

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sejak Januari hingga Maret tahun 2025, Kepolisian Republik Indonesia (RI) melaporkan pencatatan 609 kasus Tindak Pidana Perdagangan Orang (TPPO) sebanyak 1503 korban. Dalam kurun waktu 3 bulan pada tahun 2025, jumlah korban bahkan sudah melebihi setengah total korban dengan kenaikan kasus hingga 50% dari total laporan kasus tahun 2024 [1]. Dari total 843 kasus perdagangan manusia pada tahun 2024, pemerintah melaporkan 482 kasus diantaranya kasus dengan “*unspecified forms of trafficking*”, melibatkan eksploitasi pekerja migran atau anak yang mencakup *labor* atau *sex trafficking* [2].

Migrant Care mencatat bahwasanya faktor kebutuhan mendesak akan pekerjaan dikarenakan fenomena pemutusan hubungan kerja (PHK) maupun ekonomi dimana sebagian besar korban adalah orang yang mengalami kemunduran pekerjaan pascapandemi dengan diiming-imingi iklan *fake job opportunity* [3]. Proses perekrutan pun dilakukan secara sistematis, memanfaatkan berbagai media sosial seperti Facebook, Instagram, dan Twitter untuk menyebarkan iklan lowongan pekerjaan. Bahkan, para perekrut menyamar sebagai staf HR perusahaan ternama untuk menumbuhkan kepercayaan calon korbannya dengan pemalsuan kontrak kerja serta surat penugasan agar terlihat resmi [2][4].

Salah satu platform media sosial yang digunakan para pelaku *human trafficking* adalah Facebook. Indonesia sendiri memiliki sebanyak 143 juta pengguna aktif Facebook per Januari 2025, yang mencakup sekitar 50,2 persen dari total populasi nasional [5]. Namun sayangnya, pertumbuhan eksponensial ini tidak diimbangi dengan sistem verifikasi dan moderasi konten yang memadai, sehingga membuka celah lebar bagi para pelaku kejahatan untuk menyebarkan tawaran pekerjaan fiktif. Beberapa kasus terkait *human trafficking* dari media sosial Facebook diantaranya seorang wanita dari Yogyakarta yang dijanjikan pekerjaan di restoran Macau melalui Facebook justru dipaksa menjadi scammer online di Kamboja [6]. Kasus serupa juga dialami Andi Arung, pemuda 18 tahun asal Selayar, terperangkap di Kamboja setelah tergiur iklan kerja di Facebook yang menjanjikan

gaji besar di perusahaan tambang Morowali [7]. Bahkan aparat penegak hukum mengungkapkan adanya jaringan perekrut yang secara sistematis menyebarkan lowongan palsu melalui Facebook sebagaimana diungkap dalam kasus kakak beradik Hariyanto dan Sugito yang ditangkap Polres Kawasan Bandara pada Juni 2023 [8]. Kelonggaran sistem Meta dalam mendeteksi dan menghapus akun palsu, dikombinasikan dengan algoritma yang cenderung mengamplifikasi konten sensasional, telah menjadikan Facebook sebagai "*breeding ground*" eksploitasi digital [9].

Afithasari et al. mengidentifikasi bahwa dalam konteks *disaster intelligence*, pendeteksian dini dengan integrasi teknologi digital untuk mendeteksi sinyal-sinyal perekrutan yang mencurigakan secara cepat, akurat, dan berkelanjutan sebelum meningkat menjadi krisis yang lebih serius [4]. Sebagian besar iklan lowongan kerja yang teridentifikasi berisiko trafficking di media sosial ditemukan dalam bentuk poster atau gambar, bukan sekadar teks caption [10], sehingga informasi kunci seperti identitas perekrut, kontak, dan syarat kerja seringkali hanya termuat di dalam gambar. Oleh karena itu, penelitian ini turut memanfaatkan hasil ekstraksi teks dari gambar melalui *Optical Character Recognition* (OCR) sebagai sumber data tambahan, agar cakupan informasi yang dianalisis tidak terbatas pada caption semata.

Deteksi risiko *human trafficking* pada iklan lowongan pekerjaan memerlukan pemahaman bahasa yang kontekstual dikarenakan penggunaan klaimat persuasif dan eufimistik yang maknanya baru terlihat jelas setelah mempertimbangkan konteks kata sekitarnya. Karakteristik inilah yang menjadikan BERT, khususnya IndoBERT sebagai pilihan relevan [11]. Kemampuannya menangkap konteks kalimat secara dua arah dan telah terbukti efektif pada tugas prediksi risiko *human trafficking* [12]. Namun, penerapan BERT pada kasus ini menghadapi tantangan data yang terbatas dan tidak seimbang, yang berisiko menyebabkan *overfitting* serta model yang bias terhadap kelas mayoritas sehingga *recall* pada kelas berisiko menjadi rendah meski akurasi terlihat tinggi [13]. Berbagai metode seperti *resampling* [14] dan *cost-sensitive learning* [15] telah diusulkan, namun umumnya belum menyertakan pengetahuan domain eksplisit

tentang pola indikator *human trafficking*, yang terbukti mampu memberikan sinyal deteksi lebih tajam bila digabungkan dengan representasi teks [16][17].

Berdasarkan uraian tersebut, penelitian ini mengembangkan model deteksi risiko human trafficking pada iklan lowongan kerja media sosial dengan menggabungkan representasi kontekstual IndoBERT dan meta-feature berbasis indikator risiko human trafficking dalam satu arsitektur *fusion* dari teks caption maupun hasil ekstraksi *Optical Character Recognition* (OCR).

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan sebelumnya, berikut adalah rumusan masalah pada penelitian ini:

1. Bagaimana mengimplementasikan integrasi ekstraksi teks OCR dan IndoBERT dengan ekstraksi *meta-feature* untuk membangun model deteksi lowongan kerja berisiko *human trafficking* di media sosial?
2. Bagaimana kinerja model Fusion OCR-IndoBERT dan *meta-feature* pada model deteksi lowongan kerja berisiko *human trafficking* di media sosial?

1.3 Batasan Masalah

Untuk memperjelas fokus penelitian, beberapa batasan masalah ditetapkan pada penelitian ini, diantaranya:

1. Penelitian difokuskan terhadap *human trafficking* pada sektor eksploitasi pekerja atau yang juga dikenal sebagai *labor trafficking* dengan perantara iklan lowongan pekerjaan ke luar negeri yang disebarkan di media sosial Facebook yang berdominasi Bahasa Indonesia.
2. Iklan lowongan pekerjaan yang digunakan adalah data postingan dalam periode 2023 - 2026 dengan catatan postingan memiliki gambar atau poster lowongan kerja disertai *caption* maupun tidak. Pengumpulan data melalui proses filtering manual postingan lowongan kerja, penyimpanan tautan, dan *web scraping*.
3. Setiap gambar poster diperlakukan sebagai satu unit data. Postingan dengan beberapa gambar poster dipecah sehingga setiap gambar menjadi data point tersendiri meskipun memiliki *caption* postingan yang sama.

4. Model yang dikembangkan terbatas pada analisis konten teks caption dan hasil ekstraksi OCR poster lowongan kerja. Analisis visual, warna, atau elemen grafis tidak termasuk pada ruang lingkup penelitian.
5. Pengecekan legalitas atau verifikasi nama perusahaan tidak termasuk pada ruang lingkup penelitian.
6. Hasil akhir model berupa besaran skor risiko per iklan lowongan kerja yang memiliki kemungkinan mengarah ke praktik *human trafficking* bukan konfirmasi hukum bahwa postingan tersebut merupakan praktik *human trafficking* yang terbukti.
7. Pelabelan data dilakukan secara manual merujuk pada panduan anotasi yang dikembangkan dari laporan publikasi *International Organization for Migration* (IOM) dan referensi literatur terkait.

1.4 Tujuan

Mengacu pada permasalahan yang telah dirumuskan, penelitian ini menetapkan tujuan sebagai berikut:

1. Mengimplementasikan model integrasi ekstraksi teks OCR dan IndoBERT dengan ekstraksi *meta-feature* pada deteksi lowongan kerja berisiko *human trafficking*.
2. Mengetahui kinerja model *Fusion* OCR-IndoBERT dan *Meta-feature* dalam mendeteksi lowongan kerja berisiko *human trafficking* di media sosial.

1.5 Manfaat

Penelitian ini diharapkan memberikan kontribusi pencapaian target *Sustainable Development Goals* (SDGs) 16 dalam upaya menghentikan eksploitasi dan perdagangan manusia dan manfaat bagi berbagai pihak, diantaranya:

1. Bagi Pencari Kerja atau Masyarakat

Sistem yang dikembangkan diharapkan dapat digunakan para pencari kerja melakukan pengecekan awal terhadap lowongan kerja yang mereka temui di media sosial. Dalam praktiknya, sistem deteksi risiko *human trafficking* dapat digunakan para pekerja sebagai salah satu pertimbangan mereka

sebelum memutuskan untuk melamar atau memberikan data pribadi, sehingga risiko menjadi korban trafficking berkedok lowongan kerja dapat diminimalisir.

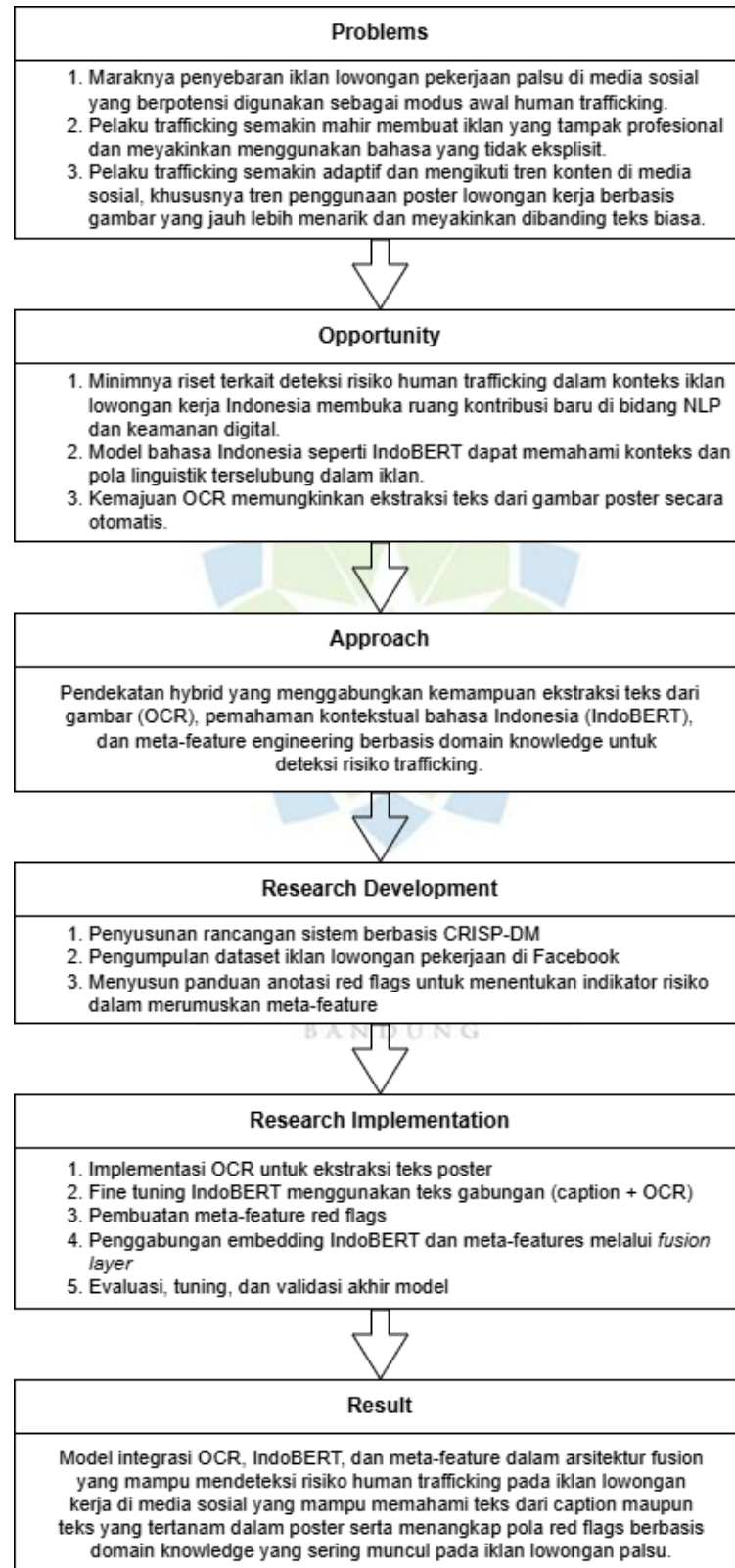
2. Bagi Peneliti

Menambah kontribusi pada pengembangan implementasi bidang *Natural Language Processing* untuk *social good* melalui pendeteksian risiko *human trafficking* pada data teks tidak terstruktur dari media sosial. Selain itu, penelitian ini dapat menjadi referensi bagi peneliti lain yang ingin menyempurnakan berbagai aspek dalam pengembangan sistem maupun mengembangkan sistem serupa.

3. Bagi Institusi dan Lembaga Terkait

Menjadi referensi pengembangan pemantauan digital iklan lowongan kerja ke luar negeri dalam memperkuat pengawasan terhadap praktik rekrutmen tenaga kerja migran yang berpotensi *human trafficking*. Institusi maupun lembaga terkait dapat memanfaatkan pendekatan ini sebagai bagian dari strategi proses identifikasi awal terhadap lowongan mencurigakan berbasis kecerdasan buatan dan tidak lagi sepenuhnya bergantung pada pelaporan manual.

1.6 Kerangka Pemikiran



Gambar 1. 1 Kerangka Pemikiran

Gambar 1.1 memperlihatkan alur konseptual pengembangan sistem deteksi risiko *human trafficking* pada iklan lowongan kerja. Bagian Problems menggambarkan tantangan utama seperti maraknya iklan palsu dan penggunaan poster visual yang sulit dideteksi. Dari pemahaman masalah tersebut, muncul peluang untuk memanfaatkan gabungan teknologi OCR, pemodelan bahasa IndoBERT, dan analisis indikator *red flags* untuk membaca sinyal risiko yang tidak tampak secara kasat mata. Alur pada gambar kemudian menunjukkan bagaimana ketiga komponen digabungkan ke dalam satu arsitektur *fusion*. Keseluruhan rangkaian ini dirancang untuk menghasilkan model yang mampu mengenali pola mencurigakan dalam iklan lowongan kerja secara lebih akurat, bahkan ketika pelaku memanfaatkan desain visual dan bahasa yang manipulatif.

