BAB V
PENUTUP

5.1. Kesimpulan

Penelitian ini berhasil mengembangkan sebuah sistem klasifikasi multi-label untuk mendeteksi informasi kebencanaan dari komentar Twitter dengan menggunakan algoritma *Categorical Boosting* (*CatBoost*) yang dioptimalkan melalui pendekatan optimasi *Bayesian*. Penelitian ini terdiri atas beberapa tahapan penting, mulai dari pengumpulan data, pra-pemrosesan teks, pelabelan otomatis, ekstraksi fitur dengan TF-IDF, hingga pembangunan dan evaluasi model klasifikasi.

a. Pengumpulan data dilakukan menggunakan *Tweet* Harvest, sebuah alat bantu yang memungkinkan ekstraksi data dari Twitter berdasarkan kata kunci tertentu. Kata kunci yang digunakan mencakup istilah yang relevan dengan bencana alam, seperti “banjir”, “bencana banjir”, dan “longsor”. Dari proses ini, diperoleh sebanyak 4.898 data komentar Twitter yang berkaitan dengan peristiwa kebencanaan.
b. Tahapan pra-pemrosesan data teks menjadi bagian krusial dalam meningkatkan kualitas fitur. Beberapa teknik yang diterapkan secara bertahap meliputi *cleansin*g, normalisasi, *stemming*, dan *stopword removal*. Tujuan dari proses ini adalah untuk menghilangkan noise, menyederhanakan bentuk kata, serta meningkatkan konsistensi dan relevansi data, sehingga menghasilkan representasi fitur yang lebih optimal untuk proses pelatihan model.
c. Data yang telah diproses kemudian digunakan dalam tahapan pelabelan otomatis, yang dilakukan dengan mengombinasikan pendekatan *keyword matching* dan *Named Entity Recognition* (NER). Proses ini menghasilkan lima jenis label, yaitu: verifikasi bencana, lokasi, kerusakan, korban, dan bantuan. Distribusi label menunjukkan bahwa label “verifiksai bencana” mendominasi dengan 4340 data, disusul oleh “lokasi” sebanyak 1825 data dan “kerusakan” sebanyak 1143 data. Sementara itu, label “korban” dan “bantuan” masing-masing berjumlah 1487 dan 1300 data.

d. Model *CatBoost* dibangun untuk melakukan klasifikasi multi-label berdasarkan fitur hasil ekstraksi dengan TF-IDF. Terdapat tujuh parameter utama yang digunakan dalam pemodelan, yaitu *iterations*, *learning_rate*, *depth*, *l2_leaf_reg*, *verbose*, *loss_function*, dan *random_state*. Dari parameter tersebut, empat di antaranya dioptimalkan menggunakan *Bayesian Optimization*, yakni *iterations*, *learning_rate*, *depth*, dan *l2_leaf_reg*. Pengujian dilakukan pada empat skenario rasio pembagian data, dan hasil terbaik diperoleh pada rasio 90:10, di mana model menunjukkan performa yang lebih optimal dibandingkan model sebelum optimasi.
e. Evaluasi performa model dilakukan dengan menggunakan tiga metrik evaluasi, yaitu *hamming loss*, *f1 weighted*, dan *accuracy*. Model terbaik, yakni model hasil optimasi dengan rasio data 90:10, menghasilkan *hamming loss* sebesar 0,0371, *f1 weighted* sebesar 95,21%, dan *accuracy* sebesar 82,45%.
f. Selain pengembangan model, penelitian ini juga berhasil merancang antarmuka pengguna (GUI) berbasis Streamlit. Antarmuka ini memfasilitasi pengguna dalam melakukan proses input data, pra-pemrosesan, pelatihan model, serta klasifikasi data baru secara interaktif dan mudah digunakan.

5.2. Saran Pengembangan

Mengacu pada hasil penelitian dan kesimpulan yang telah disampaikan, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan lebih lanjut.

a. Jumlah dataset masih terbatas, sehingga model mungkin belum mampu mengenali pola secara umum. Untuk meningkatkan kemampuan generalisasi model, sebaiknya digunakan dataset yang lebih besar dan beragam, agar representasi jenis bencana dan variasi komentar yang dianalisis menjadi lebih lengkap dan akurat.
b. Metode ekstraksi fitur masih menggunakan TF-IDF, yang hanya memperhatikan frekuensi kata tanpa memahami konteks makna kata dalam kalimat. Oleh karena itu, disarankan untuk menggunakan metode representasi teks yang lebih canggih seperti *fastText*, *Word2Vec*, atau *BERT* embeddings.

Metode ini dapat menangkap makna kata berdasarkan konteksnya, sehingga hasil klasifikasi bisa lebih akurat.

c. Teknik pelabelan otomatis masih memiliki kelemahan, terutama pada label ‘lokasi’ yang diperoleh dari model *Named Entity Recognition* (NER). Model NER yang digunakan masih menghasilkan beberapa kesalahan (*false positive*). Peningkatan akurasi bisa dilakukan dengan mempertimbangkan penggunaan model NER lain yang lebih kuat, atau melatih ulang model tersebut dengan data khusus tentang bencana. Selain itu, pelabelan otomatis berbasis kata kunci pada label seperti ‘bantuan’, ‘korban’, atau ‘kerusakan’ bisa ditingkatkan dengan pendekatan klasifikasi terawasi (*supervised*) atau semi-terawasi (*semi-supervised*) yang menggunakan fitur berbasis *embeddings* agar hasil pelabelan lebih cerdas dan tidak hanya bergantung pada keberadaan kata kunci.

Halaman ini sengaja dikosongkan