

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Konsep "berita palsu" mengacu pada penyebaran dengan sengaja informasi palsu di media sosial yang bertujuan untuk membingungkan dan menyesatkan pembaca untuk mencapai agenda ekonomi atau politik. Selain itu, pelaku industri yang semakin beragam dan semakin banyak di bidang penulisan dan penyebaran berita telah mengarah pada penciptaan artikel berita yang sulit diketahui apakah kredibel atau tidak. Berdasarkan Badan Intelijen Nasional (BIN), hampir 60% informasi di Indonesia adalah *hoax* (Ghinadya & Suyanto, 2020).

Selain itu, berita bohong berpotensi menimbulkan kerugian sosial dan politik bagi masyarakat Indonesia. penganiayaan terhadap penderita kesehatan mental (Rifa'i, 2018) dan merusak reputasi politikus Indonesia saat pemilihan umum (BBC News Indonesia, 2019) dan bukti bahaya yang ditimbulkan oleh berita palsu adalah merusak reputasi politisi Indonesia selama pemilu. Pada tahun 2016, sebuah studi oleh Central Connecticut University bertajuk "World's Most Literate Nations" menempatkan Indonesia di peringkat 60 dari 61 negara peserta (Gunawan, 2016). Menurut pemeringkatan ini, literasi media Indonesia masih kurang memiliki kemampuan menilai informasi secara kritis dan membedakan berita yang benar dan yang salah (Khidhir, 2019).

### SEKOLAH PASCASARJANA

Pemanfaatan pembelajaran mendalam dalam penemuan berita palsu telah menjadi subjek dari beberapa penelitian yang bermanfaat, termasuk yang dilakukan oleh (Nayoga dkk., 2020) di mana metode *Deep Neural Network* (DNN), *Long Short-Term Memory* (LSTM), *Bidirectional LSTM* (BI-LSTM), *Gated Recurrent Unit* (GRU), *Bidirectional GRU* (BI-GRU), *1-Dimensional Convolutional Neural Network* (1D-CNN), *Support Vector Machine* (SVM) algoritma klasifikasi, dan *Naive Bayes*.

Penelitian lain, seperti “Deteksi *Hoaks* dalam Bahasa Indonesia Menggunakan Model Memori Jangka Pendek Panjang” (Apriliyanto & Kusumaningrum, 2020), berhasil mendeteksi berita hoaks berbahasa Indonesia dengan menggunakan metode LSTM. Hasilnya menunjukkan bahwa model DNN bekerja lebih baik daripada pengklasifikasi dalam tugas klasifikasi teks yang diawasi.

Penelitian sebelumnya akan digunakan untuk membuat Augmentasi Data Berbasis Sinonim untuk Deteksi Berita Palsu Menggunakan Metode CNN untuk Mendeteksi Berita Palsu Menggunakan *Headline*. Metode *Bidirectional LSTM* adalah subjek penelitian sebelumnya, dan memiliki skor F1 0,2423 dan digunakan dengan berita dari bahasa asing (Ghinadya & Suyanto, 2020). Namun F1-score hanya 0,24 karena data yang digunakan dalam penelitian ini adalah artikel berita yang ditulis dalam bahasa asing. Studi ini akan mengimplementasikan rekomendasi dari studi sebelumnya untuk lebih meningkatkan prosedur augmentasi dengan memanfaatkan data dari Berita Indonesia dan metode CNN. Metode CNN sering digunakan selama proses augmentasi (Nayoga dkk., 2020). Pemrosesan Bahasa Alami (NLP) menggunakan teks sebagai input utama untuk deteksi berita palsu, yang membutuhkan korpus bahasa yang sama dengan informasinya. Dengan demikian, korpus teks non-Inggris termasuk korpus teks bahasa Indonesia memiliki jargon yang lebih sedikit dibandingkan dengan korpus teks bahasa Inggris sehingga beberapa pengidentifikasi scam non-Inggris dapat diakses. Hasilnya, penelitian ini menggunakan augmentasi data untuk meningkatkan akurasi CNN dalam mendeteksi berita palsu berbahasa Indonesia dengan metode CNN.

## SEKOLAH PASCASARJANA

### 1.2 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengembangkan seberapa akurat metode Augmentasi dan CNN dalam mengidentifikasi *hoax*.

### **1.3 Manfaat Penelitian**

Pengembangan sistem pendeteksi berita palsu yang menggabungkan pendekatan CNN untuk meningkatkan kesadaran masyarakat terhadap informasi berita palsu menjadi manfaat dari penelitian ini.