

Abstrak– Saat ini, sosial media menjadi tempat untuk mengekspresikan pendapat. Pendapat tersebut bisa berupa positif ataupun negatif. Namun, belakangan ini pendapat yang sering muncul yaitu pendapat negatif, seperti ujaran kebencian (*Hate Speech*). *Hate speech* sering ditemukan di sosial media, seperti komentar jahat yang ditujukan untuk menghina individu atau kelompok. Berdasarkan data *WeAreSocial* pada tahun 2021, salah satu *platform* media sosial yang paling banyak digunakan di Indonesia adalah Twitter dengan nilai 63,6% pengguna. Menurut Kepolisian Negara Republik Indonesia, kasus *hate speech* lebih mendominasi selama kurun waktu April 2020 hingga Juli 2021. Selain itu, penanggulangan pelaku kejahatan dunia maya tidaklah mudah, sehingga dibutuhkan infrastruktur dan personel. Penanggulangan kejahatan dunia maya juga sulit, sehingga dibutuhkan infrastruktur dan personel. Oleh karena itu, diperlukan upaya untuk mengidentifikasi *hate speech* di platform Twitter dalam bahasa Indonesia agar penegak hukum dapat mendeteksi penyebaran *hate speech*. *Deep learning* merupakan salah satu metode untuk mendeteksi *hate speech*. Pada penelitian ini, kami menggunakan *deep learning* model *Long Short-Term Memory* (LSTM) dengan *word embedding*. *Word embedding* yang kami gunakan yaitu FastText dan *Global Vector*(GloVe) sebagai inputan representasi kata dan klasifikasi kata. FastText *embedding* menggunakan informasi subkata untuk membuat penyematan kata dan GloVe *embedding* menggunakan metode *unsupervised learning* yang dilatih pada korpus untuk menghasilkan vektor fitur distribusi. Dari hasil evaluasi pada percobaan model LSTM-FastText menggunakan *random oversampling* memiliki keunggulan dengan *F1-score* 89.91% dibandingkan dengan LSTM-GloVe memperoleh *F1-score* 82.14%.

Kata Kunci: Hatespeech; Twitter; Fasttext; Glove; Lstm