

C7_14_SEMNASIF_2015

by Pranowo Pranowo

Submission date: 06-Jan-2018 05:06PM (UTC+0700)

Submission ID: 900576212

File name: C7_14_SEMNASIF_2015.pdf (254.3K)

Word count: 3271

Character count: 20222

LEKSIKON UNTUK DETEKSI EMOSI DARI TEKS BAHASA INDONESIA

Julius Bata¹⁾, Suyoto²⁾, Pranowo³⁾

^{1,2,3)}Program Studi Magister Teknik Informatika

Program Pascasarjana Universitas Atma Jaya Yogyakarta

Jl. Babarsari 43 Yogyakarta 55281, Telp. (0274)487711

e-mail : 145302168@students.uajy.ac.id¹⁾

Abstrak

Deteksi emosi dari teks merupakan bidang penelitian yang menarik perhatian beberapa tahun terakhir. Salah satu komponen utama dalam deteksi emosi adalah leksikon emosi. Makalah ini memaparkan proses pengembangan leksikon emosi untuk bahasa Indonesia. Pengembangan leksikon terdiri dari 2 proses utama yaitu pemilihan seed words dan perluasan leksikon. Pemilihan seed words dilakukan berdasarkan jenis emosi yaitu senang, cinta, marah, takut dan sedih. Jumlah seed words yang digunakan sebanyak 124 kata. Perluasan leksikon dilakukan menggunakan Tesaurus Bahasa Indonesia. Setiap kata dalam leksikon diberi bobot biner 1 atau 0. Leksikon emosi yang dihasilkan terdiri dari 1165 kata.

Kata Kunci : Leksikon Emosi, Tesaurus, Deteksi Emosi.

1. PENDAHULUAN

Kemampuan mengolah emosi penting dalam suatu sistem komputer. Mengenali emosi pengguna dapat meningkatkan kualitas interaksi manusia-komputer (Lopatovska & Arapakis, 2010; Calvo & Kim, 2013). Selain itu, kecerdasan sistem komputer dapat ditingkatkan dengan menambahkan faktor emosi didalamnya (Tao & Tan, 2005). Deteksi emosi merupakan proses utama dan berperan penting dalam pengolahan emosi oleh komputer. Sebelum komputer dapat memberikan tanggapan sesuai emosi dari pengguna maka proses deteksi harus dilakukan terlebih dahulu (Calvo & D'Mello, 2010). Emosi dapat dideteksi dari suara, ekspresi wajah, gestur dan teks (Calvo & D'Mello, 2010).

Teks merupakan salah satu media yang digunakan untuk berkomunikasi dan menyampaikan informasi. Tidak hanya memuat informasi, teks juga dapat mengekspresikan emosi (Hirat & Mittal, 2015; Mohammad & Kiritchenko, 2015). Teks juga menjadi media utama dalam komunikasi menggunakan komputer (*Computer-Mediated Communication*) seperti email, blog dan media sosial (Quan & Ren, 2013). Deteksi emosi teks dapat digunakan pada berbagai aplikasi seperti analisa sentimen, ekspresi verbal pada interaksi manusia-komputer, analisa konsumen, pembelajaran online dan permainan komputer (Strapparava & Mihalcea, 2008; Binali & Potdar, 2012). Beberapa hal tersebut menunjukkan pentingnya melakukan deteksi emosi dari teks.

Berbagai metode telah diusulkan untuk menyelesaikan persoalan deteksi emosi dari teks. Terdapat 3 pendekatan utama dalam deteksi emosi dari teks yaitu *keyword-spotting*, *statistical*, dan *ruled-based* (Neviarouskaya et al., 2011; Krcadinac et al., 2013; Anusha & Sandhya, 2015). Pendekatan *keyword-spotting* bekerja dengan cara menemukan kata kunci yang berhubungan dengan suatu jenis emosi. Sedangkan pendekatan *statistical* menggunakan sejumlah besar data (*corpus*) untuk melatih model deteksi emosi. Data (*corpus*) juga digunakan untuk membuat sejumlah aturan dalam pendekatan *ruled-based*.

Sebagian besar metode yang digunakan untuk menyelesaikan deteksi emosi dari teks menggunakan leksikon emosi. Leksikon emosi merupakan daftar kata (kamus) serta jenis emosi yang terkait dengan kata tersebut (Poria et al., 2013). Leksikon emosi menjadi komponen utama dalam pendekatan *keyword-spotting* (Anusha & Sandhya, 2015). Leksikon emosi juga digunakan sebagai salah satu ciri pada pendekatan *statistical*. Ciri yang diperoleh dari leksikon emosi terbukti mampu meningkatkan akurasi (Aman & Szipakowicz, 2007; Ghazi et al., 2010; Mohammad, 2012b). Terdapat beberapa leksikon emosi yang telah dikembangkan dan digunakan untuk deteksi emosi seperti *WordNet-Affect* (Strapparava & Valitutti, 2004), *Affect database* (Neviarouskaya et al., 2011) dan *EmoLex* (Mohammad & Turney, 2013). Meskipun leksikon berperan penting dalam deteksi emosi dari teks, mayoritas leksikon emosi hanya tersedia dalam bahasa Inggris.

Pada sisi yang lain, pengguna media sosial di Indonesia telah bertumbuh dengan pesat. Hal ini mendorong penelitian menyangkut analisa sentimen dan deteksi emosi dari teks bahasa Indonesia mulai banyak dilakukan. Bahasa Indonesia merupakan *under-resourced language* (Wicaksono et al., 2014) yaitu bahasa yang masih kurang sumber daya untuk pengolahan bahasa alami. Oleh karena itu diperlukan penelitian untuk mengembangkan sumber daya tersebut. Salah satu sumber daya yang penting dalam analisa sentimen dan deteksi emosi adalah leksikon emosi bahasa Indonesia.

Penelitian dalam makalah ini merupakan bagian dari penelitian pengembangan model deteksi emosi dari teks bahasa Indonesia. Leksikon emosi pada penelitian ini adalah hasil awal dalam pengembangan leksikon emosi bahasa Indonesia. Leksikon emosi dikembangkan menggunakan buku tesaurus bahasa Indonesia. Kontribusi utama penelitian ini adalah mengembangkan leksikon emosi untuk bahasa Indonesia. Proses pengembangan leksikon diberikan pada bagian 3.

2. TINJAUAN PUSTAKA

Analisa sentimen dan emosi dapat diterapkan pada berbagai jenis teks dan dokumen. Sejalan dengan meningkatnya komunikasi menggunakan internet seperti email dan media sosial maka analisa emosi difokuskan pada media tersebut. Deteksi emosi sudah dilakukan terhadap isi blog (Aman & Szpakowicz, 2007; Ghazi et al., 2010; Ghazi et al., 2014), email (Mohammad, 2012a; Gupta et al., 2013) dan micro-blog (Mohammad, 2012c; Li & Xu, 2014). Berbagai metode juga sudah digunakan untuk melakukan deteksi emosi seperti *keyword-spotting* (Hancock et al., 2007; Krcadinac et al., 2013), *rule-based* (Chaumartin, 2007; Neviarouskaya et al., 2011) dan *statistical* dengan model pembelajaran mesin (Aman & Szpakowicz, 2007; Ghazi et al., 2010; Li & Xu, 2014).

Salah satu komponen utama dalam deteksi emosi dari teks adalah leksikon emosi. Leksikon emosi seperti *WordNet-Affect* (Strapparava & Valitutti, 2004) digunakan dalam metode *keyword-spotting* (Krcadinac et al., 2013) maupun *statistical* (Aman & Szpakowicz, 2007; Ghazi et al., 2010). Selain *WordNet-Affect*, ANEW (*Affective Norm for English Words*) juga digunakan untuk deteksi emosi model *dimensional* (Calvo & Kim, 2013). Meskipun leksikon seperti *WordNet-Affect* dan ANEW dapat digunakan dalam deteksi emosi, leksikon tersebut hanya terbatas pada bahasa Inggris. Leksikon tersebut juga memiliki keterbatasan ketika digunakan pada domain yang dinamis seperti media sosial (Bandhakavi et al., 2014). Terdapat sejumlah penelitian yang berfokus dalam pengembangan leksikon untuk mengatasi kedua masalah tersebut.

Pengembangan leksikon untuk bahasa Cina telah dilakukan oleh Yang et al. (2007). Leksikon dikembangkan menggunakan data yang diambil dari Yahoo! Kimo Blog. Jenis emosi yang digunakan sesuai dengan jumlah emoticon pada Kimo Blog yaitu 40 buah. Hubungan setiap kata dengan suatu emoticon ditentukan berdasarkan nilai *collocation* yang dihitung menggunakan *pointwise mutual information* (PMI). Pada penelitian yang lain, pendekatan algoritma *graph-based* digunakan untuk mengembangkan leksikon emosi bahasa Cina (Xu et al., 2011). Jenis emosi yang digunakan yaitu senang, marah, sedih, takut dan terkejut. Leksikon emosi untuk bahasa Cina juga dikembangkan dalam penelitian Li & Ren (2011). Leksikon dikembangkan berdasarkan corpus Ren-CECps 1.0. Corpus Ren-CECps adalah corpus yang berisi data weblog yang telah dianotasi dengan 8 jenis emosi. Leksikon emosi pada penelitian – penelitian tersebut (Yang et al., 2007; Xu et al., 2011; Li & Ren, 2011) dikembangkan berdasarkan corpus emosi.

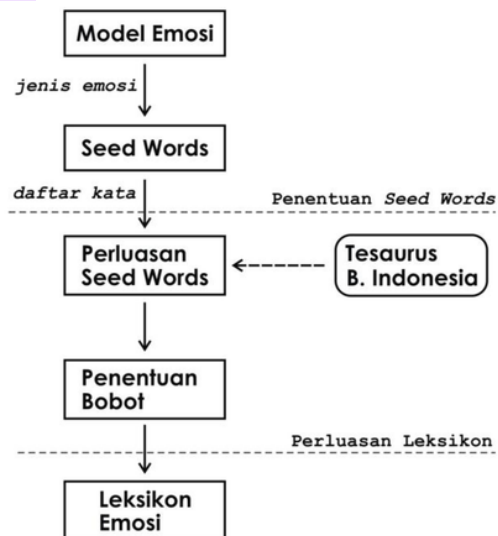
Data (corpus) emosi menjadi sumber utama ketika akan mengembangkan leksikon emosi untuk domain tertentu. Data twitter digunakan untuk mengembangkan leksikon emosi dalam penelitian Bandhakavi et al. (2014). Bobot setiap kata dihitung menggunakan jumlah kemunculan kata (*term frequency*) dan metode *iterative* yang didasarkan pada algoritma *expectation maximization* (EM). Pada penelitian lain, Muhammad (2012c) juga menggunakan data twitter (*TEC – Twitter Emotion Corpus*) dan 1000 headline berita untuk membuat leksikon emosi. Hubungan suatu kata dengan suatu jenis emosi dinyatakan dengan *Strength of Association* (*SoA*). Perhitungan *SoA* menggunakan PMI. Data dari situs berita digunakan untuk pengembangan leksikon emosi pada penelitian Staiano & Guerini (2014). Proses pembentukan leksikon terdiri dari 4 tahap yaitu membentuk matrik dokumen-emosi, lalu membentuk matrik kata-dokumen, selanjutnya membentuk matrik kata-emosi dengan melakukan perkalian antara matrik dokumen-emosi dan matrik kata-dokumen. Tahap akhir yaitu normalisasi matrik kata-emosi. Selain secara otomatis, leksikon emosi juga dikembangkan secara manual dengan melibatkan orang banyak atau yang disebut "*crowdsourcing*". Metode ini digunakan oleh Mohammad & Turney (2013)

untuk mengembangkan *EmoLex*. Proses anotasi emosi untuk setiap kata dilakukan dengan memanfaatkan layanan *Amazon's Mechanical Turk*.

Berbagai penelitian tersebut mengembangkan leksikon emosi berbahasa Inggris. Penelitian menyangkut pengembangan leksikon untuk bahasa Indonesia masih sedikit mendapat perhatian. Penelitian yang dekat dengan penelitian pada makalah ini adalah penelitian untuk mengembangkan leksikon sentimen bahasa Indonesia (Vania et al., 2014). Data-set yang digunakan adalah review yang diambil dari sejumlah situs tentang pariwisata Indonesia. Leksikon sentimen dikembangkan menggunakan *senti-patterns* dan menghitung polaritas setiap kata dengan adaptasi PMI. Penelitian Vania et al. (2014) menggunakan 2 polaritas yaitu positif dan negatif, pada makalah ini menggunakan 5 jenis emosi. Selain itu, leksikon sentimen pada Vania et al. (2014) dikhususkan pada domain pariwisata, sedangkan leksikon emosi pada makalah ini bersifat umum (*general-purpose emotion lexicon*).

3. METODE PENELITIAN

Makalah ini mengusulkan leksikon emosi untuk deteksi emosi dari teks bahasa Indonesia. Pengembangan leksikon emosi pada makalah ini terdiri dari 2 tahap utama yaitu 1) penentuan kata kunci (*seed-words*) dan 2) perluasan leksikon berdasarkan kata kunci. Aliran proses pengembangan leksikon dapat dilihat pada Gambar 1.



Gambar 1. Proses pengembangan leksikon emosi.

Berikut ini proses pengembangan leksikon sesuai Gambar 1. :

1. Penentuan *Seed Words*

Komponen utama dari leksikon emosi adalah daftar kata dan jenis emosi yang berelasi dengan kata tersebut. Oleh karena itu tahap awal dari pembentukan leksikon adalah menentukan jenis emosi yang akan digunakan. Pada penelitian ini model emosi yang digunakan sesuai dengan hasil penelitian Shaver et al. (2001). Pada model ini terdapat lima kategori emosi yaitu senang, sedih, cinta, takut dan marah.

Tahap berikut setelah menentukan jenis emosi yaitu menentukan *seed words*. *Seed words* merupakan sejumlah kata yang langsung terkait dengan satu jenis emosi tertentu. Sebagai contoh : jenis emosi senang memiliki beberapa *seed words* yaitu “*suka cita*”, “*riang*”, “*gembira*” dan “*girang*” (Shaver et al., 2001).

2. Perluasan Leksikon

Daftar kata dalam *seed words* hanya terdiri dari beberapa kata, sedangkan leksikon emosi membutuhkan kata – kata dalam jumlah besar. Kata dalam leksikon dikembangkan berdasarkan *seed*

42

words. Salah satu cara yang dapat digunakan adalah mencari sinonim dari kata dalam *seed words*. Cara ini berdasar pada pemikiran bahwa kata yang bersinonim cenderung memiliki makna yang sama, begitu juga jenis emosi dari kata tersebut.

Pendekatan ini membutuhkan sumber leksikal untuk mencari sinonim dari *seed words*. Pada penelitian ini menggunakan Tesaurus bahasa Indonesia Pusat Bahasa. Tesaurus Bahasa Indonesia Pusat Bahasa tersedia secara bebas dan dapat diunduh dari situs pusat bahasa. Tesaurus Bahasa Indonesia Pusat Bahasa memuat kata – kata yang saling berkaitan makna. Selain itu, juga mencantumkan jenis kata tersebut serta antonim (lawan kata).

Setelah dilakukan perluasan *seed words*, tahap selanjutnya adalah menentukan bobot dari setiap kata terhadap masing – masing jenis emosi. Penelitian ini menggunakan bobot biner yaitu jika suatu kata memiliki kaitan dengan suatu jenis emosi maka akan diberi bobot 1, jika tidak maka diberikan bobot 0. Bentuk akhir dari leksikon emosi ditunjukkan Gambar 2.

kata/frasa bobot₁ bobot₂ bobot₃ bobot₄ bobot₅

Gambar 2. Bentuk akhir leksikon emosi.

38

4. HASIL DAN PEMBAHASAN

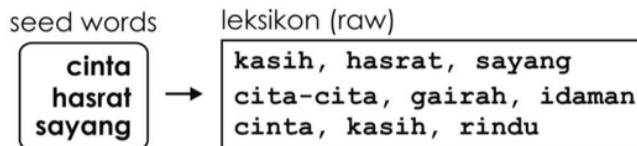
Tujuan utama dari penelitian ini adalah menghasilkan leksikon emosi. Pengembangan leksikon terdiri dari 2 tahap yaitu penentuan *seed words* dan perluasan leksikon. Kata yang digunakan sebagai *seed words* ditentukan berdasarkan Shaver et al. (2001). Jumlah total kata yang digunakan sebagai *seed words* adalah 124 kata. Kata – kata tersebut dikelompokkan kedalam 5 jenis emosi yaitu senang, cinta, marah, sedih dan takut. Persebaran *seed words* pada masing – masing jenis emosi dapat dilihat pada Tabel 1. Selain itu, contoh dari *seed words* untuk setiap jenis emosi juga diberikan pada Tabel 1.

Tabel 1. Sebaran dan Contoh *Seed Words*.

Jenis Emosi	# Kata	Contoh
Senang	26	"bangga", "kagum", "bahagia", "damai", "puas", "ceria", "riang", "sukacita"
Cinta	26	"ingin", "hasrat", "terpesona", "tertarik", "getar hati", "asmara", "mesra", "kangen"
Marah	31	"bosan", "jenuh", "muak", "benci", "emosi", "sebal", "murka", "naik darah", "naik pitam"
Takut	16	"gentar", "takut", "kecemasan", "gelisah", "risau", "bingung", "galau"
Sedih	25	"kecil hati", "simpati", "murung", "pilu", "sedih", "kasihan", "sakit hati", "putus asa"

Berdasarkan Tabel 1 terlihat bahwa *seed words* tidak hanya berupa kata melainkan juga gabungan kata atau frasa.

Selanjutnya adalah tahap perluasan leksikon. Perluasan leksikon dilakukan dengan mencari semua kata yang berhubungan dengan *seed words* di dalam Tesaurus Bahasa Indonesia. Sebagai contoh *seed words* "cinta" memiliki beberapa kata yang berhubungan seperti "kasih", "sayang" dan "hasrat", sedangkan *seed words* "sayang" berhubungan dengan kata "cinta", "kasih" dan "rindu". Semua kata yang berhubungan dengan *seed words* akan menjadi leksikon mentah (raw) seperti tampak pada Gambar 3.



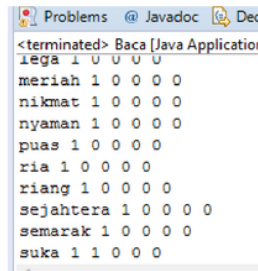
Gambar 3. Hasil leksikon mentah (raw).

Jumlah total kata dalam leksikon mentah adalah 2075 kata. Terdapat kata – kata yang sama dalam leksikon mentah, oleh karena itu dilakukan proses pembersihan. Setelah dilakukan proses pembersihan jumlah kata dalam leksikon adalah 1165 kata. Sebaran kata pada setiap jenis emosi seperti terlihat pada Tabel 2.

Tabel 2. Sebaran dan Contoh *Seed Words*.

Jenis Emosi	# Mentah	# Bersih
Senang	280	230
Cinta	389	251
Marah	628	301
Takut	421	221
Sedih	357	162

Tahap terakhir adalah penentuan bobot setiap kata. Hasil dari tahap ini adalah leksikon emosi yang berisi kata serta bobot. Bobot yang digunakan adalah biner yaitu 1 atau 0. Contoh leksikon emosi seperti terlihat pada Gambar 4.



```
Problems @ Javadoc Dec
<terminated> Baca [Java Application
iega 1 0 0 0 0
meriah 1 0 0 0 0
nikmat 1 0 0 0 0
nyaman 1 0 0 0 0
puas 1 0 0 0 0
ria 1 0 0 0 0
riang 1 0 0 0 0
sejahtera 1 0 0 0 0
semarak 1 0 0 0 0
suka 1 1 0 0 0
```

Gambar 4. Leksikon emosi.

Seperti yang terlihat pada Gambar 4, kata “riang” memiliki bobot [1 0 0 0 0]. Bobot [1 0 0 0 0] menunjukkan bahwa kata “riang” memiliki keterkaitan dengan jenis emosi senang dan tidak dengan jenis emosi yang lain. Terdapat beberapa kata yang termasuk dalam lebih dari 1 jenis emosi. Secara lebih jelas leksikon emosi dari Gambar 4 akan ditampilkan dalam bentuk tabel pada Tabel 3.

Tabel 3. Contoh Leksikon Emosi.

Kata	#Senang	#Cinta	#Marah	#Takut	#Sedih
puas	1	0	0	0	0
riang	1	0	0	0	0
suka	1	1	0	0	0

Pada Tabel 3. kolom #Senang menunjukkan bobot untuk jenis emosi senang, begitu juga dengan kolom #Cinta, #Marah, #Takut, #Sedih yang menunjukkan bobot untuk setiap jenis emosi. Terlihat pada Tabel 3. kata “suka” memiliki bobot #Senang dan #Cinta = 1, hal ini terjadi karena kata “puas” memiliki keterkaitan dengan jenis emosi senang dan cinta.

Analisa juga dilakukan terhadap kata – kata yang memiliki keterkaitan dengan lebih dari 1 jenis emosi. Terdapat beberapa kata seperti “suka”, “menyenangi”, “menaksir” dan “mencintai” yang termasuk dalam jenis emosi senang dan cinta. Sedangkan kata – kata seperti “kecewa”, “patah hati”, dan “sakit hati” termasuk dalam jenis emosi marah dan sedih. Tidak ada kata dari jenis emosi senang dan cinta yang termasuk dalam jenis emosi marah, takut dan sedih. Jenis emosi senang dan cinta termasuk jenis emosi positif sedangkan jenis emosi marah, sedih dan takut termasuk jenis emosi negatif.

5. KESIMPULAN

Makalah ini memuat hasil awal pengembangan leksikon emosi untuk bahasa Indonesia. Pada makalah ini sumber yang digunakan untuk mengembangkan leksikon adalah Tesaurus Bahasa Indonesia. Proses utama dalam pengembangan leksikon adalah pemilihan *seed words* dan perluasan leksikon. Kedua proses tersebut menghasilkan 1165 kata dalam leksikon emosi bahasa Indonesia.

Pada makalah ini hanya melakukan pengembangan leksikon. Beberapa hal yang dapat dilakukan pada penelitian selanjutnya adalah mengukur unjuk kerja dari leksikon dengan melakukan deteksi emosi dari teks. Penelitian ini hanya menggunakan bobot berupa biner, penelitian selanjutnya adalah menggunakan pembobotan lain seperti tfidf. Penelitian selanjutnya perlu menggunakan pendekatan lain seperti *corpus-based* untuk mengembangkan leksikon emosi.

DAFTAR PUSTAKA

- 33 Aman, S. & Szpakowicz, S., 2007. Identifying Expressions of Emotion in Text. In Text, Speech and Dialogue, Lecture Notes in Artificial Intelligence Vol 4629, pp. 196-205. 36
- 22 Anusha, V. & Sandhya, B., 2015. A Learning Based Emotion Classifier With Semantic Text Processing. Advances in Intelligent Systems and Computing Vol 320, pp. 371-382.
- Bandhakavi, A., Wiratunga, N., Deepak, P., & Massie, S., 2014. Generating a Word-Emotion Lexicon from #Emotional Tweets. Proc of the Third Joint Conference on Lexical and Computational Semantics (*SEM 2014), pp. 12-21.
- 45 Binali, H., & Potdar, V., 2012. Emotion Detection State of the Art. Proc of the CUBE International Information Technology Conference on CUBE 2012, ACM Press, pp. 501-507.
- 7 Calvo, R. A., & D'Mello, S., 2010. Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. IEEE Transactions on Affective Computing 1(1), pp. 18-37.
- 21 Calvo, R. A., & Kim, M. K., 2013. Emotions In Text: Dimensional And Categorical Models. Computational Intelligence 29(3), pp. 527-543.
- 6 Chaumartin, F.R., 2007. UPAR7: A knowledge-based system for headline sentiment tagging. Proc of the Fourth International Workshop on Semantic Evaluations (SemEval-2007), pp. 422-425.
- 18 Ghazi, D., Inkpen, D. & Szpakowicz, S., 2010. Hierarchical approach to emotion recognition and classification in texts. Advances in Artificial Intelligence LNCS Vol 6085, pp. 40-50.
- 26 Ghazi, D., Inkpen, D. & Szpakowicz, S., 2014. Prior and contextual emotion of words in sentential context. Computer Speech and Language 28, pp. 76-92.
- 10 Gupta, N., Gilbert, M., & Di Fabrizio, G., 2013. Emotion Detection in Mail Customer Care. Computational Intelligence 29(3), pp. 489-505.
- 24 Hancock, J.T., Landrigan, C., & Silver, C., 2007. Expressing emotion in text-based communication. Proc of the SIGCHI Conference on Human Factors in Computing Systems, ACM, pp. 929-932.
- Hirat, R., & Mittal, N., 2015. A Survey On Emotion Detection Techniques using Text in Blogposts. International Bulletin of Mathematical Research Vol 2, Issue 1, pp. 180-187.
- 9 Krcadinac, U., Pasquier, P., Jovanovic, J. & Devedzic, V., 2013. Synesketch: An Open Source Library for Sentence-Based Emotion Recognition. IEEE Transactions on Affective Computing 4(3), pp. 312-325.
- 35 Li, J., & Ren, F., 2011. Creating a Chinese Emotion Lexicon Based on Corpus Ren-CECps. Proc of IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), pp. 80-84. 31

- Li, W., & Xu, H., 2014. Text-based emotion classification using emotion cause extraction. *Expert System with Applications* 41, pp. 1742-1749.
- Lopatovska, I., & Arapakis, I., 2010. Theories, methods and current research on emotions in library and information science, information retrieval and human-computer interaction. *Information Processing and Management* 47(4), pp. 575-592.
- Mohammad, S.M., 2012a. From once upon a time to happily even after: Tracking emotions in mail and books. *Decision Support Systems* 53, pp. 730-741.
- Mohammad, S.M., 2012b. Portable features for classifying emotional text. *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 587-591.
- Mohammad, S.M., 2012c. #Emotional Tweets. *Proc of First Joint Conference on Lexical and Computational Semantics (*SEM)*, pp. 246-255.
- Mohammad, S.M., & Turney, P.D., 2013. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence* 29(3), pp. 436-465.
- Mohammad, S.M., & Kiritchenko, S., 2015. Using Hashtags to Capture Fine Emotion Categories from Tweets. *Computational Intelligence* 31(2), pp. 301-326.
- Neviarouskaya, A., Prendinger, H., & Ishizuka, M., 2011. Affect Analysis Model: novel rule-based approach to affect sensing from text. *Natural Language Engineering* 17, pp. 95-135.
- Poria, S., Gelbukh, A., Das, D., & Bandyopadhyay, S., 2013. Fuzzy Clustering for Semi-supervised Learning – Case Study: Construction of an Emotion Lexicon. I. Batyrshin and M. González Mendoza (Eds.): *MICAI 2012, Part I, LNAI 7629*, pp. 73–86.
- Quan, C., & Ren, F., 2013. Finding Emotional Focus for Emotion Recognition at Sentence Level. *Chinese Journal of Electronics* 22(1), pp. 99-103.
- Shaver, P.R., Murdaya, U., & Fraley, R.C., 2001. Structure of the Indonesian emotion lexicon. *Asian Journal of Social Psychology* Vol 4, pp. 201-224.
- Staiano, J., & Guerini, M., 2014. DepecheMood: a Lexicon for Emotion Analysis from Crowd-Annotated News. *Proc of the 52nd Annual Meeting of the Association for Computational Linguistics*, pp. 427-433.
- Strapparava, C., & Valitutti, A., 2004. WordNetAffect: an affective extension of WordNet. *Proc of the Conference on International Language Resources and Evaluation (LREC)*, pp. 1083-1086.
- Strapparava, C., & Mihalcea, R., 2008. Learning to identify emotions in text. *SAC'08: Proc of the 2008 ACM symposium on Applied computing*, pp. 1556-1560.
- Tao, J., & Tan, T., 2005. Affective Computing: A Review. J. Tao, T. Tan, and R.W. Picard (Eds.): *ACII 2005, LNCS 3784*, pp. 981-995.
- Vania, C., Ibrahim, M., & Adriani, M., 2014. Sentiment Lexicon Generation for an Under-Resourced Language. *International Journal of Computational Linguistic and Applications* 5(1), pp. 59-72.

- 30
Wicaksono, A. F., Vania, C., Distiawan, T. B. & Adriani, M., 2014. Automatically Building a Corpus for Sentiment Analysis on Indonesian Tweets. Proc of the 28th Pacific Asia Conference on Language, Information and Computation, pp. 185-194
- 3
Xu, G., Meng, X., & Wang, H., 2011. Build Chinese Emotion Lexicons Using A Graph-based Algorithm and Multiple Resources. Proc of the 23rd International Conference on Computational Linguistics (Coling 2010), pp. 1209-1217.
- 12
Yang, C., Lin, K.H-Y., & Chen, H-H., 2007. Building Emotion Lexicon from Weblog Corpora. Proc of the ACL 2007 Demo and Poster Sessions, pp. 133-136.

C7_14_SEMNASIF_2015

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

ORIGINALITY REPORT

19%

SIMILARITY INDEX

14%

INTERNET SOURCES

12%

PUBLICATIONS

10%

STUDENT PAPERS

PRIMARY SOURCES

- 1** Ahmad M. Abd Al-Aziz, Mervat Gheith, Ahmed Sharf Eldin. "Lexicon Based and Multi-Criteria Decision Making (MCDM) Approach for Detecting Emotions from Arabic Microblog Text", 2015 First International Conference on Arabic Computational Linguistics (ACLing), 2015
Publication 1%
- 2** Submitted to Università di Bologna
Student Paper 1%
- 3** "Joint Sentiment and Emotion Classification with Integer Linear Programming", Lecture Notes in Computer Science, 2015.
Publication 1%
- 4** mixedemotions-project.eu
Internet Source 1%
- 5** 68.180.206.247
Internet Source 1%
- 6** Submitted to Hong Kong Baptist University
Student Paper 1%

7	docs.di.fc.ul.pt Internet Source	1%
8	Submitted to Texas A&M University, College Station Student Paper	1%
9	sewaproject.eu Internet Source	1%
10	Submitted to University of Leeds Student Paper	1%
11	repository.upnyk.ac.id Internet Source	1%
12	www.mindigo.com Internet Source	1%
13	Ika Alfina, Dinda Sigmawaty, Fitriasari Nurhidayati, Achmad Nizar Hidayanto. "Utilizing Hashtags for Sentiment Analysis of Tweets in The Political Domain", Proceedings of the 9th International Conference on Machine Learning and Computing - ICMLC 2017, 2017 Publication	1%
14	www.aipo.es Internet Source	1%
15	Chang, Moon-Soo. "Empirical Sentiment Classification Using Psychological Emotions and Social Web Data", Journal of Korean	1%

16	ci.hfut.edu.cn Internet Source	1%
17	www.limbicsignal.com Internet Source	1%
18	Submitted to University of Newcastle upon Tyne Student Paper	1%
19	Submitted to University of South Africa Student Paper	<1%
20	Wang, Mengmeng Zuo, Wanli Wang, Ying. "A multidimensional nonnegative matrix factorization model for retweeting behavior prediction.(Resear", Mathematical Problems in Engineering, Annual 2015 Issue Publication	<1%
21	Sul, Hong Kee, Alan R. Dennis, and Lingyao Ivy Yuan. "Trading on Twitter: Using Social Media Sentiment to Predict Stock Returns : Trading on Twitter", Decision Sciences, 2016. Publication	<1%
22	www.tojet.net Internet Source	<1%
23	Submitted to University of Kent at Canterbury Student Paper	<1%

24	etd.lib.utexas.edu Internet Source	<1%
25	Bello-García, Marilyn, María Matilde García-Lorenzo, and Rafael Bello. "A Method for Building Prototypes in the Nearest Prototype Approach Based on Similarity Relations for Problems of Function Approximation", Lecture Notes in Computer Science, 2013. Publication	<1%
26	Submitted to Ryerson University Student Paper	<1%
27	www.experts.scival.com Internet Source	<1%
28	www.gsid.nagoya-u.ac.jp Internet Source	<1%
29	Submitted to University of Hertfordshire Student Paper	<1%
30	Submitted to Surabaya University Student Paper	<1%
31	www.ijsacs.org Internet Source	<1%
32	biometrics.cse.msu.edu Internet Source	<1%
33	openresearchsoftware.metajnl.com Internet Source	<1%

<1%

34

nactem.ac.uk

Internet Source

<1%

35

dblp1.uni-trier.de

Internet Source

<1%

36

Savargiv, Mohammad, and Azam Bastanfard. "Persian speech emotion recognition", 2015 7th Conference on Information and Knowledge Technology (IKT), 2015.

Publication

<1%

37

mahendraputra.net

Internet Source

<1%

38

hubsasia.ui.ac.id

Internet Source

<1%

39

journal.uny.ac.id

Internet Source

<1%

40

Ptaszynski, Michal, Jacek Maciejewski, Pawel Dybala, Rafal Rzepka, and Kenji Araki. "CAO: A Fully Automatic Emoticon Analysis System Based on Theory of Kinesics", IEEE Transactions on Affective Computing, 2010.

Publication

<1%

41

www.rug.nl

Internet Source

<1%

42 oputu.blogspot.com Internet Source <1%

43 www.thecrowdvoice.com Internet Source <1%

44 manuardirenata.blogspot.com Internet Source <1%

45 Chakraborty, Tamal, and Iti Saha Misra. "Design and analysis of channel reservation scheme in Cognitive Radio Networks", *Computers & Electrical Engineering*, 2015. Publication <1%

Exclude quotes Off

Exclude matches Off

Exclude bibliography Off