

---

## 1. Pendahuluan

Kemajuan teknologi internet telah membuat pencarian informasi lebih mudah melalui media online dibandingkan dengan sumber fisik seperti koran, majalah, dan buku. Masyarakat dapat dengan mudah mencari informasi melalui media online seperti situs web dan Twitter karena kemajuan teknologi yang pesat saat ini. Media sosial seperti Twitter memungkinkan pengguna untuk dengan cepat dan bebas berbagi informasi yang mereka miliki. Indonesia adalah negara peringkat kelima dalam hal pengguna Twitter secara global, dengan total 24 juta pengguna [1]. Dalam skala global, Twitter mengakumulasi 500 juta tweet di seluruh dunia, dengan total sekitar 200 miliar tweet per tahun [2]. Meskipun Twitter memfasilitasi komunikasi dan berbagi informasi, namun ada juga aspek negatifnya, seperti penyebaran hoax di media sosial [3].

*Hoax* adalah sebuah berita yang dibuat dan disebarluaskan oleh siapa saja, di mana saja, dan kapan saja tanpa memperhatikan kebenaran atau kekeliruan informasi yang disampaikan [4]. Tujuan dari penyebaran *hoax* adalah untuk merusak reputasi seseorang, kelompok, rekan kerja, atau bahkan teman, dan dapat menghasilkan kerugian finansial [5]. Seiring kemajuan teknologi, masyarakat telah beralih ke media sosial untuk berkomunikasi, yang secara tidak sengaja mengarah pada penyebaran hoaks dan berita bohong di kalangan pengguna. Di era media sosial, di mana setiap orang bisa menjadi "penerbit informasi", gelembung filter dan algoritme yang mempromosikan konten sensasional dapat memperkuat informasi yang salah [5]. Akibatnya, terjadi banyak penyebaran *hoax* terkait berita-berita viral di media sosial yang tidak bertanggung jawab dan menyebabkan disinformasi serta meresahkan masyarakat.

Metode deep learning yang berbeda seperti *Long Short-Term Memory* (LSTM), *Gated Recurrent Unit* (GRU), *Bidirectional GRU* (BI-GRU), dan *1-Dimensional Convolutional Neural Network* (1D-CNN) digunakan dalam analisis berita palsu, seperti yang dijelaskan dalam penelitian Nayoga dkk. mengenai analisis berita palsu pada berita Indonesia [6]. Dalam penelitian lain yang dilakukan oleh Bahad dkk., ditemukan hasil yang cukup baik dalam mengidentifikasi berita palsu menggunakan metode yang sama pada beberapa dataset berbeda. Pendekatan ini melibatkan penggunaan GloVe untuk mengukur hubungan antara judul dan konten artikel berita, serta penyetelan hyperparameter menggunakan data validasi guna mencapai tingkat akurasi yang baik [7].

Jaringan saraf *Bi-Directional Long Short-Term Memory* (Bi-LSTM) terdiri dari unit LSTM yang bekerja di kedua arah untuk menggabungkan informasi konteks masa lalu dan masa depan, sehingga dapat mempelajari dependensi jangka panjang tanpa menyimpan informasi konteks yang sama [8]. Berbeda dengan jaringan LSTM, jaringan Bi-LSTM memiliki dua lapisan paralel yang merambat dalam dua arah, yaitu lintasan maju dan mundur, yang memungkinkannya untuk menangkap ketergantungan dalam dua konteks [8], [9]. *Recurrent Neural Network* (RNN) merupakan sebuah jaringan saraf tiruan *feed-forward* yang dapat memproses input urutan panjang yang bervariasi. Dengan menggunakan lapisan tersembunyi berulang, aktivasi setiap waktu bergantung pada waktu sebelumnya, sehingga RNN lebih sesuai untuk memproses informasi kontekstual jarak jauh [10].

Tujuan dari penelitian ini adalah untuk mengembangkan sebuah sistem pendeteksi berita bohong yang dapat mengidentifikasi apakah sebuah berita yang tersebar adalah *hoax* atau bukan. Penelitian ini menggunakan empat metode deep learning yaitu Bi-LSTM, RNN, *hybrid* RNN-Bi-LSTM, dan *hybrid* Bi-LSTM-RNN. Metode *hybrid* memiliki keunikan dengan menggabungkan fungsi RNN untuk memproses data berurutan dan menjaga kemampuan Bi-LSTM dalam menyimpan informasi jangka panjang [8]–[10]. Untuk membangun model dari keempat metode tersebut, digunakan TF-IDF N-gram sebagai *baseline* atau ekstraksi fitur dan GloVe sebagai ekspansi fitur dalam pembangunan korpus.