

BAB II

TINJAUAN PUSTAKA

Penelitian dalam tesis ini melakukan deteksi emosi. Deteksi emosi menjadi salah satu persoalan penting dalam bidang *affective computing*. Deteksi emosi dapat dilakukan melalui beberapa media seperti ekspresi wajah, nada suara dan teks (Calvo & D'Mello, 2010). Perkembangan penggunaan teks sebagai media utama dalam komunikasi berbasis komputer menyebabkan deteksi emosi dari teks menjadi penting (Binali & Potdar, 2012; Anusha & Sandhya, 2015). Oleh karena itu berbagai penelitian telah dilakukan untuk menyelesaikan persoalan deteksi emosi dari teks.

Pada bagian ini akan dipaparkan sejumlah penelitian yang terkait dengan deteksi emosi dari teks. Penelitian terkait difokuskan pada pendekatan *keyword-spotting* dan *learning based*. Penelitian yang menggunakan kedua pendekatan tersebut dijelaskan pada bagian pertama. Selanjutnya dipaparkan penelitian terkait deteksi emosi dari teks bahasa Indonesia. Rangkuman penelitian terkait diberikan pada bagian terakhir.

2.1. Deteksi Emosi dari Teks

Deteksi emosi dari teks dapat diterapkan pada berbagai bidang seperti *costumer service*, media pembelajaran online, *video game* dan antar muka (Binali & Potdar, 2012; Anusha & Sandhya, 2015). Deteksi emosi dari teks juga merupakan persoalan penelitian yang menarik, oleh karena itu telah dilakukan berbagai penelitian untuk menyelesaikan persoalan tersebut. Berbagai pendekatan telah

digunakan untuk menyelesaikan persoalan deteksi emosi dari teks. Berdasarkan cara deteksi maka pendekatan tersebut dapat dikelompokkan menjadi *keyword-spotting*, *learning-based* dan *ruled-based* (Neviarouskaya, et al., 2011; Krcadinac, et al., 2013; Anusha & Sandhya, 2015).

Deteksi emosi dari teks dapat dilakukan dengan menemukan kata – kata yang terkait dengan emosi dari sebuah teks. Pendekatan ini disebut *keyword-spotting* (Krcadinac, et al., 2013). Pendekatan *keyword-spotting* menggunakan kamus yang berisi kata atau frasa yang terkait dengan suatu jenis emosi. Kamus ini sering disebut dengan leksikon emosi (Mohammad & Kiritchenko, 2015). Pendekatan *keyword-spotting* digunakan untuk mengembangkan *library Synesketch* (Krcadinac, et al., 2013). Proses deteksi emosi pada *Synesketch* menggunakan dua leksikon yaitu leksikon kata emosi dan leksikon emoticon. Leksikon kata emosi dikembangkan dari *WordNet*. Kelemahan utama metode *keyword-spotting* adalah tidak mampu melakukan deteksi jika kalimat tidak memiliki kata kunci yang sesuai dengan leksikon emosi (Krcadinac, et al., 2013; Anusha & Sandhya, 2015).

Pendekatan *keyword-spotting* terbatas pada leksikon emosi yang digunakan. Oleh karena itu deteksi emosi juga dilakukan tanpa menggunakan leksikon emosi. Deteksi emosi dari teks dapat dipandang sebagai suatu permasalahan klasifikasi. Suatu *classifier* dilatih menggunakan sejumlah data latih yang kemudian digunakan untuk melakukan deteksi (Anusha & Sandhya, 2015). Pendekatan yang menggunakan data latih untuk melatih *classifier* ini disebut *statistical* atau *learning based* (Krcadinac, et al., 2013; Anusha & Sandhya, 2015).

Pendekatan *learning-based* telah digunakan untuk mendeteksi emosi dari teks pada domain cerita anak – anak (Alm, et al., 2005). Pada penelitian yang lain dilakukan deteksi emosi terhadap isi blog (Aman & Szpakowicz, 2007). Emosi ditentukan dengan pendekatan *statistical* menggunakan metode *Support Vector Machines* dan *Naive Bayes*. Ciri yang diekstrak terdiri dari dua yaitu *corpus-based* dan *lexical-based*. Metode *Naive Bayes* memberikan tingkat akurasi 72,08 %. Analisis dan deteksi jenis emosi dari blog juga dilakukan pada penelitian Neviarouskaya, et al. (2011). Berbeda dengan penelitian Aman & Szpakowicz (2007), pada penelitian Neviarouskaya, et al. (2011) menggunakan pendekatan *rule-based* dan mengusulkan sebuah model deteksi emosi dari teks yang disebut dengan *Affect Analysis Model (AAM)*. Model AAM mampu untuk melakukan deteksi dari teks formal maupun informal.

Penelitian – penelitian tersebut diatas, proses deteksi dilakukan tanpa memperhatikan struktur hirarki dari emosi. Deteksi secara hirarki telah dilakukan pada penelitian Ghazi, et al. (2010). Penelitian ini menggunakan tujuh kelas emosi yang dibagi menjadi tiga tingkat. Pendekatan yang digunakan yaitu *learning-based* dengan metode SVM dan dua jenis ciri. Hasil penelitian menunjukkan bahwa model hirarki dapat meningkatkan akurasi. Model hirarki juga digunakan untuk klasifikasi emosi dari teks *micro-blog sina weibo* (Xu, et al., 2015). Model hirarki yang digunakan terdiri dari empat tingkat dengan 19 jenis emosi pada level empat. Hasil penelitian Xu, et al. (2015) juga menunjukkan peningkatan akurasi dengan menggunakan model deteksi hirarki.

Berbagai penelitian yang telah dijelaskan tersebut umumnya menggunakan data teks berbahasa Inggris dan Mandarin. Sedangkan penelitian tesis ini menggunakan data teks berbahasa Indonesia. Deteksi emosi dari teks bahasa Indonesia sudah dilakukan pada penelitian terdahulu. Penelitian – penelitian tersebut diberikan pada bagian berikut.

2.2. Deteksi Emosi dari Teks Bahasa Indonesia

Penelitian tesis ini melakukan deteksi emosi dari teks bahasa Indonesia. Metode Naive Bayes digunakan untuk melakukan klasifikasi emosi teks bahasa Indonesia dalam penelitian pemodelan ekspresi wajah 3D (Sumpeno, et al., 2011). Data yang digunakan merupakan terjemahan bahasa Indonesia dari data ISEAR. Pada penelitian yang lain metode *k-Nearest Neighbor* (k-NN) digunakan untuk melakukan klasifikasi artikel berita (Arifin & Purnama, 2012). Artikel berita diklasifikasikan menjadi enam jenis emosi.

Deteksi emosi juga sudah dilakukan menggunakan data tweet. Data tweet berbahasa Indonesia digunakan untuk melakukan deteksi lima jenis emosi (Arifin, et al., 2014). Deteksi emosi dilakukan berdasarkan pendekatan *learning-based* menggunakan metode k-NN. Pada penelitian ini metode *Non-Negative Matrix Factorization* (NMF) digunakan untuk mereduksi ciri. Jenis emosi yang digunakan yaitu takut, marah, terkejut, sedih dan senang. Selain deteksi emosi, tweet juga digunakan dalam analisis sentimen untuk mengukur rating acara televisi (Tiara, et al., 2015). Gabungan metode *lexicon-based* dan SVM digunakan untuk melakukan analisis sentimen.

2.3. Rangkuman Penelitian Terkait

Pada bagian ini telah diberikan sejumlah penelitian terkait dengan deteksi emosi dari teks. Umumnya deteksi emosi dari teks dilakukan tanpa memperhatikan bentuk hirarki dari emosi. Penelitian tesis ini melakukan deteksi emosi secara hirarki seperti pada penelitian Ghazi, et al. (2010) dan Xu, et al. (2015). Bentuk hirarki yang digunakan berdasarkan bentuk hirarki leksikon kata emosi bahasa Indonesia (Shaver, et al., 2001).

Deteksi emosi pada penelitian ini menggunakan pendekatan *keyword-spotting* dan *learning based*. Berbeda dengan penelitian sebelumnya yang menggunakan *lexicon-based* untuk melakukan proses pelabelan data latih metode SVM (Tiara, et al., 2015), pada penelitian ini metode *keyword-spotting* digunakan dalam proses deteksi. Pendekatan *keyword-spotting* yang digunakan berdasarkan *library Synesketch* (Krcadinac, et al., 2013). Penelitian ini juga melakukan pengembangan leksikon emosi berdasarkan buku Tesaaurus Bahasa Indonesia dan data tweet seperti pada penelitian Mohammad & Kiritchenko (2015). Rangkuman penelitian terkait dapat dilihat pada Tabel 2.1.

Tabel 2.1. Rangkuman Penelitian Terkait

Peneliti	(Ghazi, et al., 2010)	(Krcadinac, et al., 2013)	(Arifin, et al., 2014)	(Tiara, et al., 2015)	(Bata, 2016)*
Masalah	Deteksi emosi	Deteksi emosi	Deteksi emosi	Analisis sentimen	Deteksi emosi
Jenis emosi	7	6	5	2	5
Model deteksi	Hirarki	Flat	Flat	Flat	Hirarki
Pendekatan	<i>learning-based</i>	<i>Keyword-spotting</i>	Statistical	<i>Keyword-spotting, learning-based</i>	<i>Keyword-spotting, learning-based</i>
Bahasa	Inggris	Inggris	Indonesia	Indonesia	Indonesia
Data	Blog	Teks	Tweet	Tweet	Tweet
Metode	SVM	-	k-NN, NMF	SVM	NB