

Identifikasi Fitur Laptop beserta Orientasinya dengan Metode Apriori dan *Lexicon-Based*

Try Satria Amanattullah^{1*}, Hilda Widyastuti^{2*}, Festy Winda Sari^{3*}

* Teknik Informatika, Politeknik Negeri Batam
hilda@polibatam.ac.id²,

Article Info

Article history:

Received 02-10-2017
Revised 19-10-2017
Accepted 24-11-2017

Keyword:

Identifikasi Fitur
Apriori
Orientasi Fitur
Lexicon-Based
Opinion Mining

ABSTRACT

Perkembangan laptop saat ini sangat pesat. Para produsen menggunakan media sosial misalnya fan page di Facebook untuk mempromosikan produknya. Semakin banyaknya pilihan membuat seseorang kesulitan dalam menentukan suatu laptop bagus atau tidak, direkomendasikan atau tidak. Review-review dan komentar-komentar yang ada di fan page tentang merek-merek laptop baru bisa dijadikan sebagai pengetahuan untuk menentukan apakah laptop baru tersebut bagus atau tidak. Jumlah review dan komentar yang ada di fan page sangat banyak sehingga diperlukan proses otomatisasi. Untuk keperluan tersebut digunakan opinion mining mencakup identifikasi target opini dan penentuan orientasinya. Identifikasi target opini digunakan untuk mengetahui fitur-fitur laptop yang dibicarakan dalam sebuah komentar sedangkan penentuan orientasi digunakan untuk menentukan apakah komentar bersifat positif atau negatif. Data yang digunakan untuk penelitian ini diambil dari data fans page Facebook yang kemudian dianalisis menggunakan metode Apriori untuk menghasilkan fitur laptop sebagai target opini dan metode Lexicon-based untuk menentukan orientasi fitur laptop, apakah berorientasi positif atau negatif. Penelitian juga menghasilkan kesimpulan dari data review dan komentar yang telah diproses.

Copyright © 2017 Journal of Applied Informatics and Computing.
All rights reserved.

I. PENDAHULUAN

Menurut [3] opini adalah status seseorang yang mencerminkan ide, kepercayaan, penilaian, dan evaluasi mengenai suatu subjek atau topik atau item. Opini masyarakat mempunyai pengaruh yang besar dan menyediakan petunjuk bagi seorang individu, pemerintah, organisasi, dan komunitas sosial dalam proses pengambilan keputusan[4]. Dengan menggunakan opini masyarakat, seseorang bisa mengintegrasikan pandangan yang berbeda-beda, pengalaman, kebijaksanaan dan pengetahuan yang dimiliki banyak orang pada proses pembuatan keputusan. Dalam kehidupan sehari-hari orang biasanya bertanya kepada teman-temannya, anggota keluarganya, dan ahli suatu bidang untuk mendapatkan informasi yang dibutuhkannya, dan opini dan perspektif mereka didasarkan kepada pengalaman, observasi, konsep, dan kepercayaan. Berikut contoh opini, "ugasnya sangat sulit", "makanan di hotel sangat lezat", "harga baju Ali murah sekali", "pemandangan gunung Bromo sangat indah".

Opinion mining adalah analisis review entitas-entitas[9]. *Opinion mining* merepresentasikan sebuah model untuk mengklasifikasikan polaritas sebuah dokumen ke dalam kelompok direkomendasikan atau tidak direkomendasikan. Menurut [3] *opinion mining* adalah salah satu bidang data

mining yang menggunakan teknik pembelajaran mesin yang bersifat statistik untuk membedakan teks-teks yang bersifat opini dengan teks-teks yang bersifat fakta. Ruang lingkup *opinion mining* mencakup identifikasi opini, klasifikasi opini, identifikasi target opini, identifikasi sumber, dan pembuatan rangkuman opini. *Opinion mining* diterapkan pada mesin pencarian, sistem pemberi rekomendasi, penyaring email, penyaring iklan di web, sistem tanya jawab, dan lain-lain.

Menurut[3] target opini mengacu kepada orang, objek, fitur, *event* atau topik tentang opini yang dibicarakan. Menurut [4], analisis setiap aspek dari sebuah produk berdasarkan opini konsumen itu penting bagi publik, baik yang berperan sebagai pembeli maupun pembuat produk. Untuk membandingkan review, proses identifikasi dan ekstrak fitur-fitur yang dibicarakan di dalam review harus dilakukan secara otomatis. Di sebagian besar kalimat, target opini ditampilkan secara eksplisit, tetapi di beberapa kalimat target opini ditampilkan implisit sehingga membutuhkan proses identifikasi lebih lanjut. Identifikasi fitur secara eksplisit dilakukan dengan identifikasi kata benda[7].

Perkembangan laptop saat ini sangat pesat. Setiap tahunnya dikeluarkan produk laptop terbaru yang siap bersaing di pasaran. Para produsen menggunakan media sosial, salah satunya *fan page* di Facebook untuk

mempromosikan produknya. Pengguna Facebook bisa mengetahui produk-produk yang telah dikembangkan oleh produsen. Selain itu produsen juga bisa mengetahui pendapat masyarakat tentang produk laptop yang dibuat perusahaannya melalui komentar-komentar pengguna Facebook di *fan page* tersebut. Semakin banyaknya pilihan membuat seseorang kesulitan dalam menentukan suatu laptop bagus atau tidak, direkomendasikan atau tidak. *Review-review* dan komentar-komentar yang ada di *fan page* tentang suatu merek laptop baru bisa dijadikan sebagai pengetahuan untuk menentukan apakah laptop baru tersebut bagus atau tidak. Jumlah review dan komentar yang ada di *fan page* sangat banyak sehingga diperlukan proses otomatisasi. Untuk keperluan tersebut digunakan *opinion mining* mencakup identifikasi target opini dan penentuan orientasinya. Identifikasi target opini digunakan untuk mengetahui fitur-fitur laptop yang dibicarakan dalam sebuah komentar sedangkan penentuan orientasi opini digunakan untuk menentukan apakah fitur-fitur laptop bersifat positif atau negatif.

Aplikasi fitur atau topik pembahasan suatu produk pada Twitter telah dibuat oleh [2]. Akan tetapi aplikasi ini memiliki kekurangan, tidak bisa menentukan sentimen sebuah komentar. Sedangkan aplikasi untuk menentukan orientasi telah dibuat oleh [1], dengan metode *Lexicon-based*. Aplikasi ini digunakan menentukan orientasi sebuah komentar apakah positif atau negatif. Tetapi aplikasi ini tidak bisa menentukan fitur apa yang dibahas pada komentar tersebut sehingga produsen yang menggunakan aplikasi tersebut tidak mengetahui topik permasalahan apa yang sering dikomentari masyarakat.

Pada penelitian ini dibuat pengembangan aplikasi [2] dan [1], berupa aplikasi yang bisa menentukan orientasi pada fitur atau topik produk laptop. Untuk penentuan fitur atau topik digunakan metode Apriori, sehingga produsen bisa lebih mudah mengetahui topik permasalahan yang sering dibicarakan masyarakat. Sesudah itu digunakan metode *Lexicon-based* untuk menentukan orientasi komentar apakah bersifat positif atau negatif, sehingga pengguna Facebook bisa mengetahui fitur yang sering dikomentari beserta orientasinya.

II. METODOLOGI PENELITIAN

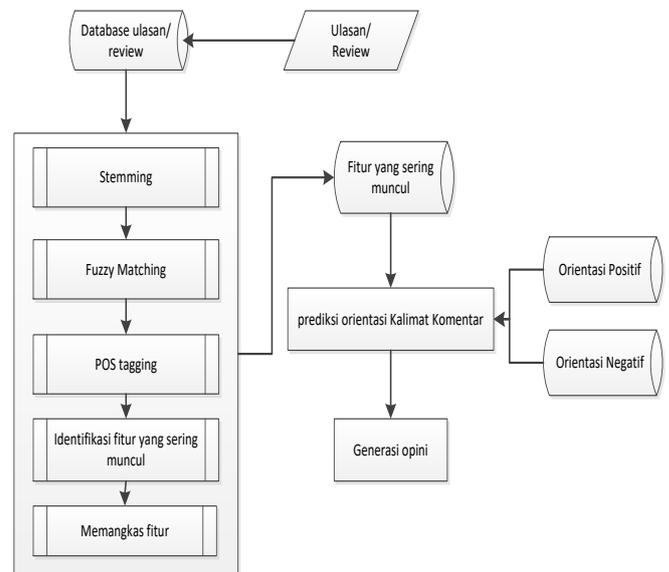
Alur proses identifikasi fitur atau topik beserta orientasinya yang dikembangkan berdasarkan hasil penelitian [8] dan [1] dapat dilihat di gambar 1 yaitu:

1. Ekstraksi fitur, terdiri dari stemming, *Fuzzy Matching*, dan *Part of Speech Tagging (POS Tagging)*
 Stemming adalah proses perubahan bentuk suatu kata menjadi bentuk dasarnya. *Stemming* bertujuan menghilangkan imbuhan-imbuhan baik berupa awalan, akhiran, maupun sisipan pada setiap kata. Di dalam proses ini dilakukan *tokenizing* terlebih dahulu untuk memisahkan kalimat menjadi sekumpulan token, yang di penelitian ini token berupa kata-kata. Langkah selanjutnya adalah *Fuzzy Matching* yang bertujuan memperbaiki kata-kata yang salah eja. Metode ini

bekerja dengan cara menemukan kata yang cocok dengan sekumpulan kata atau frase di database. Jika persentase kemiripannya melebihi ambang batas yang ditetapkan, maka kata tersebut akan diubah menjadi dengan kata termirip yang ada pada database. Langkah terakhir adalah *POS Tagging*. Fitur atau topik suatu produk biasanya kata benda[7]. Untuk mengetahui jenis katanya setiap kata harus diberi label jenis katanya dengan *POS Tagging*. Pelabelan ini menggunakan *NL processor linguistik parser* untuk menghasilkan *POS tag* setiap kata, misalnya kata benda, kata kerja, kata sifat, dan lain-lain. Data kata-kata beserta *POS tag*-nya disimpan di dalam file transaksi.

2. Identifikasi fitur yang sering muncul

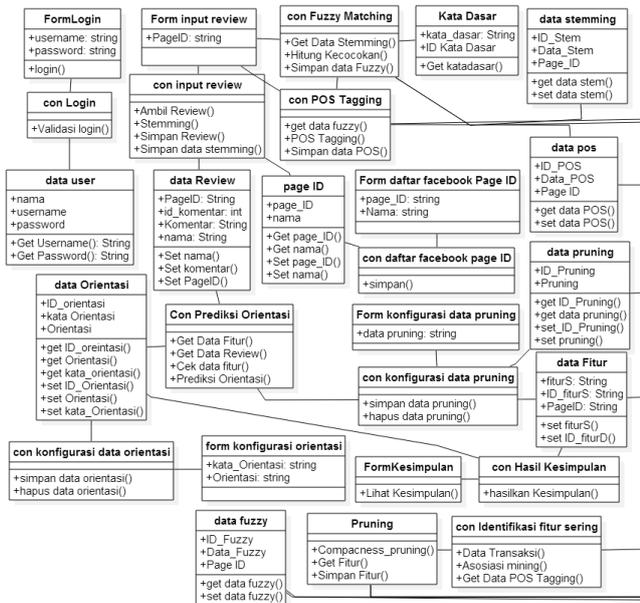
Bagian ini menemukan fitur yang paling banyak dibicarakan dengan menggunakan *association rule mining* dengan algoritma Apriori untuk menemukan semua *itemset* yang sering muncul dan memangkas fitur yang mempunyai nilai *support* lebih rendah dari *minimum support* yang ditetapkan[6][8].



Gambar 1 Alur proses identifikasi fitur dan orientasinya

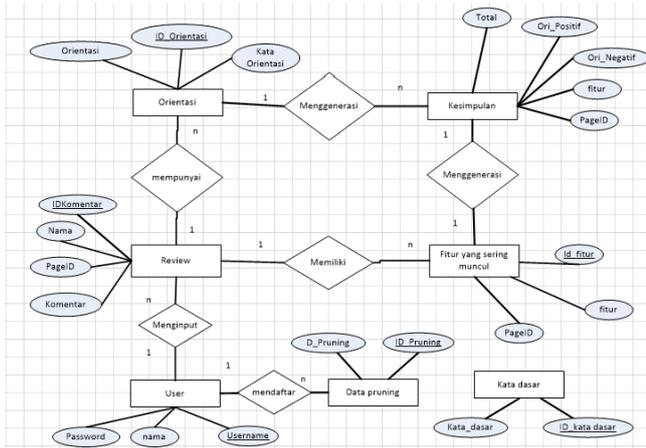
3. Identifikasi orientasi kalimat opini dengan *lexicon-based*.
 Caranya dengan menghitung jumlah kemunculan kata opini positif dan kata opini negatif yang ada di komentar berdasarkan daftar kata opini positif dan kata opini negatif yang sudah terlebih dahulu disimpan di kamus. Kata opini di sini berjenis kata sifat yang sudah ditandai saat *POS tagging*. Jika jumlah kata opini lebih besar dari nol, maka sentimen untuk kalimat tersebut adalah positif, namun jika jumlah kata opini kurang dari nol maka nilai sentimen untuk kalimat tersebut adalah negatif [1]. Bagian ini digunakan untuk memprediksi orientasi kalimat opini apakah positif atau negatif

III. HASIL DAN PEMBAHASAN



Gambar 3 Class diagram

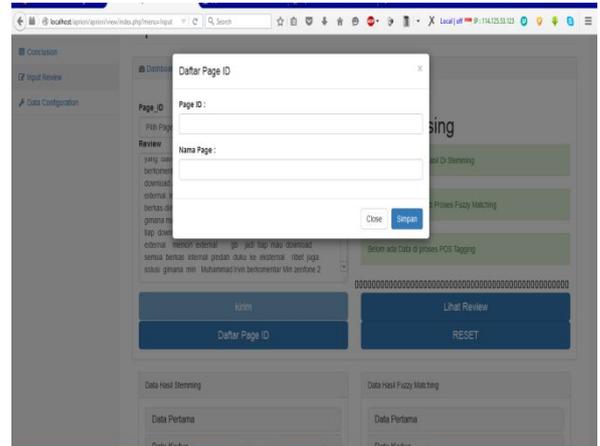
Class diagram menjelaskan hubungan antar class di mana masing-masing class memiliki atribut dan metode sendiri-sendiri. Pada class diagram di atas terdapat beberapa form Graphic User Interface yaitu login, input review, konfigurasi data pruning, konfigurasi data orientasi, kesimpulan yang terhubung dengan controller class dan entity class masing-masing. Sedangkan untuk perancangan penyimpanan data digunakan diagram Entity Relationship (ER) pada gambar 4, yang merupakan sebuah konsep yang mendeskripsikan hubungan antara penyimpanan dan didasarkan pada pedoman persepsi dari sekumpulan objek yang disebut entity dan relasi.



Gambar 4 Diagram ER

3.1. Hasil Implementasi Input (Daftar Page ID)

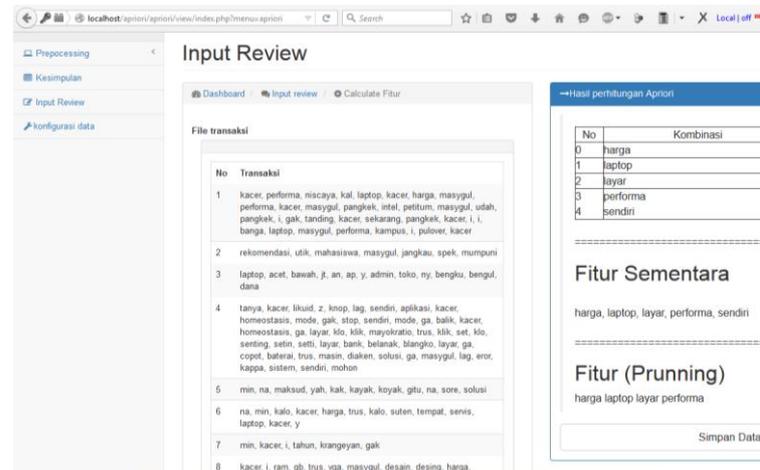
Implementasi antarmuka input ini digunakan admin untuk memasukkan komentar Facebook ke dalam database, melihat data komentar, melakukan proses *Stemming*, *fuzzy matching*, dan *POS Tagging* pada data komentar yang telah dimasukkan ke dalam database.



Gambar 5 Implementasi antarmuka input (Daftar page ID)

3.2. Hasil Implementasi Antarmuka Fitur yang sering Muncul

Implementasi fitur yang sering muncul berfungsi untuk mencari fitur dari data komentar yang telah dimasukkan ke dalam database dengan menggunakan metode Apriori. Fitur yang dianggap tidak sesuai akan dipangkas.



Gambar 6 Implementasi antarmuka fitur yang sering muncul

3.3. Implementasi Kesimpulan

Implementasi kesimpulan digunakan admin dan end user untuk melihat hasil kesimpulan yang telah diproses.

Fitur	Positif	Negatif	Kesimpulan	Komentar Pengguna
harga	3	4	Negatif	Dihitung Dari 5 orang komentar Pengguna
laptop	5	4	Positif	Dihitung Dari 5 orang komentar Pengguna

Gambar 7 Implementasi antarmuka kesimpulan

V. KESIMPULAN

Kesimpulan penelitian ini adalah:

1. Aplikasi dapat mengambil data komentar Facebook menggunakan *page ID*.
2. Aplikasi dapat melakukan *text preprocessing* berupa *Stemming*, *Fuzzy Matching*, dan *POS Tagging*.
3. Aplikasi dapat menentukan fitur dari komentar atau review dengan algoritma Apriori.
4. Aplikasi dapat menentukan orientasi positif dan negatif dari sebuah *review* dengan algoritma *Lexicon-based*.
5. Aplikasi dapat menghasilkan kesimpulan dari data *review* yang telah diproses.

Sedangkan untuk pengembangan aplikasi di masa yang akan datang saran yang dapat diberikan adalah:

1. Ditambahkan fitur pencarian *page ID* Facebook, karena pada penelitian ini untuk mencari *Page ID* Facebook harus membuka web www.findmybid.com terlebih

dahulu.

2. Sumber data yang digunakan bisa lebih beragam tidak hanya Facebook, tapi media sosial yang lain seperti Twitter.

DAFTAR PUSTAKA

- [1] J. Bahisindra, 2015, "Aplikasi Analisis Sentimen dengan Metode *Lexicon-Based* Untuk Menilai Produk Smartphone Baru", *Teknik Informatika, Politeknik Negeri Batam*, 2015.
- [2] R. Gusnaldy, "Aplikasi Menentukan Topik Pembahasan suatu Produk pada Twitter", *Teknik Informatika, Politeknik Negeri Batam*, 2015.
- [3] K.Khan, B. Baharudin, A. Khan, A. Ullah, "Mining Opinion Components from Unstructured Reviews: A review", *Journal of King Saud University- Computer and Information Sciences*, 2014.
- [4] L. Zhang, B. Liu, "Identifying noun product features that imply opinions", in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers, vol. 2. Association for Computational Linguistics*, 2011, pp. 575-580.
- [5] B. Liu, "Sentiment analysis and subjectivity", in *Indurkha, N., Damerou, F.J., (Eds.), Handbook of Natural Language Processing*, 2010, pp. 627-666.
- [6] Erwin, "Analisis Market Basket dengan Algoritma Apriori dan FP-Growth", *Jurnal Generic*, 2009.
- [7] L. Ferreira, N. Jakob, I. Gurevych, A comparative study of feature extraction algorithms in customer reviews. Paper presented at the Semantic Computing, 2008.
- [8] M. Hu, B. Liu, "Mining and summarizing customer reviews", in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2004, pp. 168-177.
- [9] K. Dave, S. Lawrence, D. M. Pennock, "Mining the peanut gallery: opinion extraction and semantic classification of product reviews", in *www'03 Proceedings of the 12th international conference on World Wide Web*, 2003.