

## BAB V

### SIMPULAN DAN SARAN

#### 5.1 Simpulan

Setelah dilakukan serangkaian eksperimen melalui 9 skenario pengujian, model IndoBERT mengalami overfitting ketika proses pelabelan dilakukan secara manual menggunakan metode else-if. Sebagai solusi atas overfitting tersebut, metode pelabelan digantikan dengan pendekatan Fuzzy Matching menggunakan threshold sebesar 80%. Pergantian metode ke Fuzzy Matching berhasil mengatasi overfitting dan learning curve menjadi lebih seimbang, dilihat dari akurasi dan validasi yang lebih representatif terhadap data sebenarnya.

Model gabungan antara IndoBERT dan Fuzzy Matching menunjukkan performa baik dalam mendeteksi komentar judi online. Namun, performa model menurun drastis jika disajikan dataset baru, dengan metrik evaluasi seperti precision, recall, dan F1-Score berada di bawah batas minimal yang dapat diterima.

Kelemahan utama model gabungan IndoBERT dan Fuzzy Matching terletak pada keterbatasan kata kunci saat training yang bersifat pre-defined dan hanya diambil dari satu sumber dataset, yaitu "youtube\_chat\_jogja\_clean.csv". Hal ini menyebabkan model kesulitan dalam mendeteksi variasi kata baru yang muncul pada data baru, karena tidak ada dalam training.

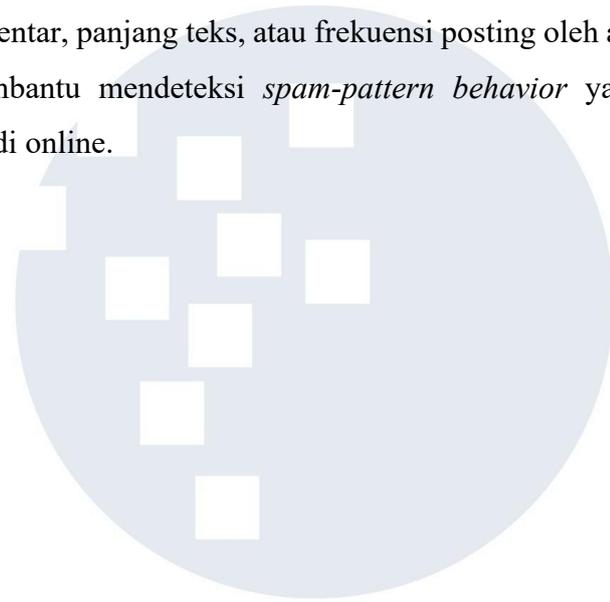
#### 5.2 Saran

Berdasarkan hasil dan keterbatasan yang ditemukan selama penelitian, beberapa saran yang dapat diberikan adalah sebagai berikut:

1. Integrasi dengan fitur anomaly detection untuk daftar kata kunci. Karena metode frequency analysis hanya mengandalkan kata-kata yang sering muncul, sistem berisiko gagal mendeteksi istilah baru atau variasi penulisan langka. Oleh karena itu, metode unsupervised anomaly detection dapat

menjadi pilihan untuk menangkap pola teks yang berbeda secara statistik dari mayoritas komentar.

2. Penggunaan teknologi deteksi real-time pada sistem seperti Kafka, Socket.IO, atau Webhook API untuk menerima komentar yang masuk secara langsung dari platform, misalnya YouTube API. Hal ini akan lebih memperkaya kata kunci untuk training model IndoBERT.
3. Selain teks, model IndoBERT dapat diperluas dengan fitur tambahan seperti waktu komentar, panjang teks, atau frekuensi posting oleh akun yang sama, untuk membantu mendeteksi *spam-pattern behavior* yang khas dalam promosi judi online.



UMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA