

IMPLEMENTASI RECOMMENDER SYSTEM BASED ON SENTIMENT CLASSIFICATION MELALUI OPINION EXTRACTION (STUDI KASUS REVIEW USER DI GOOGLE PLAY)

Muhammad Arief Kurniawan¹, Ema Rachmawati², Bambang Pudjoatmodjo³

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Sentiment Classification dapat memberikan nilai pada suatu teks apakah teks tersebut termasuk ke dalam “negatif”, “netral” atau “positif”. Dari Sentiment Classification tersebut dapat memberikan suatu rekomendasi kepada pengguna dalam bentuk tekstual. Recommender System dapat dijadikan cara untuk memberikan rekomendasi suatu produk baru kepada pengguna. Sebagian besar deskripsi aplikasi/produk, opini dari pengguna dan sebagainya disajikan dalam bentuk tekstual di website. Ada banyak cara dalam menilai suatu produk yang ditawarkan, contohnya ialah dengan pengguna memberi penilaian dengan memberi “Like” atau “Dislike” atau pengguna memberikan bintang dari skala 1(kurang bagus) sampai 5 (sangat bagus). Pada Tugas akhir ini, Sentiment Classification akan dilakukan melalui Opinion Extraction, dimana dalam tahap ini akan dilakukan pengolahan kata fitur produk dan kata opini, proses parsing dengan menggunakan Stanford Parser untuk mendapatkan hubungan gramatikal dalam setiap kalimat, menentukan pasangan kata fitur produk dan opini, menentukan kekuatan dan polaritas kata opini melalui SentiWordNet lalu mengakumulasi nilai akhir untuk setiap review. Jika nilai akhir Sentiment tersebut positif maka pengguna tersebut merekomendasikan aplikasi yang dibicarakan, jika negatif sebaliknya. Kemudian dengan menggunakan teknik Item Based Collaborative Filtering Recommender System, kita dapat memberikan suatu rekomendasi kepada pengguna berdasarkan dari aplikasi-aplikasi yang pernah mereka nilai sebelumnya. Berdasarkan hasil pengujian menggunakan metode Mean Opinion Score (MOS), Recommender System yang dibangun dengan Sentiment Classification melalui Opinion Extraction dapat menjamin hasil nilai total sentimen dengan akurasi sebesar 92% dan akurasi untuk hasil aplikasi rekomendasi sebesar 83%.

Kata Kunci : Opinion Extraction, Sentiment Classification , Recommender System, Collaborative Filtering

Telkom
University

Abstract

Sentiment Classification can provide value in a text is the text belong to the "negative", "positive" or "neutral". The Classification of Sentiment can provide a recommendation to the user in the form of a textual. Recommender System can be used as a way to give a new product recommendations to users. Most of the application/product descriptions, opinions from users and so on are presented in textual form on the website. There are many ways in assessing a product offered, for example, is to give users an assessment by giving it a "Like" or "Dislike" or the user gives the stars from a scale of 1 (not good) to 5 (very good).

In this final Task, Sentiment Classification will be done through Opinion Extraction, where in this phase will be done word processing product features and opinions, said the process of parsing using the Stanford Parser to get the grammatical relationship in every phrase, word pairs to determine product features and opinions, determines the strength and polarity of opinion through SentiWordNet words then accumulate final value to each review. If the value of the end of the positive Sentiment that user then recommend applications that talk about, if negative otherwise. Then by using the techniques of Item

Based Collaborative Filtering Recommender System, we can provide a recommendation to the user based on the applications they've ever scored before. Based on the results of testing method using Mean Opinion Score (MOS), the Recommender System based on Sentiment Classification through Opinion Extraction can guarantee the result of a total value of sentiment with the accuracy of 92% for accuracy and results of application recommended by 83%.

Keywords : Opinion Extraction, Sentiment Classification, Recommender System, Collaborative Filtering



1. PENDAHULUAN

1.1 Latar Belakang Masalah

Ledakan informasi pada jaringan internet menyebabkan masyarakat mengalami kesulitan mendapatkan informasi yang cepat dan relevan yang sesuai dengan kebutuhannya. Untuk itu diperlukan sistem yang dapat mengakses dan menemukan informasi seperti yang dikehendaki [14]. Dalam proses pencarian informasi terdapat dua tipe pencari. Tipe pertama adalah pencari yang sudah memiliki referensi yang jelas informasi apa yang hendak dicari. Tipe seperti ini cukup dibantu dengan mesin pencari (*search engine*). Sedangkan tipe kedua adalah seseorang yang tidak berbekal referensi, namun hanya memiliki topik tertentu. Pencari tipe kedua ini dapat dipastikan mengalami kesulitan untuk mendapatkan informasi yang dimaksud. Oleh karena itu, dia tidak cukup dibantu hanya dengan *search engine* tetapi juga memerlukan suatu sistem penyedia rekomendasi (*recommender system*) dalam menemukan informasi yang diinginkan [14].

Recommender System merupakan sebuah cara bagaimana mengolah koleksi data yang besar dengan sebuah mekanisme yang dapat mengembalikan atau menghasilkan analisa sehingga didapatkan susunan koleksi yang sesuai dengan kemauan pengguna yang menggunakannya [10], [8].

Penulis menggunakan suatu metode *Sentiment Classification* untuk memberikan nilai positif atau negatif kepada *review* setiap pengguna terhadap suatu aplikasi. *Sentiment Classification* ini juga dilakukan melalui *Opinion Extraction* dimana pembobotan berdasarkan bagaimana hubungan/keterkaitan antara kata fitur produk dengan kata opini pengguna dalam mengulas suatu aplikasi. *Review* yang tergolong ke dalam *Sentiment Classification* positif akan memiliki nilai total sentimen lebih dari 0 (positif), sementara *Review* yang tergolong ke dalam *Sentiment Classification* negatif akan memiliki nilai total sentimen kurang dari 0 (negatif). Kemudian setelah didapat nilai keseluruhan *Sentiment Classification* untuk setiap pengguna, penulis akan membangun suatu *Item Based Collaborative Filtering Recommender System*, dimana setiap nilai tersebut akan dicari nilai kemiripan/similaritasnya berdasarkan *item*/aplikasi,

kemudian akan ditentukan nilai prediksinya dan nilai prediksi dengan nilai tertinggi di setiap pengguna akan menjadi rekomendasi kepada pengguna tersebut (akan diambil 5 besar).

Kelebihan dari sistem ini ialah dapat memiliki kemampuan untuk menyediakan rekomendasi yang tidak terduga atau tidak disengaja, misalnya dapat merekomendasikan aplikasi yang relevan kepada pengguna sekaligus tidak mengandung konten dari profil pengguna tersebut.

Kekurangan atau kelemahan dari sistem ini ialah jika kebanyakan pengguna tidak menggunakan bahasa Inggris yang baik dan benar, akan mengakibatkan sistem tidak menilai *review* tersebut.

1.2 Perumusan Masalah

Permasalahan yang akan dibahas dalam tugas akhir ini antara lain:

1. Bagaimana cara membangun suatu *Recommender System Based on Sentiment Classification* melalui *Opinion Extraction* yang dapat menjamin keterkaitan antara kata fitur produk dengan kata opini dapat menentukan *review* tersebut positif atau negatif?
2. Seberapa besar tingkat akurasi sistem yang telah dibangun tersebut?

1.3 Batasan Masalah

Adapun batasan masalah dalam pengerjaan tugas akhir ini. Diantaranya adalah sebagai berikut:

1. Dataset yang akan dianalisis menggunakan bahasa Inggris yang sesuai dengan ejaan yang benar
2. Data yang akan di-*retrieve* berupa teks
3. Pengujian data dilakukan secara *offline learning* (tidak menggunakan sambungan *internet*)
4. Sampel data berupa *review* pengguna yang diambil hanya dari *Google Play* (<http://play.google.com>), dimana terdapat sampel data 70 aplikasi (20 aplikasi gratis, 20 *games* gratis, 15 aplikasi berbayar dan 15 *games* berbayar). Setiap aplikasi terdiri dari 50 *review* pengguna.

1.4 Tujuan

Tujuan dari tugas akhir ini adalah:

1. Mengimplementasikan *Recommender System Based on Sentiment Classification* melalui *Opinion Extraction*.
2. Mengevaluasi tingkat akurasi dari sistem yang dibangun.

1.5 Metodologi Penyelesaian Masalah

Metodologi yang digunakan untuk memecahkan permasalahan-permasalahan yang ada dalam tugas akhir ini terdiri dari 6 tahap, yaitu:

1. Tahap Studi Literatur
Mempelajari mengenai *Sentiment Classification* melalui *Opinion Extraction* dan mengenal tata Bahasa Inggris terutama dalam kalimat opini.
2. Tahap Pencarian dan Pengumpulan Data
Melakukan pencarian dan pengumpulan data aplikasi beserta *review* pengguna dari *Google Play*. Sedangkan SentiWordNet 3.0 yang berisi kumpulan kata-kata sentimen berbahasa Inggris beserta dengan skornya.
3. Tahap Perancangan Sistem
Melakukan analisis kebutuhan dan merancang sistem dari metode yang digunakan.
4. Tahap Implementasi
Melakukan implementasi berdasarkan dari perancangan sistem yang telah dibuat ke dalam bentuk aplikasi.
5. Tahap Pengujian dan Analisis Hasil Pengujian
Sistem yang telah dibuat kemudian akan diuji pada tahap ini untuk mengukur akurasi hasil klasifikasi serta akurasi performansi keseluruhan sistem dan melakukan analisa terhadap hasil yang diberikan.
6. Tahap Pengambilan Kesimpulan dan Pembuatan Laporan
Pada tahap ini dilakukan pengambilan kesimpulan terhadap analisa yang telah dilakukan dan pembuatan laporan tugas akhir untuk mendokumentasikan tahap-tahap kegiatan.

5. PENUTUP

5.1 Kesimpulan

Berdasarkan hasil pengujian yang dilakukan, dapat disimpulkan beberapa poin, yaitu sebagai berikut.

1. Pasangan kata fitur produk dan kata opini (FO) dan kualitas *review* pengguna berpengaruh besar untuk mendapat nilai total sentimen dalam tiap *review*, karena dengan adanya pasangan FO maka *review* tersebut akan menghasilkan nilai sentimen dari pasangan FO tersebut.
2. Pada pengujian jumlah pasangan FO dengan menggunakan metode MOS, menghasilkan nilai akurasi rata-rata untuk keseluruhan pengujian yang sangat baik, yaitu 4.6 yang terdiri dari aspek relevansi pasangan FO dan hasil akhir nilai total sentimen.
3. Pada pengujian aplikasi yang direkomendasikan, menghasilkan nilai akurasi rata-rata untuk keseluruhan pengujian sebesar 4.15.

5.2 Saran

Memberikan penanganan tambahan untuk kasus jika *review* mendapatkan nilai total sentimen *null*, dengan melakukan penghitungan nilai didasarkan pada aspek-aspek tertentu.

Telkom
University

6. DAFTAR PUSTAKA

- [1] Bo Pang and Lillian Lee. 2008. *Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval*, 2(1-2):1–135.
- [2] Catherine de Marneffe, Marie and Christopher D. Manning. “*Stanford typed dependencies manual*”. United States: *Stanford NLP Group*.
- [3] Feng, S., Ming, Zhang,. Yanxing, Zhang,. Zhihong, Deng,. 2010. *Recommended or Not Recommended? Review Classification through Opinion Extraction*. Beijing: *Peking University*.
- [4] G. Linden, B. Smith, and J. York, “Amazon.com recommendations: Item-to-item collaborative filtering,” *IEEE Internet Computing*, vol. 7, pp. 76–80, January 2003.
- [5] Hu, Minqing, Bing Liu. (2004). *Mining and Summarizing Customer Reviews. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004)*.
- [6] Infogistic – Tag-Set. 2000. <http://www.infogistics.com/tagset.html> (diakses tanggal 10 Juli 2013)
- [7] J. Dietmar, Z. Markus, F. Alexander, F. Gerhard. 2011. “*Recommender System: An Introduction*”. Cambridge: *Cambridge University*.
- [8] L. Candillier, F. Meyer, F. Fessant, and K. Jack, “*State-of-the-art recommender systems*,” 2009.
- [9] NLProcessor: *Text Analysis Toolkit*. 2000. <http://www.infogistics.com/textanalysis.html> (diakses tanggal 13 Februari 2013)
- [10] P. Resnick and H. R. Varian, “*Recommender systems*,” *Commun. ACM*, vol. 40, pp. 56–58, March 1997.
- [11] Sarwar, B., George, Karypis., Joseph, Konstan., John, Riedl., “*ItemBased Collaborative Filtering Recommendation Algorithms*”. Minneapolis: *University of Minnesota*.
- [12] *SentiWordNet: A Publicity Available Lexical Resource for Opinion Mining*. 2010. <http://sentiwordnet.isti.cnr.it/> (diakses tanggal 13 Juli 2013)

- [13] *The Stanford Parser: A Statistical Parser*. 2011.
<http://nlp.stanford.edu/software/lex-parser.shtml#Download> (diakses
tanggal 9 Juli 2013)
- [14] Wiranto. “Mengenal Recommender System”.
<http://wir.staff.uns.ac.id/2010/02/07/mengenal-recommender-system/>
(diakses tanggal 6 Juli 2013)

