

## 1. Pendahuluan

### Latar Belakang

Keberadaan teknologi informasi yang semakin pesat telah memberikan nuansa baru terhadap semua aspek kehidupan. Dampak yang diberikan tidak selalu menguntungkan masyarakat, karena teknologi telah menyebabkan tindak kejahatan juga semakin mudah dilakukan[1]. Salah satunya adalah penyebaran ujaran kebencian di Indonesia yang sudah sangat umum dilakukan oleh pengguna media sosial dengan memanfaatkan teknologi informasi, biasanya menyebabkan informasi yang buruk tentang seseorang atau sekelompok suku tertentu melalui media sosial.

Contoh ujaran kebencian yang paling populer di media sosial Twitter yaitu cuitan dari musisi Ahmad Dhani pada 1 Maret 2017, secara tidak langsung merujuk pada Gubernur DKI nonaktif pada saat itu, Basuki Tjahaja Purnama. Kasus ujaran kebencian ini telah ditangani oleh pihak kepolisian setelah Basuki Tjahaja Purna melaporkan[2]. Kerugian yang terjadi pada reputasi dan integritas seseorang serta potensi permusuhan dan konflik dalam bentuk SARA membuat kerugian akibat kejahatan ini tidak signifikan. Penegak hukum menggunakan pendekatan manual yang mana akan menghabiskan banyak waktu.

Pada penelitian sebelumnya membahas mengenai klasifikasi sentiment terhadap performa model *Convolutional Neural Network*(CNN) dengan berbagai jenis *word embedding*[3]. Penulis ingin mengkaji dan melihat bagaimana penerapan *Convolutional Neural Network*(CNN) dan *FastText* sebagai penyematan kata untuk mendeteksi komentar sebanyak 13.169 tweet Indonesia yang tersedia pada kumpulan data publik. Kumpulan data sudah diberi label positif dan negatif. Percobaan dilakukan dengan menggunakan *cross validation k-fold* dan menyesuaikan nilai *dropout* untuk menilai dan mengukur seberapa baik kinerja model dalam mendeteksi ujaran kebencian.

### Topik dan Batasannya

Topik yang dibahas pada penelitian ini yaitu mengenai penggunaan *FastText* sebagai *word embeddings* dengan *Convolutional Neural Network* sebagai model klasifikasi terhadap sistem deteksi ujaran kebencian pada komentar Twitter Indonesia. Dataset yang sudah diperoleh memiliki peran penting sebagai proses analisa apakah teks didalam dataset tersebut mengandung ujaran kebencian atau tidak.

Adapun batasan masalah dalam penelitian ini sebagai berikut :

- a. Data yang bersumber dari *public dataset*
- b. Keterbatasan data komentar yang digunakan hanya Bahasa Indonesia

### Tujuan

Tujuan penelitian ini yaitu untuk menilai seberapa baik kinerja *FastText* sebagai *word embedding* dengan model klasifikasi *Convolutional Neural Network* dalam mendeteksi ujaran kebencian pada komentar Twitter Indonesia.

### Organisasi Tulisan

Urutan yang diterapkan pada penelitian ini yaitu sebagai berikut: pada bagian ke 2 menunjukkan studi literatur mengenai teori yang digunakan, kemudian pada bagian ke 3 menguraikan terkait sistem yang akan dibangun yaitu diawali oleh tahap *preprocessing* data awal hingga menjadi data siap dilatih untuk sistem deteksi ujaran kebencian menggunakan *Convolutional Neural Network* sebagai model klasifikasi dan *FastText* sebagai *word embeddings*. Bagian ke 4 akan menguraikan terkait hasil dari pengujian yang telah dilakukan, dan pada bagian ke 5 mengenai kesimpulan dan saran untuk penelitian berikutnya.