

## BAB II

### TINJAUAN PUSTAKA

Pada tahap ini peneliti akan menjabarkan penelitian terdahulu sebagai penunjang pengerjaan tugas akhir yang diperoleh dari skripsi dan jurnal yang berkaitan dengan penelitian

#### 2.1 Penelitian Terdahulu

Di bawah ini penelitian sebelumnya yang dijadikan acuan dalam tugas akhir. Literature Review Dapat di lihat pada Tabel 1.

**Tabel 1.** Literature Review

No	Judul	Metode	Dataset	Hasil Akurasi
1.	Analisis Self-Attention pada Bidirectional LSTM dengan FastText dalam Mendeteksi Emosi Berdasarkan Teks	Bi-LSTM, FastText, Self-Attention	Emotion DW (40.000 tweet) dan RoBERTa dataset (416.809 tweet)	Bi-LSTM (8LTLi: 79%, 5LDLi: 99%, 2LDLi: 100%), Bi-LSTM + Self-Attention (79%, 96%, 99%), Self-Attention (77%, 91%, 95%)
2.	Sentiment Analysis of COVID-19 Vaccines in Indonesia on	Bi-LSTM, FastText, GloVe	6.547 tweet dari Twitter	GloVe (92.5%) pada dataset tanpa

	Twitter Using Pre-Trained and Self-Training Word Embeddings			stemming, FastText (92.3%) pada dataset dengan stemming
3.	Perbandingan CNN dan Bi-LSTM pada Analisis Sentimen dan Emosi Masyarakat Indonesia di Media Sosial Twitter Selama Pandemi Covid-19	Bi-LSTM, CNN, Word2Vec	58.984 tweet dari Twitter dan Kaggle	Bi-LSTM: 69.48% (sentimen), 84.36% (emosi); CNN: 68.58% (sentimen), 84.21% (emosi)
4.	Model Deteksi Berita Palsu Menggunakan Pendekatan Bidirectional LSTM dan GloVe	Bi-LSTM, GloVe	23.481 berita palsu dan 21.417 berita asli	97.5%

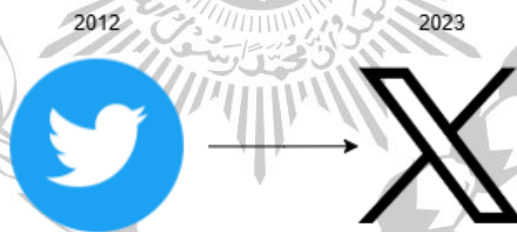
## 2.2 Landasan Teori

### 2.2.1 Aplikasi X

Twitter telah berkembang menjadi salah satu sumber informasi saat ini. Pengguna Twitter dapat mengumpulkan data dari tweet yang mereka publikasikan. Trending di Twitter dapat digunakan untuk mengumpulkan data tentang pendapat pengguna. Indonesia adalah negara kelima dengan jumlah

pengguna Twitter terbanyak. Pengguna media sosial Twitter dapat berekspresi dengan mengirimkan tweet kepada pengikutnya atau membalas satu sama lain dengan menyebutkan teman atau orang lain. Komentar di Twitter dapat digunakan sebagai kritik dan rekomendasi yang dapat diterima oleh calon kepala daerah [4]. Dalam situasi seperti ini, perlu dilakukan analisis sentimen, yang juga disebut opinion mining. Untuk mengetahui persepsi masyarakat terhadap penyelenggaraan pilkada [3].

Pada 22 Juli 2023, Twitter resmi berganti nama menjadi X sebagai bagian dari visi Elon Musk untuk menjadikannya "aplikasi segalanya" yang mencakup media sosial, audio, video, pesan, pembayaran, dan perbankan. Musk menilai nama "Twitter" tidak lagi relevan seiring berkembangnya fitur-fitur baru, serta mencerminkan sejarah pribadinya dengan huruf "X". Sejak perubahan ini, X menghadirkan berbagai fitur, seperti model berlangganan untuk verifikasi, unggahan video berdurasi panjang, perubahan nama tab navigasi, serta program bagi hasil iklan bagi kreator. Transformasi ini menegaskan komitmen X untuk menjadi platform interaktif tanpa batas, didukung kecerdasan buatan [26]. Timeline perubahan Twitter menjadi X dapat dilihat pada gambar 1



**Gambar 1.** Timeline Twitter Menjadi X

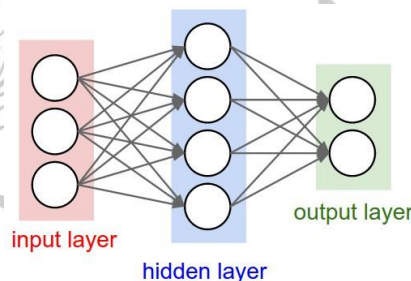
### 2.2.2 Analisis Sentimen

Analisis sentimen adalah bagian dari pengolahan teks yang dapat mengklasifikasikan teks untuk memahami arah perasaan dalam sebuah tweet. Ini digunakan untuk mengidentifikasi pandangan masyarakat terhadap suatu topik, yang mendapatkan tanggapan dari pengguna Twitter [6]. Analisis sentimen biasa digunakan untuk mengklasifikasi komentar pengguna X ke dalam beberapa kategori, di mana penelitian ini terdapat tiga kategori sentimen positif, netral, dan negatif [5].

### 2.2.3 Deep Learning

Seperti Artificial Neural Networks, model Pembelajaran Mendalam sangat efektif dalam mengolah volume data besar dengan mengekstrak pola-pola yang rumit dari data teks. Model Pembelajaran Mendalam juga menguasai berbagai tingkat representasi dari data, yang menghasilkan prediksi yang cukup tepat. [8].

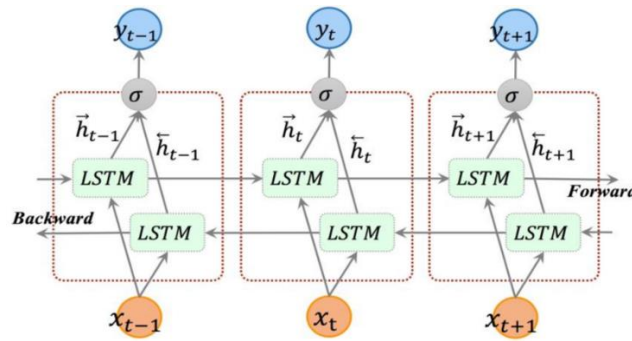
Dalam berbagai studi analisis sentimen, algoritma Deep Learning seperti Recurrent Neural Network dan Long Short Term Memory yang sering digunakan dalam analisis sentimen, mengindikasikan bahwa Deep Learning lebih unggul dibandingkan Machine Learning [8]. Neural Network Layer Dapat di lihat pada Gambar 2.



**Gambar 2.** Neural Network Layer

### 2.2.4 Bidirectional LSTM

Pemodelan dalam analisis sentimen melibatkan pengembangan dan implementasi mesin pelatihan atau model pembelajaran mendalam untuk memprediksi sentimen atau polaritas emosional dari data tekstual. Tujuan dari pemodelan adalah untuk membuat model yang secara akurat mengklasifikasikan teks ke dalam kategori sentimen positif, negatif, atau netral berdasarkan konten dan konteksnya. Metode yang digunakan untuk pemodelan menggunakan Bi-LSTM. Bi-LSTM memasukkan lapisan tambahan yang memproses urutan input dalam arah sebaliknya dari kanan ke kiri. Aliran dua arah ini memungkinkan model untuk menangkap ketergantungan dan informasi kontekstual dari kata-kata sebelum dan sesudahnya dalam urutan, yang menghasilkan pemahaman yang lebih baik tentang teks input [22]. Arsitektur dari Bi-LSTM dapat dilihat pada gambar 3.



**Gambar 3.** Arsitektur Bidirectional LSTM

Bi-LSTM menggunakan dua LSTM yaitu memproses urutan maju  $\vec{h}_t$  dan memproses urutan mundur  $\overleftarrow{h}_t$ . Adapun Rumus Bi – LSTM dapat dilihat pada rumus 1 :

$$\vec{h}_t = LSTM_{forward}(x_t, \vec{h}_{t-1}) \quad (1)$$

$$\overleftarrow{h}_t = LSTM_{forward}(x_t, \overleftarrow{h}_{t+1})$$

Kemudian, kedua hidden state digabungkan menjadi rumus berikut yang dapat dilihat pada rumus 2 :

$$\vec{h}_t = [\vec{h}_t; \overleftarrow{h}_t] \quad (2)$$

Gabungan ini bisa dilakukan dengan concatenation  $([:, :])$  penjumlahan, atau metode lain tergantung implementasi.

### 2.2.5 Word Embedding

Word Embedding adalah salah satu cabang proses pengolahan bahasa natural (NLP) yang mengolah kata menjadi sebuah vektor kata, yang mewakili sebuah kata di dalam ruang vektor. Dengan vektor ini, mesin dapat memahami bahasa yang dimaksud pengguna. Word Embedding telah digunakan dalam banyak hal, salah satunya adalah Word Similarity Task. Word Similarity Task mengkalkulasi nilai Word Similarity ketika nilainya mendekati 1. Jika nilainya mendekati -1, maka dua kata tersebut tidak memiliki kemiripan. Word Embedding dapat mengambil informasi semantik dan sintaktik kata dari

korpus. Dalam Word Embedding, vektor yang sama untuk kata-kata yang memiliki arti yang sama dan vektor yang berdekatan untuk memiliki kata kata yang memiliki arti yang sama [7].

