

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara yang memiliki tingkat kerawanan bencana alam yang tinggi, seperti banjir, gempa bumi, tanah longsor, dan letusan gunung berapi. Dalam situasi pascabencana, fase tanggap darurat menjadi tahap krusial yang menentukan keselamatan serta kesejahteraan korban terdampak. Pada fase ini, identifikasi kebutuhan dasar korban harus dilakukan secara cepat, akurat, dan terstruktur agar distribusi bantuan dapat berjalan efektif.

Badan Nasional Penanggulangan Bencana (BNPB) telah menetapkan standar kebutuhan dasar korban bencana yang harus dipenuhi selama masa tanggap darurat, yaitu:

1. Pangan
2. Pelayanan Kesehatan
3. Kebutuhan Air Bersih dan Sanitasi
4. Sandang
5. Penampungan dan Tempat Hunian Sementara
6. Pelayanan Psiko-sosial

Dalam praktik di lapangan, proses identifikasi kebutuhan korban sering kali mengalami keterlambatan akibat keterbatasan akses wilayah terdampak, keterbatasan sumber daya manusia, serta kendala komunikasi. Namun, perkembangan teknologi informasi telah mengubah pola penyebaran informasi masyarakat. Media sosial seperti X (sebelumnya Twitter) dan TikTok kini menjadi medium utama masyarakat dalam menyampaikan kondisi darurat, permintaan bantuan, serta laporan situasional secara real-time.

Informasi yang disampaikan melalui media sosial tersebut memiliki potensi besar sebagai sumber data untuk mendukung pengambilan keputusan dalam manajemen bencana. Akan tetapi, karakteristik data media sosial yang tidak terstruktur, masif, serta mengandung variasi bahasa non-formal dan slang menyebabkan proses identifikasi kebutuhan secara manual menjadi tidak efektif. Informasi penting berisiko tenggelam di tengah arus percakapan yang sangat cepat.

Sebagai contoh, pernyataan seperti “kami butuh air bersih dan makanan karena banjir merendam rumah” mengandung indikasi kebutuhan pada kategori Pangan dan Kebutuhan Air Bersih dan Sanitasi. Sementara itu, pernyataan seperti “anak-anak trauma dan sulit tidur setelah banjir” mengarah pada kategori Pelayanan Psiko-sosial. Tanpa sistem klasifikasi otomatis, proses pemetaan kebutuhan tersebut akan memakan waktu dan rentan terhadap kesalahan.

Dalam bidang *Natural Language Processing* (NLP), pendekatan berbasis Transformer seperti *Bidirectional Encoder Representations from Transformers* (BERT) telah terbukti unggul dalam memahami konteks semantik bahasa. Khusus untuk Bahasa Indonesia, model IndoBERT mampu menangkap makna kontekstual secara lebih akurat dibandingkan model generik. Namun, penggunaan BERT sebagai *classifier end-to-end* memerlukan sumber daya komputasi yang besar.

Oleh karena itu, penelitian ini mengusulkan pendekatan **Hibrida IndoBERT-SVM**, di mana IndoBERT digunakan sebagai pengekstraksi fitur kontekstual (*feature extractor*), sedangkan *Support Vector Machine* (SVM) digunakan sebagai algoritma klasifikasi multi-kelas. Kombinasi ini diharapkan mampu menggabungkan kekuatan representasi semantik dari IndoBERT dengan stabilitas dan kemampuan generalisasi SVM dalam membentuk hyperplane pemisah yang optimal.

Penelitian ini bersifat research-based dan berfokus pada pengembangan serta evaluasi model klasifikasi teks untuk mengidentifikasi kebutuhan dasar korban pascabencana berdasarkan standar BNPB. Dengan pendekatan ini, sistem yang dikembangkan diharapkan mampu menghasilkan informasi terstruktur yang dapat digunakan sebagai instrumen pendukung keputusan dalam penanggulangan bencana.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana merancang model klasifikasi teks berbasis Hibrida IndoBERT-SVM untuk mengidentifikasi kebutuhan dasar korban pascabencana sesuai standar BNPB?
2. Bagaimana mengimplementasikan model tersebut untuk mengklasifikasikan

data teks media sosial ke dalam enam kategori kebutuhan dasar BNPB?

3. Bagaimana performa model klasifikasi berdasarkan metrik evaluasi *Accuracy*, *Precision*, *Recall*, dan *F1-Score*?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Mengembangkan model klasifikasi teks berbasis Hibrida IndoBERT–SVM untuk kebutuhan dasar korban pascabencana.
2. Mengimplementasikan IndoBERT sebagai pengekstraksi fitur semantik dan SVM sebagai algoritma klasifikasi multi-kelas.
3. Mengevaluasi kinerja model menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score*.

1.4 Manfaat Penelitian

1.4.1 Implikasi Praktis

1. Membantu lembaga penanggulangan bencana dalam mengidentifikasi kebutuhan dasar korban secara otomatis dan cepat.
2. Mendukung optimalisasi distribusi bantuan agar sesuai dengan kategori kebutuhan yang teridentifikasi.
3. Menjadi dasar pengembangan sistem monitoring kebutuhan korban berbasis media sosial.

1.4.2 Implikasi Ilmiah

1. Memberikan kontribusi pada pengembangan NLP Bahasa Indonesia dalam domain kebencanaan.
2. Memberikan validasi empiris terhadap arsitektur Hibrida IndoBERT–SVM pada klasifikasi multi-kelas.
3. Menjadi referensi bagi penelitian selanjutnya terkait pemanfaatan media sosial dalam manajemen bencana.

1.5 Ruang Lingkup Penelitian

Agar penelitian lebih terarah, maka ditetapkan batasan sebagai berikut:

1. Data yang digunakan bersumber dari media sosial X dan komentar TikTok yang berkaitan dengan peristiwa pascabencana banjir.

2. Data berupa teks dalam Bahasa Indonesia dan Bahasa Inggris.
3. Klasifikasi dibatasi pada enam kategori kebutuhan dasar sesuai standar BNPB, yaitu:
 1. Pangan
 2. Pelayanan Kesehatan
 3. Kebutuhan Air Bersih dan Sanitasi
 4. Sandang
 5. Penampungan dan Tempat Hunian Sementara
 6. Pelayanan Psiko-sosial
4. Model yang digunakan adalah Hibrida IndoBERT–SVM.
5. Penelitian tidak mencakup analisis sentimen, prediksi waktu kejadian bencana, maupun implementasi sistem respons *real-time* penuh.

1.6 Kebaruan Penelitian

Kebaruan penelitian ini terletak pada beberapa aspek berikut:

1. Penerapan arsitektur Hibrida IndoBERT–SVM pada domain kebencanaan di Indonesia yang masih terbatas eksplorasinya.
2. Pengembangan klasifikasi multi-kelas fungsional berdasarkan standar resmi BNPB, sehingga output model bersifat *actionable* dan relevan secara operasional.
3. Validasi model bahasa pra-latih khusus Bahasa Indonesia (IndoBERT) dalam konteks teks kebencanaan yang memiliki karakteristik bahasa non-formal dan situasional.
4. Integrasi data media sosial sebagai sumber informasi alternatif dalam pemetaan kebutuhan korban pascabencana.