

## BAB II

### TINJAUAN PUSTAKA

Berdasarkan referensi dari penelitian-penelitian sebelumnya, topik pembahasan dalam penelitian ini bukan hal yang baru. Penelitian ini memiliki beberapa kesamaan dengan penelitian-penelitian tersebut. Cen *et al.* (2020) melakukan penelitian untuk membentuk dua model menggunakan dua jenis arsitektur *Artificial Neural Network* (ANN), yaitu *Convolutional Neural Network* (CNN) dan *Recurrent Neural Network* (RNN). Penelitian ini menggunakan *dataset* yang berasal dari *Internet Movie Database* (IMDb) dengan jumlah data ulasan sebanyak 50 ribu. *Dataset* tersebut terbagi menjadi dua kelas sentimen, yaitu sentimen negatif dan sentimen positif. Dalam proses *data splitting*, *dataset* dibagi menjadi data latih dan data tes dengan rasio pembagian data yang sama, yaitu 50%. Penelitian ini menghasilkan model berarsitektur RNN dengan skor akurasi sebesar 68,64% dan model berarsitektur CNN dengan 88,22% akurasi [10].

Penelitian oleh Basiri *et al.* (2021) membentuk sebuah model dengan arsitektur gabungan antara *Bidirectional Long Short-Term Memory* (Bi-LSTM) dengan *attention layer* untuk melakukan analisis sentimen. Gabungan arsitektur tersebut dinamakan *An Attention-based Bidirectional CNN-RNN Deep Model* (ABCDM). *Dataset* yang digunakan pada penelitian ini berjumlah delapan dengan dua kelas sentimen, positif dan negatif. Untuk mempersingkat penulisan, tinjauan Pustaka ini hanya menyertakan tiga *dataset* beserta hasilnya. Ketiga *dataset* tersebut, yaitu: ulasan aplikasi Android, ulasan produk di Kindle Store, dan ulasan produk elektronik. Jumlah data ulasan untuk setiap *dataset* masing-masing adalah 752.937, 982.619, dan 1.689.188. Nilai *specificity*, *recall*, dan akurasi model yang dilatih dengan menggunakan *dataset* ulasan aplikasi Android masing-masing sebesar 94,91%, 89,45%, dan 92,18%. Model yang dilatih dengan *dataset* ulasan produk di Kindle Store mendapatkan nilai *specificity*, *recall*, dan akurasi masing-masing sebesar 95,92%, 90,88%, dan 93,40%. Performa model pada *dataset* ulasan produk elektronik, yaitu: *specificity* sebesar 94,29%, *recall* sebesar 87,01%, dan akurasi sebesar 90,65% [11].

Penelitian oleh Iqbal *et al.* (2022) menggunakan arsitektur LSTM untuk melakukan tugas analisis sentimen terhadap dua *dataset* yang berbeda, meliputi *amazon fine-food review* dan *cell phones and accessories reviews*. *Dataset amazon fine-food reviews* berjumlah 25 ribu ulasan dan *dataset cell phones and accessories reviews* terdiri dari 100 ribu ulasan. Kedua *dataset* ini masing-masing dibagi menjadi dua kelas sentimen, yaitu sentimen positif dan negatif. Performa model yang dilatih dengan menggunakan *dataset* pertama, yaitu: skor *recall* 55% dan skor akurasi 87%. Model yang dilatih dengan menggunakan *dataset* kedua mendapatkan skor *recall* dan akurasi masing-masing sebesar 63% dan 88% [12].

Berdasarkan penelitan-penelitian tersebut, penelitian ini diputuskan untuk menggunakan arsitektur *Long Short-Term Memory* (LSTM) dalam melakukan analisis sentimen terhadap ulasan produk kecantikan. Berbeda dengan penelitian sebelumnya, penelitian ini menggunakan *dataset* berisi ulasan berbahasa Indonesia. *Dataset* yang digunakan merupakan gabungan dari ulasan produk kecantikan yang didapatkan dengan cara *Application Programming Interface* (API) *scraping* dengan *dataset* yang telah tersedia secara daring. Jumlah ulasan yang berhasil dikumpulkan sebanyak 1.989.583. *Dataset* yang didapatkan melalui API *scraping* dan diunduh masing-masing berjumlah 1.589.590 dan 399.993 ulasan. Ulasan terbagi berdasarkan dua kelas sentimen, yaitu: positif dan negatif. Ulasan bersentimen positif terdiri dari 56,6% dari total data atau sebanyak 1.126.726, sedangkan ulasan bersentimen negatif sebanyak 862.857 atau 43,4% dari total data.

Tabel 2.1 Referensi penelitian

Peneliti		Cen, P., dkk. [10]		Basiri, M., dkk. [11]			Iqbal, dkk.[12]	
Tahun Penelitian		2020		2021			2022	
Sentimen		Positif dan negatif		Positif dan negatif			Positif dan negatif	
Jenis Dataset		IMDb review		Ulasan aplikasi android	Ulasan produk di kindle store	Ulasan produk elektronik	<i>Amazon fine-food reviews</i>	<i>Cell phones and accessories</i>
Jumlah Data Ulasan		50.000		752.937	982.619	1.689.188	25.000	100.000
Persentase Jumlah Data Ulasan per Sentimen		Positif: 50% Negatif: 50%		-			Positif: 87% Negatif: 13%	Positif: 89% Negatif: 11%
Arsitektur Model		RNN	CNN	ABCDM			LSTM	LSTM
Evaluasi Performa Model	Specificity	-		94,91%	95,92%	94,29%	-	-
	Recall	-		89,45%	90,88%	87,01%	55%	63%
	Akurasi	68,64%	88,22%	92,18%	93,40%	90,65%	87%	88%