

Analisis Sentimen pada Twitter untuk Menilai Performansi Program Televisi dengan Kombinasi Metode *Lexicon-Based* dan *Support Vector Machine*

Tiara ⁽¹⁾, Mira Kania Sabariah S.T., M.T ⁽²⁾, Veronikha Effendy S.T., M.T ⁽³⁾

(1) Mahasiswa Teknik Informatika, (2) Staf Pengajar Telkom University, (3) Staf Pengajar Telkom University

Abstrak

Perkembangan media sosial, khususnya twitter berkembang sangat pesat. Twitter biasanya digunakan untuk mengomentari suatu produk, tokoh ataupun suatu program televisi. Komentar yang ditulis oleh para pengguna twitter dapat mencapai ratusan ribu bahkan jutaan tiap harinya. Dengan menggunakan komentar yang didapat dari twitter dapat melengkapi penilaian suatu program televisi yang selama ini dilakukan dengan menggunakan *rating*, dimana hal tersebut hanya mewakili dari segi kuantitas dan tidak dapat sepenuhnya dijadikan acuan dalam suatu penilaian. Maka dari itu, dengan menganalisis komentar dari media sosial twitter dirasa dapat melengkapi penilaian dari segi kualitas. Adapun cara yang dilakukan dalam menganalisis komentar Twitter adalah dengan melakukan proses analisis sentimen, dan metode yang digunakan pada penelitian ini adalah kombinasi dari metode *lexicon-based* dan *Support Vector Machine*. Berdasarkan hasil pengujian, kombinasi dari kedua metode tersebut terbukti dapat di implementasikan dalam menganalisa sentimen mengenai program televisi dengan nilai akurasi yang mencapai 80%. Nilai tersebut tidak dipengaruhi oleh perbandingan dari jumlah data latih dan data uji yang digunakan. Namun pada penelitian ini, data tweet yang di didominasi oleh sentimen positif cenderung memiliki nilai akurasi yang lebih tinggi daripada data tweet yang memiliki jumlah sentimen yang seimbang ataupun yang di didominasi oleh sentiment negatif.

Kata Kunci: *Twitter, program televisi, analisis sentimen, lexicon-based, Support Vector Machine*

1. Pendahuluan

Televisi merupakan media massa yang sangat berpengaruh dan paling akrab di kalangan masyarakat. Tayangan televisi menyediakan berbagai macam program acara televisi yang biasanya memiliki ciri khas tersendiri dalam penyajiannya, dimana hal itu juga akan memberikan pengaruh yang berbeda-beda kepada penontonnya. Menurut [10] perdebatan mengenai acara televisi kebanyakan lebih banyak didominasi oleh berapa banyak suatu program acara ditonton, program apa yang paling banyak ditonton (*rating*), dan seterusnya. Namun pemikiran seperti baik atau tidaknya acara tersebut untuk ditonton, apakah ramah bagi anak-anak, atau apakah acara tersebut menambahkan informasi penonton belum

banyak ditanyakan kepada penonton terkait. Selain itu menurut [10] para penonton tidak memiliki sarana untuk bersuara dan menyatakan pendapatnya atas kualitas berbagai acara televisi yang mereka tonton.

Beberapa tahun terakhir media sosial khususnya Twitter berkembang sangat pesat. Indonesia merupakan negara dengan pengguna Twitter yang terbilang aktif, terbukti dari data situs SemioCast Dot Com pada 1 Juli 2012 yang menyatakan jumlah pengguna twitter di Indonesia mencapai 29,5 juta [1]. Twitter biasa digunakan untuk mengutarakan pendapat mengenai hal personal, produk, ataupun suatu program acara di televisi. Biasanya hal itu dilakukan

untuk meningkatkan popularitas pengguna di dunia Twitter (dunia maya).

Ocehan di Twitter mungkin terlihat hanya sebatas kata-kata yang tidak memiliki nilai. Namun sebenarnya ocehan tersebut dapat memiliki arti, terlebih lagi sebagai penyampaian opini yang dapat diteliti dengan suatu cara yaitu analisis sentimen. Analisis sentiment adalah bidang studi yang menganalisa pendapat, sentiment, evaluasi, penilaian, dan emosi banyak orang mengenai entitas seperti produk, pelayanan, organisasi, individu, permasalahan, acara, topic, dan atributnya. [4]

Banyak metode yang digunakan dalam analisis sentimen ini seperti *Naïve Bayes*, *K-Nearest*, *Maximum Entropy*, dan lain-lain. Namun, pada penelitian ini digunakan kombinasi metode *Lexicon-Based* dengan metode *Support Vector Machine* (SVM) dikarenakan metode *Lexicon-based* dapat menghasilkan nilai *precision* yang tinggi, namun menghasilkan nilai *recall* yang rendah. Nilai *recall* yang rendah itu dapat ditingkatkan dengan menggunakan metode SVM. Selain itu, dengan menggunakan metode *lexicon-based* dapat mengurangi proses *manual labeling*, yang biasanya dilakukan untuk proses *training data*. [3]

Twitter yang biasa digunakan untuk penyampaian opini ataupun kritik mengenai suatu produk sekiranya dapat dijadikan sarana yang baik untuk menampung suara atau pendapat mengenai suatu acara televisi. Selain itu, sifat Twitter yang selalu *up-to-date* juga memudahkan dalam penerapan metode analisis sentimen dalam segi data. Penelitian ini diharapkan dapat memperkaya penilaian terhadap suatu program televisi melalui sisi kualitas yang diambil dari suara para penontonnya.

2. Dasar Teori

2.1 Twitter

Twitter merupakan situs *micro blogging* yang dioperasikan oleh Twitter, Inc. Twitter disebut *microblogging* karena situs ini memungkinkan penggunanya mengirim dan membaca blog pada umumnya. Pesan yang dikirimkan melalui Twitter disebut dengan *tweets*, yaitu teks tulisan yang berkapasitas maksimal 140 karakter yang ditampilkan pada halaman profil pengguna.

Dalam twitter terdapat konten-konten yang bisa kita gunakan dengan mudah. Konten-konten tersebut antara lain adalah *Home*, *Profile*, *Following*, *Follower*, *Mentions (@)*, *Direct Message*, *Hashtags*, *Trending Topics*, dan sebagainya. [3]

Pada penelitian ini data yang digunakan adalah tweet (URL dan hashtag diabaikan) yang dituliskan oleh user mengenai program televisi yang dianalisis.

2.2 Program Televisi

Program televisi adalah sebuah perencanaan dasar dari suatu konsep acara televisi yang akan menjadi landasan kreatifitas dan desain produksi yang akan terbagi dalam berbagai kriteria utama yang disesuaikan dengan tujuan dan target pemirsa acara tersebut. Suatu program televisi biasanya memiliki tujuan dalam apa yang disiarkan olehnya, antara untuk memberikan hiburan ataupun untuk memberikan informasi yang berguna untuk penontonnya. [6]

Di Indonesia, jumlah orang yang menonton program acara di stasiun televisi mengambil peran penting untuk bertahannya suatu program televisi. Maka nilai *rating* dinilai sangat penting dan sangat menentukan peluang suatu program acara itu untuk mendapatkan dari banyaknya iklan yang menjadi sponsornya. Namun permasalahan di Indonesia, acara dengan nilai *rating* yang tinggi belum tentu memiliki

kualitas yang tinggi (hasil penelitian Program Studi Televisi IKJ, 2006). Maka penelitian ini dilakukan untuk memperkaya penelitian kuantitas (*rating*) yang selama ini dilakukan dengan cara menganalisis pendapat orang mengenai program televisi yang dituliskan melalui *media social* twitter. Adapun lima program televisi yang digunakan sebagai data untuk dianalisis, yaitu : Mata Najwa, Kick Andy, Tukang Bubur Naik Haji (TBNH), Yuk Keep Smile (YKS), dan juga Pesbuker.

2.3 Analisis sentimen

Analisis sentiment atau dapat disebut juga sebagai opinion mining adalah bidang studi yang menganalisa pendapat, sentiment, evaluasi, penilaian, dan emosi banyak orang mengenai entitas seperti produk, pelayanan, organisasi, individu, permasalahan, acara, topic, dan atributnya [4]. Sebelum melakukan analisis sentimen dengan metode *lexicon-based*, data twitter yang masih mentah (*raw*) harus diubah menjadi data yang dapat diolah melalui proses *preprocessing*. Proses *preprocessing* yang dilakukan dalam tugas akhir ini adalah *data cleaning*, tokenisasi, *Part-of-speech Tagging* (*POS Tagging*).

Ekstraksi Opini dan Fitur

Setelah dilakukan proses *preprocessing*, sistem akan mengekstraksi kalimat opini. Kalimat opini merupakan kalimat yang mengandung kata opini (kata dengan label *adjective*, *adverb*, dan juga *verb*) [3]. Sedangkan untuk mengekstraksi fitur dilakukan dengan mengambil semua kata dengan label *noun* (kata benda).

2.4 Lexicon-based

Lexicon-based termasuk dalam kategori *unsupervised techniques* dimana metode ini melakukan

Metode *lexicon-based* ini sendiri menggunakan bantuan kamus untuk melakukan klasifikasi tweet menjadi tweet positif ataupun negatif. Ada beberapa tahapan dari *lexicon-based* yang dilakukan pada penelitian ini, yaitu penentuan polaritas kata, penanganan negasi, serta pemberian skor terhadap setiap fitur atau entitas. Adapun rumus yang digunakan dalam pemberian skor tersebut dapat dilihat pada persamaan (2-1)

$$w_i = \sum_{j=1}^n w_{ij} \cdot s_{ij} \tag{2-1}$$

penggolongan/klasifikasi data menjadi dua kelas yaitu positif atau negative [5][6].

2.5 Support Vector Machine

Sebelumnya kita telah membahas metode *lexicon-based* yang merupakan metode dari *unsupervised techniques*, dan pada sub-bab ini akan dijelaskan metode *Support Vector Machine* yang merupakan metode dari *supervised techniques*. Mungkin perlu diingatkan kembali bahwa *Supervised*

merupakan sebuah metode yang melatih sebuah sentimen klasifier yang diambil berdasarkan frekuensi kemunculan dari macam-macam kata yang terdapat pada dokumen, teks, ataupun tweet [7]. Dengan melakukan pelatihan yang menggunakan data inputan berupa data numerik seperti nomor indeks kata, dan juga bobot (biasanya didapatkan dari perhitungan TF-IDF, Term Presence, dsb) akan menghasilkan sebuah nilai atau pola yang nantinya digunakan pada proses testing untuk proses pelabelan tweet. Nilai atau pola yang dihasilkan dari Metode *Support Vector Machine* sebenarnya adalah sebuah garis pemisah yang disebut dengan *Hyperplane*, dimana pada kasus ini garis tersebut berperan dalam memisahkan tweet dengan sentimen positif (berlabel +1) dengan tweet yang memiliki sentimen negatif (berlabel -1). Adapun rumus dari *hyperplane* dapat dilihat pada persamaan (2-2)

Learning (atau *machine learning techniques*) dalam analisis sentimen

Dimana $w \cdot x + b = 0$ (2-2)
 w = parameter *hyperplane* yang dicari (garis yang tegak lurus antara garis *hyperplane* dan titik *support vector*)
 x = data input SVM (x_1 = index kata, x_2 = bobot kata)
 b = parameter *hyperplane* yang dicari (nilai bias)
 f = fungsi *Hyperplane*

2.6 Pengujian

Setelah sistem berhasil dibuat, diperlukan suatu pengukuran atau perhitungan untuk menilai apakah sistem tersebut sudah sesuai dengan yang diharapkan atau belum. Pada penelitian ini akan diadakan suatu penilaian dari keberhasilan sistem berdasarkan keakuratan sistem memprediksi tweet yang mengandung opini, memprediksi apakah tweet tersebut termasuk tweet positif atau negatif, dsb. Adapun beberapa cara perhitungan dalam pengujian yang dilakukan pada Penelitian ini adalah sebagai berikut : [2]

$$TP + FP$$

$$A = \frac{TP}{TP + FP + FN + TN} \quad (2-1)$$

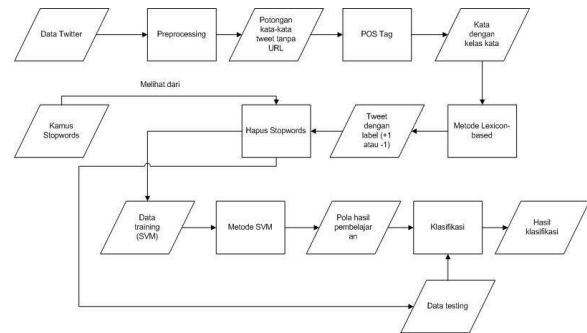
$$P = \frac{TP}{TP + FP} \quad (2-4)$$

$$F1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (2-5)$$

$$F1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}$$

3. Perancangan Sistem

Pada umumnya, sistem yang akan dibuat pada penelitian ini merupakan sebuah sistem yang dapat menganalisa sentimen yang terdapat pada media sosial Twitter mengenai suatu acara televisi. Analisis sentimen pada penelitian ini berada pada kategori *Fined-grained sentiment analysis* dimana sentimen berupa suatu kalimat (tweet) bukan suatu dokumen. Analisis sentimen yang dilakukan pada suatu acara televisi ini dilakukan menggunakan kombinasi antara metode *Lexicon-based* dengan *Support Vector Machine (SVM)*. Gambaran umum dari sistem yang akan dibuat pada penelitian adalah seperti Gambar 3-1:



Gambar 3-1 Flowchart Proses Secara Umum

Adapun penjelasan secara rinci dari Gambar 3-1 di atas:

1. Pengambilan data mentah yang berasal dari Twitter yang diambil menggunakan NodeXL [8]
2. Kemudian dilakukan proses *preprocessing* yaitu proses pembersihan data mentah yang dilanjutkan dengan proses tokenize, dan POS-tagging

$$TP + FP + FN + TN + R \quad (2-6)$$

3. Setelah itu, data diproses dengan menggunakan metode *lexicon-based*. Pada proses ini, Ada beberapa cara yang digunakan seperti memberikan penanganan terhadap kata negasi, dan juga menghitung skor polaritas (+1 atau -1) yang dimiliki oleh sebuah fitur atau entitas.

4. Setelah itu, dilakukan klasifikasi sentimen dengan menggunakan metode *Support Vector Machine* (SVM). Data yang digunakan untuk data latih merupakan tweet yang telah diberi label saat proses *lexicon-based* berhasil dilakukan, sedangkan untuk data uji (*testing*) digunakan data tweet yang sudah melalui proses *preprocessing* namun belum diberikan label pada tweet tersebut. Data yang digunakan (data latih ataupun data uji) harus melalui proses penghapusan *stopwords* terlebih dahulu sebelum diproses ke dalam metode *Support Vector Machine*.

Dari keseluruhan proses di atas akan didapatkan hasil klasifikasi tweet mengenai suatu acara televisi.

4. Analisis dan Pengujian

4.1 Dataset

Dalam melakukan proses pengujian ini digunakan 5 dataset yang berupa jenis program acara yang telah dipilih yaitu: Mata Najwa, Kick Andy, Tukang Bubur Naik Haji (TBNH), Yuk Keep Smile (YKS), dan juga Pesbukers. Pengujian ini dilakukan pada masing-masing program televisi yang ada, pembagian dataset ini dapat dilihat pada Tabel 4.1.

Tabel 4-1 Dataset yang digunakan

ACARA	Jumlah Data		JUMLAH TWEET
	negatif	positif	
Kick Andy	77	223	300
Mata najwa	78	222	300
TBNH	298	202	500
yks	252	248	500
Pesbukers	277	323	600

4.2 Analisis Hasil Pengujian

Pengujian ini dilakukan untuk menguji atau menilai hasil yang didapatkan dari sistem, dimana jika hasil yang didapatkan tersebut memiliki nilai akurasi yang tinggi maka hasil tersebut dapat diakui

kebenarannya. Adapun pengujian yang dilakukan dalam penelitian ini adalah untuk melihat hasil akurasi, presisi, recall, dan juga f-measure. Pada Tabel 4-2 sampai Tabel 4-6 merupakan hasil penilaian dari masing-masing acara dengan menggunakan kombinasi metode *lexicon-based* dan svm. Tabel 4-2 menunjukkan hasil penilaian terhadap sistem dalam menangani tweet yang membicarakan tentang program televisi Kick Andy.

Tabel 4-2 Hasil penilaian sistem untuk program televisi Kick Andy

KICK ANDY				
Latih : Uji	Akurasi	Presisi	Recall	f-measure
60% : 40%	78.33	78.15	100	87.74
	75	75.42	98.89	85.58
70% : 30%	72.22	73.86	97.01	83.87
	70	70.79	98.44	82.35
80% : 20%	75	75	100	85.71
	73.33	72.88	100	84.31
90% : 10%	70	70	100	82.35
	70	70	100	82.35

Tabel 4-3 menunjukkan hasil penilaian terhadap sistem dalam menangani tweet yang membicarakan tentang program televisi Mata Najwa.

Tabel 4-3 Hasil penilaian sistem untuk program televisi Mata Najwa

MATA NAJWA				
Latih : Uji	Akurasi	Presisi	Recall	f-measure
60% : 40%	69.17	69.17	100	81.77
	73.33	73.33	100	84.62
70% : 30%	72.22	73.86	97.01	83.87
	72.22	72.22	100	83.87
80% : 20%	71.67	71.67	100	83.5
	68.33	68.33	100	81.19
90% : 10%	80	80	100	88.89
	80	80	100	88.89

Tabel 4-4 menunjukkan hasil penilaian terhadap sistem dalam menangani

tweet yang membicarakan tentang program televisi TBNH.

Tabel 4-4 Hasil penilaian sistem untuk program televisi TBNH

TBNH				
Latih : Uji	Akurasi	Presisi	Recall	f-measure
60% : 40%	61	52.63	12.66	20.41
	61.5	47.06	10.53	17.2
70% : 30%	64	87.5	11.67	20.59
	59.33	62.5	15.38	24.69
80% : 20%	71	100	21.62	35.56
	64	66.67	15.38	25
90% : 10%	64	100	5.26	10
	60	50	10	16.67

Tabel 4-5 menunjukkan hasil penilaian terhadap sistem dalam menangani tweet yang membicarakan tentang program televisi YKS.

Tabel 4-5 Hasil penilaian sistem untuk program televisi YKS

YKS				
Latih : Uji	Akurasi	Presisi	Recall	f-measure
60% : 40%	57.5	80	16.49	27.35
	50	46.67	7.07	12.28
70% : 30%	62	76.92	28.17	41.24
	50.67	85.71	7.59	13.95
80% : 20%	56	87.5	14	24.14
	62	76.19	32.65	45.71
90% : 10%	60	76.92	37.04	50
	56	100	4.35	8.33

Tabel 4-6 menunjukkan hasil penilaian terhadap sistem dalam menangani tweet yang membicarakan tentang program televisi Pesbukers.

Tabel 4-6 Hasil penilaian sistem untuk program televisi Pesbukers

PESBUKERS				
Latih : Uji	Akurasi	Presisi	Recall	f-measure
60% : 40%	53.33	70	26.52	38.46
	47.08	56.82	18.8	28.25
70% : 30%	48.89	59.26	16.49	25.81
	59.44	75.51	37.76	50.34
80% : 20%	49.17	81.82	13.24	22.78
	50	69.57	23.19	34.78
90% : 10%	58.33	70	42.42	52.83
	46.67	50	25	33.33

Dari hasil keseluruhan diatas dapat ditarik kesimpulan bahwa perbedaan jumlah dat latih dan data uji yang digunakan tidak memberikan pengaruh terhadap nilai pengujian (akurasi, presisi, recall, dan f-measure). Selain itu nilai pengujian (akurasi, presisi, recall, dan f-measure) cenderung memiliki nilai yang lebih tinggi ketika sentimen yang dianalisis lebih banyak mengandung sentimen yang positif. Pada program acara yang banyak mengandung sentiment negatif cenderung memiliki nilai akurasi yang rendah dikarenakan banyaknya kesalahan dalam menganalisa tweet positif. Sedangkan hal yang menyebabkan nilai recall, presisi, dan juga f-measure yang rendah pada kasus ini cenderung dikarenakan perhitungan nilai recall, dan presisi yang dilakukan dalam penelitian ini merupakan nilai recall dan presisi untuk kelas positif. Jadi karena data cenderung banyak mengandung tweet negatif menyebabkan banyaknya kesalahan dalam mengklasifikasikan tweet positif sehingga menyebabkan akurasi, presisi, recall, dan f-measure rendah dalam kasus program televisi yang di dominasikan oleh sentimen negatif.

5. Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan hasil pengujian dan analisa yang telah dilakukan sebelumnya, maka dapat diambil beberapa kesimpulan sebagai berikut :

1. Metode kombinasi Lexicon-based dan SVM dapat mengklasifikasikan sentimen yang terdapat dari suatu tweet dengan keluaran positif atau negatif, dengan akurasi berkisar 50%-80% dimana acara dengan sentiment positif memiliki nilai akurasi yang lebih tinggi.
2. Dalam penelitian dihasilkan kesimpulan bahwa acara Kick Andy dan Mata najwa memiliki sentimen positif, sedangkan acara tukang Bubur Naik Haji (TBNH), Yuk Keep Smile, dan juga Pesbukers memiliki sentiment yang negatif.

5.2 Saran

Saran yang ingin disampaikan guna mengembangkan Tugas Akhir ini adalah sebagai berikut:

1. Pengembangan kamus yang digunakan untuk meningkatkan hasil pengujian
2. Data Tweet yang dianalisis diambil secara real-time (online)
3. Penggunaan pembobotan fitur lain seperti Term Presence (TP), Term Frequency (TF)
4. Penggunaan SVM dengan kernel lain seperti kernel polynomial, RBF, atau Sigmoid

DAFTAR PUSTAKA

- [1] Campagne, J.C., Dux, J., Guyot, P. dan Julien, D.(2012). Twitter reaches half a billion accounts more than 140 millions in the U.S. http://semicast.com/publications/2012_07_30_Twitter_reaches_half_a_billion_accounts_140m_in_the_US [16 November 2013]
- [2] Khairnar, J., Kinikar, M. (2013). "Machine Learning Algorithms for Opinion Mining and Sentiment Classification". *International Journal of Scientific and Research Publications*, 3 (6), 1-6.
- [3] Ley, Z., Riddhiman, G., Mohamed, D., Meichun, H., & Bing, L. (2011). "Combining lexicon-based and learning-based methods for twitter sentiment analysis". *HP Laboratories*, Technical Report HPL-2011, 89.
- [4] Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.
- [5] Lun-Wei Ku, Yu-Ting Liang, & Hsin-Hsi Chen. (2006). "Opinion Extraction, Summarization and Tracking in News and Blog Corpora ." *Proceedings of AAAI-CAAW-06, the Spring Symposia on Computational Approaches to Analyzing Weblogs* (pp. 100-107).
- [6] Morissan. (2005). *Media Penyiaran : Strategi Mengelola Radio dan Televisi*. Jakarta : Ramdina Prakarsa.
- [7] Pang, B., Lee, L., & Vithyanathan, S. (2002). "Thumbs Up ? SentimentClassification Using Machine Learning Techniques." *Proceedings of The ACL-02 conference on Empirical methods in natural language processing* (pp. 79-86). Stroudsburg: Association for ComputationalLinguistic.
- [8] Smith, M., Milic-Frayling, N., Shneiderman, B., Mendes Rodrigues, E., Leskovec, J., Dunne, C., (2010). NodeXL: a free and open network overview, discovery and exploration add-in for Excel 2007/2010, <http://nodexl.codeplex.com/> from the Social Media Research Foundation, <http://www.smrfoundation.org>
- [9] Stone, P.J., Hunt, E.B. (1963). "A Computer Approach to Content

Analysis: Studies using the General Inquirer System". *In Proceedings of the Spring Joint Computer Conference* (pp.241–256), New York, NY, USA: ACM.

- [10] Yayasan, SET., Yayasan, TIFA., IJTI. The Habibie Center & LSPR. "Menuju Televisi yang Ramah Keluarga". Rating Publik

