

## 1. Pendahuluan

### Latar Belakang

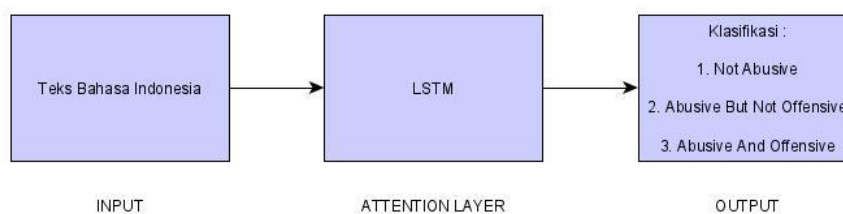
Adanya internet dan diikuti dengan munculnya berbagai jenis jejaring sosial merupakan salah satu hasil dari berkembangnya teknologi informasi[1]. Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) melalui buletin APJII November 2020 mengumumkan, jumlah pengguna internet di Indonesia mencapai 196,7 juta pengguna atau setara dengan 73,7% dari total populasi RI yang mencapai 266,9 juta penduduk[2]. Keberadaan internet seharusnya dipergunakan untuk mempermudah pengguna dalam mendapatkan informasi, menjalin hubungan dengan pengguna lain dan memperluas relasi antar pengguna[3]. Namun kemudahan ini tidak serta merta memberikan dampak baik kepada seluruh pengguna internet. Adanya anggapan “ini media sosial saya, terserah saya mau ngomong apa!” kerap kali memicu konflik diantara para pengguna internet. Penggunaan kalimat *abusive* seringkali ditemukan pada unggahan seperti ini dengan tujuan untuk menyerang pihak tertentu atau bahkan hanya sebagai bahan lelucon. Jika kalimat *abusive* dapat dengan mudah ditemukan di internet maka tidak menutup kemungkinan dapat mempengaruhi pola pikir pengguna khususnya yang masih berusia remaja bahwa penggunaan kalimat *abusive* dalam kehidupan sehari-hari tidak masalah[4].

Kalimat *abusive* merupakan ekspresi yang memuat kata-kata kasar atau kotor baik dalam lisan ataupun tulisan. Penyebab banyaknya penggunaan kalimat *abusive* pada internet atau jejaring sosial dikarenakan belum adanya *tools* yang efektif untuk menyaring penggunaan kalimat *abusive*, kurangnya rasa empati antar sesama pengguna internet dan kurangnya pengawasan orang tua[4]. Sebagai contoh, *Hate Speech* atau ujaran kebencian mengandung kalimat *abusive* yang seringkali dapat memicu konflik sosial karena dapat menimbulkan emosi bagi pihak yang dituju maupun siapa saja yang membacanya[5]. Penggunaan kalimat *abusive* juga dapat memberikan dampak buruk bagi kesehatan mental pihak yang dituju terutama jika pihak yang dituju masih tergolong usia remaja[6]. Penggunaan kalimat *abusive* dapat mengarah pada tindakan *cyberbullying* yang mana tingkat depresi yang akan dialami korban dapat lebih tinggi daripada depresi yang didapat dari kekerasan secara fisik[7].

Dalam mengatasi masalah tersebut, dibutuhkan tindakan preventif dan penegakkan hukum yang tegas dan sesuai dengan hukum yang berlaku[8]. Salah satu tindakan preventif bisa dilakukan dengan cara mendeteksi penggunaan kalimat *abusive* untuk penulisan pada media sosial. Deteksi yang dilakukan secara manual akan memakan banyak waktu[9], maka hal ini mendorong peneliti untuk menciptakan cara-cara otomatis yang dapat diterapkan dalam mendeteksi penggunaan kalimat *abusive* pada media sosial. Hal ini dapat dilakukan dengan cara membangun sebuah sistem yang dapat mendeteksi penggunaan kalimat *abusive*. Sistem ini akan dibuat menggunakan salah satu algoritma dari *Recurrent Neural Network (RNN)* yaitu *Long Short Term Memory (LSTM)*. Algoritma ini biasa digunakan pada masalah-masalah yang berkaitan dengan *Deep Learning*. Algoritma ini memiliki mekanisme internal yang disebut *gates* atau gerbang yang dapat mengatur aliran informasi. Gerbang ini dapat mempelajari data mana yang penting untuk disimpan atau yang perlu dilupakan dalam sebuah *sequence*[10]. Algoritma LSTM menggunakan *memory cell* yang dapat bekerja lebih baik dibanding dengan jaringan saraf rekuren biasa[11]. Algoritma LSTM juga biasa digunakan untuk prediksi dan klasifikasi[12]. Algoritma LSTM juga cocok digunakan pada data yang memiliki urutan seperti data *time series*, teks dan DNA karena setiap datanya terhubung satu sama lain [13]. Terdapat *imbalanced* atau ketidakseimbangan jumlah antar label pada dataset yang digunakan sehingga terdapat kelas mayoritas dan minoritas. Untuk itu peneliti juga ingin mengetahui pengaruh terhadap adanya penambahan jumlah dataset pada arsitektur yang dibangun.

### Topik dan Batasannya

Topik penelitian yaitu mendeteksi penggunaan kalimat *abusive* dengan masukan berupa teks bahasa indonesia lalu diproses dengan algoritma LSTM yang mana hasil keluaran berupa klasifikasi jenis kalimat. Terdapat 3 jenis label dalam klasifikasi hasil keluaran seperti gambar berikut:



**Gambar 1. Gambaran Topik Penelitian**

Batasan masalah dalam penelitian ini yaitu arsitektur hanya memproses kalimat bahasa indonesia karena *dataset* yang digunakan berupa teks bahasa indonesia sehingga struktur kalimat yang dipelajari hanya struktur kalimat bahasa indonesia. Jumlah *dataset* yang digunakan terdapat imbalanced atau ketidakseimbangan antara jumlah data pada setiap labelnya.

**Tujuan**

Tujuan dari penelitian ini yaitu mengevaluasi penerapan algoritma LSTM untuk melakukan deteksi penggunaan kalimat *abusive* pada teks bahasa indonesia dan untuk mengetahui pengaruh penambahan data pada dataset kelas minoritas terhadap performansi sistem dalam perhitungan nilai *F1 Score* untuk masing-masing kelas.

**Organisasi Tulisan**

Pada bagian 2 menjelaskan dasar teori yang digunakan sebagai pedoman dalam penelitian. Pada bagian 3 menjelaskan alur penelitian yang dilakukan. Pada bagian 4 menjelaskan hasil yang didapat melalui algoritma dan sistem yang sudah dibangun. Pada bagian 5 menjelaskan kesimpulan yang didapat dan saran yang bisa dikembangkan untuk penelitian selanjutnya.