

# PERAN *OPINION MINING* DAN *SENTIMENT ANALYSIS* UNTUK MENDENTIFIKASIKAN SENTIMEN PUBLIK DALAM SISTEM *E-GOVERNANCE*

Sherly Christina <sup>1)</sup>, Deddy Ronaldo <sup>2)</sup>

<sup>1</sup> Fakultas Teknik, Universitas Palangka Raya  
sherly.christina.upr@gmail.com

<sup>2</sup> Fakultas Teknik, Universitas Palangka Raya  
deddy.ronaldo@gmail.com

## *Abstract*

*People use text to communicate their emotion and opinion about some subjects. Currently government online media provides the facility for the citizen to express their opinions through text. E-governance system that built by the government will function better if the data that contains citizen's opinion not only saved but can be managed to be more valuable information.*

*Opinion Mining and Sentiment Analysis is the science of natural language processing that could identify emotions, sentiments and main ideas expressed in the text. The government can take advantage by embedding the opinion mining and sentiment analysis methods in the e-governance system.*

*Citizen's opinion about the services or policies of the government is an information that required by the Government in running the good government system. In this study, will be discussed some methods of opinion mining and sentiment analysis that have been performed by some previous researchs.*

**Keywords :** *opinion mining, sentiment analysis, e-governance*

## 1. PENDAHULUAN

Salah satu informasi yang dibutuhkan untuk membuat strategi dan mengambil keputusan adalah pendapat orang lain. Apa yang dipikirkan oleh orang lain terhadap suatu subjek dapat memberikan kontribusi terhadap optimalnya strategi, pengambilan keputusan atau tindakan terkait subyek tersebut.

Media *online* yang difasilitasi internet dan web seperti media sosial, kolom opini atau buku tamu memberikan kemudahan untuk mengumpulkan opini publik terhadap suatu subjek. Menurut kamus besar bahasa Indonesia [7] opini memiliki arti pendapat, pikiran dan pendirian. Pada bidang *e-commerce*, opini publik terhadap suatu produk adalah salah satu informasi yang dibutuhkan untuk kepentingan analisis dan strategi pemasaran produk. Sedangkan dalam bidang pemerintahan, opini publik dapat menjadi tolak ukur dari kinerja pemerintah.

Saat ini instansi-instansi pemerintahan di Indonesia menggunakan kemajuan teknologi dan internet untuk memfasilitasi pelayananan terhadap publik dalam bentuk *E-governance*. Tujuan utama dari *E-governance* adalah peningkatan efisiensi, kenyamanan serta aksesibilitas terhadap layanan publik. *E-governance* selain memberikan informasi dan layanan terhadap publik, juga telah dilengkapi

dengan sarana untuk menampung opini dari publik.

Opini publik terhadap layanan yang diberikan pemerintah adalah salah satu informasi yang dibutuhkan oleh pemerintah untuk menganalisis kebijakan yang telah dibuat, membuat perencanaan serta pengambilan keputusan. Sistem *E-governance* yang ada sekarang akan bermanfaat lebih optimal bila opini publik yang ditampung tidak hanya disimpan tetapi bisa diolah menjadi informasi yang lebih bernilai dan mudah dipahami misalnya dalam bentuk grafik.

*Opinion mining* adalah bidang ilmu yang menganalisis pendapat, sentimen, evaluasi, penilaian, sikap dan emosi seseorang terhadap suatu entitas seperti produk, layanan, individu, peristiwa dan isu tertentu [6]. Opini yang disampaikan oleh publik dapat menunjukkan sentimen negatif atau positif terhadap suatu entitas. *Sentiment Analysis* adalah salah satu *natural language processing* untuk melacak *mood* dari seseorang terhadap suatu produk atau jasa [1][3][4].

Pada penelitian ini dilakukan studi literatur terhadap beberapa penelitian di bidang *opinion mining* dan *sentiment analysis*. Beberapa penelitian tersebut memberikan data dan informasi bagaimana kinerja pemrosesan teks pada bidang ilmu *opinion mining* dan

*sentiment analysis*. Sehingga dari beberapa penelitian yang telah dilakukan bisa dikumpulkan informasi yang menunjukkan bagaimana *opinion mining* dan *sentiment analysis* dapat meningkatkan manfaat dari *E-governance*.

## 2. PEMBAHASAN

Nasukawa dan Yi [9] melakukan penelitian untuk mengidentifikasi sentimen yang dinyatakan seseorang terhadap suatu subyek (*favorability analysis*) dan menentukan polaritas atau kecenderungan dari sentimen tersebut. Menurut [9], esensi dari *Sentiment Analysis* adalah mengidentifikasi sentimen yang dinyatakan seseorang dalam bentuk teks dan apakah sentimen tersebut menyatakan kesukaan (*favor*) atau ketidaksukaan (*unfavor*). Dalam *framework* yang diusulkan oleh [9], terlebih dahulu didefinisikan kata sifat, kata keterangan kata benda, dan kata kerja yang mengekspresikan polaritas positif atau negatif. Kemudian menghubungkannya dengan argumen seperti subyek dan obyek yang berisi sentimen.

Pernyataan-pernyataan atau ungkapan-ungkapan sentimen didefinisikan dalam bentuk kamus dengan menggunakan notasi-notasi yang berisi informasi:

- Polaritas, *positif/good* (p), *negatif/bad* (b) atau netral (n)
- Part of speech* (POS), kata sifat (JJ), kata keterangan (RB) dan kata benda (NN)
- Istilah-istilah dalam bentuk kanonik
- Argumen seperti subyek(sub) dan obyek(obj) yang menerima sentimen dari suatu kata kerja

Algoritma Nasukawa dan Yi melakukan pendekatan *sentiment analysis* dengan cara mencari ekspresi yang menyatakan sentimen dan secara semantik terkait dengan *subject term* dalam teks, kemudian mengidentifikasi polaritas dari ekspresi sentimen tersebut. Mengidentifikasi ekspresi sentimen dan analisa hubungan semantik menggunakan POS *tagging*. Selanjutnya dilakukan penguraian kalimat untuk mengidentifikasi frasa dan ketergantungan antar kata. Kemudian polaritas sentimen ditentukan berdasarkan definisi dari kamus bila tidak ditemukan ekspresi negatif seperti "not" atau "never". Polaritas sentimen akan bernilai +1 (*positif=favorable*) atau -1 (*negatif=unfavorable*).

Berikut adalah contoh eksperimen yang dilakukan oleh Nasukawa dan Yi. Inputan pada

eksperimen ini adalah *subject term*. *Subject term* adalah entitas yang menjadi target analisis oleh sistem dari suatu potongan teks (*text fragment*). Output sistem adalah polaritas dari *sentiment term* yang muncul dalam *text fragment* terhadap *subject term*. *Subject term* dan *sentiment terms* dihubungkan oleh "---" dalam bentuk kanonik yang terkait dengan seluruh frasa dari *text fragment*.

### Contoh 1:

<input> (subject="MDX")

*For 2002, the MDX features the same comfort and exhilaration, with an even quieter ride.*

<output>

+1 MDX (the MDX)---feature (features)---comfort

(the same comfort and exhilaration)

### Contoh 2:

<input> (subject="IBM")

*Of the seven stakeholder groups, IBM received the highest score in the ranking for its policies and programs for minorities and women.*

<output>

+1 IBM (IBM)---receive (received)---high score (the highest score in the ranking)

Eksperimen Nasukawa dan Yi menghasilkan nilai presisi 95% dan *recall* 20%. Metode yang dirancang oleh Nasukawa dan Yi memiliki kelemahan untuk menganalisis polaritas sentimen pada struktur kalimat yang kompleks, contohnya "*It's not that it's a bad camera*". Walaupun begitu metode yang dirancang oleh Nasukawa dan Yi telah membantu mengidentifikasi kata dalam teks yang mengekspresikan sentimen dan mengetahui polaritas dari sentimen tersebut. Kelak dengan mengidentifikasi fitur pada kalimat yang menyebabkan ambiguitas, kinerja dari sistem dapat ditingkatkan.

Binali dkk [5] menyatakan bahwa emosi sangat mempengaruhi pengambilan keputusan. Menurut Wordnet Search [10] emosi memiliki makna "*any strong feelings*". Binali dkk melakukan penelitian untuk mendeteksi emosi yang tertuang dalam teks dengan menggunakan arsitektur hybrid dan algoritma *Support Vector Machine* untuk menguji arsitektur tersebut.

Pendeteksian emosi dalam teks bertujuan untuk mengetahui emosi yang mempengaruhi penulis teks. Premis yang mendasarinya adalah jika seseorang bahagia maka ia akan menggunakan kata-kata yang positif. Demikian juga jika seseorang sedih, frustrasi atau marah, ia cenderung menggunakan kata negatif.

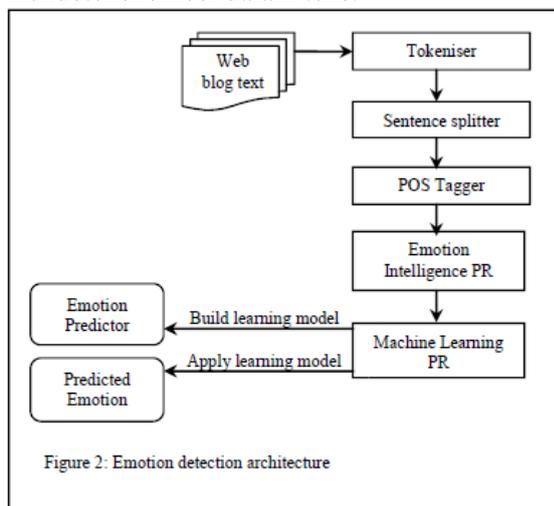
Arsitektur yang diusulkan oleh [5] bertujuan untuk memenuhi kebutuhan:

1. Pendeteksian *hidden phrase pattern*, kalimat-kalimat yang mengekspresikan emosi harus diidentifikasi secara sintatik dan semantik.
2. Pendeteksian emosi, mengidentifikasi emosi berdasarkan sumber pengetahuan linguistik dan *trained classifiers*.

Untuk memenuhi kebutuhan tersebut, dibuat desain yang terdiri atas:

1. mekanisme untuk mengidentifikasi kata kunci
2. contoh dari *hidden pattern* yang diperoleh dari seorang pakar untuk membantu dalam *trained classifiers*.

Gambar 2.1 menunjukkan arsitektur hybrid yang dibangun oleh [5] untuk mendeteksi emosi dalam teks.



Gambar 2.1 Hybrid Arsitektur dari Emotion Detection (Binali dkk, 2010).

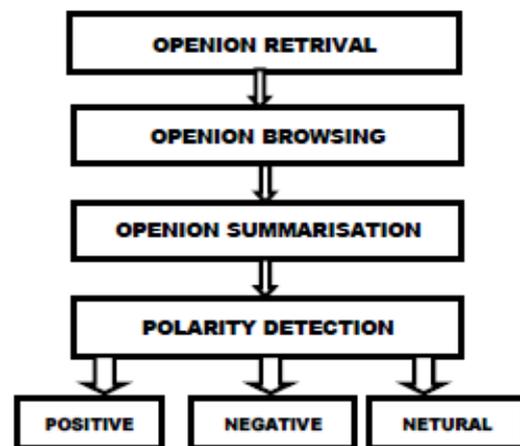
Tahapan pertama pada arsitektur adalah melakukan prapemroses terhadap teks menjadi token menggunakan tokeniser. Kemudian teks dipecah menjadi kalimat-kalimat menggunakan *sentence splitter*. Kemudian dari kalimat-kalimat tersebut diidentifikasi kata kerja, kata benda, kata keterangan dan kata sifat menggunakan *POS Tagger*. Kemudian informasi ini digunakan oleh *Emotion Intelligence PR* untuk membuat anotasi baru. *Intelligence PR* akan mengambil informasi semantik dari kata-kata tersebut dari kamus yang telah dibuat oleh seorang pakar untuk mengidentifikasi emosi-emosi yang terdapat dalam kalimat tersebut. Langkah terakhir adalah mengkategorikan emosi positif dan emosi negatif. *Machine Learning PR* bekerja dalam dua langkah, pertama kali

dibangun *training model* menggunakan *emotion predictor*. Langkah kedua mengaplikasikan *trained model* pada pengujian dataset, untuk memprediksikan emosi positif atau emosi negatif.

Hasil eksperimen [5] menunjukkan tingkat akurasi 96,43% pada data dari blog web. Arsitektur hybrid untuk mendeteksi emosi yang menunjukkan hasil pengujian dalam informasi sintatik dan semantik dapat meningkatkan akurasi untuk memprediksi emosi seseorang atau sentimen publik terhadap suatu entitas.

Kumar dkk [6] menyatakan bahwa *opinion mining* merupakan kesempatan untuk menemukan sumber permasalahan lebih cepat sebelum masalah itu terjadi. Warga negara mengharapkan pelayanan yang cepat, tindakan yang tepat dan tanggapan yang segera terhadap setiap keluhan dan kebutuhan. Menurut [6] interaksi antar pemerintah dan warganya tidak selalu mudah dan menyenangkan karena tingkatan pemahaman atau kelas yang berbeda dalam masyarakat dan dampak terhadap setiap kebijakan dan proyek juga berbeda. Pemerintah diharapkan bersikap transparan, bertanggung-jawab dan cepat dalam melaksanakan tugas dan memberikan layanan terhadap warganya. Oleh karena itu *opinion mining* dan *sentiment analysis* dapat menjembatani interaksi antar pemerintah dan warganya melalui sistem *E-governance*.

Penelitian oleh [6] bertujuan menemukan sentimen dari berbagai kelas masyarakat terhadap *E-governance* India khususnya pada proyek *Smart City*. Gambar 3.2 menunjukkan metode *opinion mining* yang dirancang oleh Kumar dkk[6].



Gambar 3.2 Metode *Opinion Mining* (Kumar dkk, 2015).

Tahap pertama dalam metode *opinion mining* tersebut adalah mengumpulkan pendapat dari warga terhadap proyek *Smart City*. Pada penelitian [6] warga India dikelompokkan menjadi 3 kelas:

1. *Upper Class*: kelompok warga yang sangat memahami detail proyek *Smart City* dan kebijakan-kebijakan *E-governance*. Kelompok warga yang sangat aktif menggunakan internet
2. *Middle Class*: kelompok warga yang cukup tahu tentang proyek *Smart City* dan merasa terbantuan dengan fasilitas di internet.
3. *Lower Class*: kelompok warga yang hampir sama sekali tidak memahami proyek *Smart City* dan tidak menggunakan fasilitas internet.

Implementasi dari tiap langkah pada metode *Opinion Mining* dijelaskan secara singkat seperti berikut:

1. *Extraction of Information*, sebuah algoritma untuk mengidentifikasi semua relasi dan informasi yang dibutuhkan.
2. *Categorizing the opinions*, mengidentifikasi ide utama dari suatu dokumen, mencari pendapat, sentimen dalam dokumen tersebut terkait suatu entitas. Pada proses ini dilakukan penghitungan kata-kata yang muncul pada teks. Kemudian pengkategorian dilakukan berdasarkan sebuah glosarium dan relasi seperti sinonim serta istilah lain yang terkait.
3. *Clustering*: Sebuah algoritma akan menciptakan berbagai topik bagi tiap dokumen dan mengukur kecocokan dokumen dalam tiap kluster.
4. *Tracking of Topic*: Sebuah sistem *text mining* digunakan untuk menemukan minat dari penulis teks.
5. *Summarization: text summarization* dilakukan untuk merangkum hal-hal utama dari dokumen, seperti fitur-fitur dari entitas dan pendapat dari penulis teks terhadap entitas tersebut.

*Naive Bayes Classifier* digunakan untuk menentukan sentimen dari kalimat-kalimat dalam teks. *Sentiment Analysis* dibagi dalam kelas positif dan kelas negatif. Hasil percobaan terhadap data-data yang digunakan dalam penelitian [6] menunjukkan bahwa *sentiment analysis* membantu mengidentifikasi sikap dari warga terhadap kebijakan dan proyek yang sedang dikerjakan pemerintah. *Sentiment*

*analysis* membantu memahami sikap publik terhadap berbagai konteks, produk maupun layanan.

Pada penelitian lain oleh Anurachalam dkk [8] dinyatakan dalam judul penelitiannya bahwa ilmu *Sentiment Analysis* adalah *New eye of Government*. Jika instansi pemerintah dapat mengikuti detak kehidupan warganya secara konstan, maka dapat terbuka jalan untuk terbentuknya sistem pemerintahan yang lebih baik.

Anurachalam dkk[8] membangun sebuah arsitektur yang didukung oleh IBM System-Cognos Consumer Insight (CCI) untuk memodelkan topik dan mengidentifikasi sentimen dari suatu teks. Hasil percobaan terhadap sejumlah dataset menunjukkan bahwa analisis yang diperoleh dari metode dalam arsitektur tersebut dapat menyimpulkan polaritas sentimen dari publik.

### 3. KESIMPULAN

Penelitian yang dilakukan oleh [5][6][8][9] menunjukkan beberapa metode *Opinion Mining* dan *Sentiment Analysis* yang berhasil mengidentifikasi sentimen publik dari pemrosesan teks. Bidang ilmu ini memungkinkan pengetahuan dan informasi tentang kecenderungan publik terhadap suatu subyek atau entitas dapat dikumpulkan lebih dini sebelum masalah terjadi. Pengetahuan dan informasi tersebut dapat digunakan untuk menyusun strategi pemecahan masalah maupun mengelola suatu entitas.

Dengan demikian *Opinion mining* dan *sentiment analysis* dapat memberikan manfaat lebih bagi sistem pemerintahan. Sistem *E-governance* yang dapat mengumpulkan dan mengelola informasi dari opini dari warganya, dapat meningkatkan kinerja dari pemerintah dalam memberikan layanan secara cepat dan tepat.

### 4. REFERENSI

- [1] Bing Liu, *Sentiment Analysis and Opinion Mining*, 2012. Morgan & Claypool Publishers
- [2] Bo Pang, Lillian Lee, *Opinion Mining And Sentiment Analysis*, Foundations and Trends® in Information Retrieval: Vol.2 No.1-2, 2008
- [3] Erick Cambria, Bjorn Chuller, Yunqing Xia, Catherine Havasi, *New Avenue in Opinion Mining and Sentiment Analysis*, 2013. IEEE Computer Society

- [4] G.Vinodhini, R.M. Chandrasekaran, *Sentiment Analysis and Opinion Mining: A Survey*, 2012. International Journal of Advanced Research in Computer Science and Software Engineering Vol.2
- [5] Haji Binali, Chen Wu, Vidyasagar Potdar, *Computational Approach for Emotion Detection in Text*, 2010. 4th IEEE International Conference on Digital Ecosystems and Technologies (IEEE DEST 2010)
- [6] Pankaj Kumar, Alisha Zaidi, Ankita Srivastava, Geet Jain, *Evolving Sentiment towards E-Governance using Opinion Mining*, 2015. International Journal of Computer Applications(0975-8887), International Conference on Distributed Computing and Internet Technology
- [7] Pusat Bahasa Departemen Pendidikan Nasional, 2008, Kamus Besar Bahasa Indonesia, Jakarta
- [8] Ravi Arunachalam, Sandipan Sarkar, 2013. *The New Eye of Government: Citizen Sentiment Analysis in Social Media*, 2013. Workshop on Natural Language Processing for Social Media (Social NLP) pages 23-28, Nagoya, Japan
- [9] Testsuya Nasukawa, Jeonghee Yi, *Sentiment Analysis: Capturing Favorability Using Natural Language Processing*, 2003. Proceedings of The 2nd International Conference on Knowledge Capture Pages 70-77, ACM ©2003 ISBN:1-58113-583-1
- [10] Wordnet 2.1 Browser