan bahwa penelitian ini mampu melakukan *multilingual text classification* dengan akurasi yang sangat baik menggunakan *model* yang melalui proses *training* dengan *dataset* Mixed News Dataset berukuran besar (108.190). *Model* ini berhasil mendapatkan nilai tertinggi saat dilakukan pengujian pada English News Dataset berukuran besar (37.886) dengan nilai MCC 0,892, nilai *accuracy*, *precision*, *recall* 0,911, dan nilai *F1* 0,910. Pada Indonesian News Dataset berukuran besar (70.304), *model* ini dapat mendapatkan nilai MCC 0,864, nilai *accuracy*, *precision*, *recall*, dan *F1* sebesar 0,902.
3. Hasil dari uji coba menunjukkan bahwa *model* XLM-RoBERTa yang digunakan mampu melakukan *zero-shot transfer learning* dari bahasa Inggris ke bahasa Indonesia dengan cukup baik. Hal ini dibuktikan pada *model* yang di-*training* sepenuhnya dengan English News Dataset

berukuran besar (37.886) dan dilakukan pengujian terhadap Indonesian News Dataset berukuran besar (70.304) dan mampu mendapatkan nilai MCC sebesar 0,422. Walaupun nilai akurasi tersebut masih tergolong rendah, hal ini membuktikan bahwa *model* XLM-RoBERTa memiliki potensi yang sangat baik untuk melakukan *multilingual text classification* pada bahasa lain yang tergolong *low-resource languages* dengan memanfaatkan *transfer learning*.

4. Hasil dari uji coba menunjukkan bahwa jumlah data yang digunakan saat proses *training* dan kombinasi bahasa dalam sebuah *dataset* sangatlah berpengaruh terhadap kemampuan *model* dalam melakukan *multilingual text classification*. Pada penelitian ini, ditemukan bahwa hasil dari *training model* dengan *dataset* campuran (bahasa Inggris dan bahasa Indonesia) memiliki kemampuan *multilingual text classification* (terutama untuk bahasa Indonesia) yang lebih baik dibandingkan *model* dengan *training dataset* satu bahasa saja. Selain itu, jumlah data yang dikombinasikan dalam sebuah *training dataset* juga mempengaruhi kemampuan *model* dalam melakukan *multilingual text classification*.

5.2 Saran

Berdasarkan penelitian yang dilakukan, dapat dijabarkan beberapa saran untuk pengembangan penelitian lanjutan sebagai berikut.

1. Mengembangkan *multilingual text classification* agar dapat digunakan ke lebih banyak *low-resource languages* lainnya, misalnya untuk bahasa-

bahasa daerah di Indonesia seperti bahasa Jawa, bahasa Sunda, bahasa Batak, dan lain-lain.

2. Mencoba pengaturan *hyperparameters* yang berbeda untuk mendapatkan *model* yang paling optimal dalam melakukan *multilingual text classification*
3. Memperbesar ukuran *dataset* yang digunakan baik dalam proses *training model* maupun proses uji coba untuk mendapatkan *model* yang paling optimal dalam melakukan *multilingual text classification*.