

1. PENDAHULUAN

Di masa lalu, komunikasi membutuhkan seseorang untuk bertemu secara personal atau mengirim pesan melalui pos. Namun, saat ini, semuanya menjadi lebih mudah karena kita hanya perlu menghubungi seseorang secara daring di internet. Saat ini berkomunikasi dengan media sosial begitu populer. Melimpahnya media sosial dan kemudahan mengaksesnya telah mengakibatkan pengaruh negatif. Misalnya, dalam beberapa tahun terakhir, ujaran kebencian, terutama kebencian perempuan, telah meningkat secara eksponensial di ruang media sosial Seperti Instagram, Twitter, Youtube, dan media sosial lainnya platform [1].

Berdasarkan data statistik penggunaan media sosial di Indonesia di tahun 2021, Indonesia memiliki banyak media sosial yang aktif pengguna, sebanyak 170 juta orang, dengan penggunaan rata-rata media sosial selama tiga jam empat belas menit. Youtube menempati urutan pertama sebagai media sosial yang paling banyak digunakan dengan rasio 92,8%, peringkat kedua adalah Whatsapp dengan rasio 87,6%, dan peringkat ketiga adalah Instagram dengan rasio 86,6%.¹

Misogini adalah suatu bentuk kebencian terhadap perempuan yang dapat diartikan sebagai segala sesuatu upaya ekspresi atau pandangan terhadap perempuan yang dapat berasal dari laki-laki atau sesama perempuan [2]. Misogini sendiri dapat berupa ekspresi verbal maupun nonverbal yang dapat dilakukan dalam bentuk seperti pengucilan sosial, diskriminasi terhadap perempuan, permusuhan, androsentrisme, fitnah terhadap perempuan, pelecehan seksual, dan kekerasan [2].

Ada beberapa jenis misogini, diantaranya *sexual harassment* adalah bentuk misogini yang merupakan tindakan pelecehan seksual, baik secara verbal maupun non verbal. *Discredit* adalah salah satu bentuk misogini yang menunjukkan rasa tidak hormat atau merendahkan perempuan, yang dapat berupa hinaan atau hinaan. *Stereotype* adalah misogini yang terlalu menggeneralisasikan pandangan tertentu tentang perempuan, seperti kepribadian, preferensi, dan kemampuan [2].

Kebencian atau ketidaksukaan terhadap perempuan dapat mendorong seseorang untuk melakukan kekerasan seksual karena menganggap perempuan sebagai makhluk yang lemah dan laki-laki sebagai makhluk yang superior² Pada tahun 2021 Komnas Perempuan Indonesia, mengeluarkan laporan catatan tahunan (CATAHU) 2021. CATAHU menjelaskan bahwa kekerasan dan pelecehan terhadap perempuan di jejaring sosial telah meningkat secara signifikan. Seperti yang terlihat pada CATAHU, terjadi peningkatan kasus dari 241 kasus pada tahun 2019 menjadi 940 kasus pada tahun 2020³. Ini menunjukkan bahwa masalah misogini perlu mendapat perhatian lebih.

¹ <https://andi.link/hootsuite-we-are-social-indonesian-digital-report-2021/>

² <https://dnktv.uinjkt.ac.id/index.php/misoginis-menjadi-salah-satu-faktor-kekerasan-seksual/>

³ <https://komnasperempuan.go.id/uploadedFiles/1466.1614933645.pdf>

Pendekatan *text classification* dapat digunakan untuk mengidentifikasi misogini, salah satu bidang *Natural Language Processing* atau NLP yang dapat membantu mengkategorikan atau mengklasifikasikan jenis kalimat ke dalam kategori tertentu [3]. Salah satu metode klasifikasi teks yang saat ini populer dan terbukti memiliki performa tinggi adalah *Bidirectional Encoder From Transformers* (BERT) [4].

Fine-tuning atau *transfer learning* adalah kemampuan suatu model untuk mentransfer pengetahuan dari model yang dilatih ke model baru untuk memecahkan masalah atau tugas baru [5]. Tidak mudah untuk melatih model dari awal menggunakan BERT karena BERT membutuhkan korpus yang besar dan spesifikasi *graphics processing unit* (GPU) yang tinggi [5]. Oleh karena itu, *fine-tuning* adalah solusi untuk masalah ini.

Kelemahan dari studi misogini yang ada adalah sebagian besar menggunakan dataset bahasa Inggris, dan hanya sedikit studi dengan bahasa Indonesia yang ditemukan. Bahasa Indonesia merupakan salah satu bahasa yang populer karena sering ditemukan di internet [6]. Namun, pengembangan resource atau sumber daya untuk pengembangan model natural language processing masih sangat lambat [6]. Oleh karena itu, penelitian ini akan fokus membangun model identifikasi misogini dengan bantuan *pre-trained* model IndoBert yang disediakan oleh *Indonesian natural language understanding* (IndoNLU), karena model BERT dari IndoNLU berfokus pada penggunaan bahasa Indonesia dan memiliki telah dilatih dengan dataset yang bersih dan besar (Indo4B). Dataset Indo4B berukuran sekitar 23 GB [6].

Berdasarkan uraian masalah dan penelitian sebelumnya, penelitian ini akan fokus untuk mengidentifikasi *sex harassment*, *discredit*, dan *stereotype* menggunakan algoritma BERT, sebuah algoritma deep learning dan transformer-based. Pendekatan BERT adalah metode penyempurnaan dengan model IndoBert yang telah dilatih sebelumnya.