

Pendekatan Metode Collaborative Filtering pada Sistem Rekomendasi Pariwisata Kota Bengkulu

Aan Erlansari^{1, a)} and Boko Susilo², Yudi Setiawan³, Iit Pranata⁴⁾

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Bengkulu.

Jl.WR. Supratman Kandang Limun Bengkulu 38371A INDONESIA

E-mail: laan_erlansari@unib.ac.id; 2bokosusilo@unib.ac.id; 3ys.teknik@unib.ac.id; 4iitpranataip@gmail.com

Abstrak. Sistem Rekomendasi objek pariwisata merupakan salah satu contoh bisnis area yang digunakan untuk membantu para wisatawan dalam membuat keputusan bagi perjalanan. Objek wisata yang direkomendasikan merupakan hasil dari rating dan ulasan dari pengguna yang kemudian diinformasikan kepada wisatawan agar lebih nyaman dalam pemilihan objek wisata. Pada artikel ini, sistem rekomendasi menggunakan metode Collaborative Filtering bekerja dengan cara mengumpulkan timbal balik dalam bentuk perankingan (rating) dari suatu objek wisata pada cakupan tertentu dan kemudian mencari kemiripan pada tingkah laku rating beberapa pengguna untuk menentukan rekomendasi terhadap suatu objek wisata. Untuk dapat menghitung nilai ulasan (sentiment analysis) dari pengguna digunakan metode bantuan K-Nearest Neighbor. Hasil yang diperoleh adalah rekomendasi terhadap suatu objek pariwisata yang akan dipilih oleh wisatawan serta perbandingannya dengan objek wisata yang lain.

Kata kunci: Collaborative filtering, rating, sistem rekomendasi

PENDAHULUAN

Pariwisata merupakan keseluruhan rangkaian kegiatan yang berhubungan dengan gerakan manusia yang melakukan perjalanan atau persinggahan sementara dari tempat tinggalnya ke tempat objek wisata (www.bps.go.id, 2017).

Kota Bengkulu yang memiliki potensi lokasi wisata yang beraneka ragam, secara garis besar, daya tarik objek wisata dibagi menjadi 3 kategori: objek buatan manusia, daya tarik sejarah dan daya tarik alam. Dengan banyaknya potensi wisata tersebut, tidak sedikit wisatawan lokal dan mancanegara yang tertarik untuk datang. Kondisi informasi yang kurang dan jumlah objek wisata yang tidak sedikit membuat wisatawan mengalami kesulitan untuk mengetahui lokasi objek wisata dan bagaimana situasi objek wisata tersebut.

Pada artikel ini, akan dibahas sistem rekomendasi yang menimbang dan memberikan rekomendasi berdasarkan rating yang diberikan oleh pengguna dan menganalisis hasil ulasan dari masing-masing ulasan dari pengguna terhadap sebuah lokasi objek wisata.

Adapun metode yang diterapkan dalam penelitian yang telah dilakukan yakni dengan menggunakan menggunakan metode collaborative filtering untuk menilai rating dari objek wisata dan sentimen analisis untuk dapat menghitung ulasan dari para pengguna.

LANDASAN TEORI

a. Collaborative Filtering

Konsep dasar algoritma collaborative filtering adalah untuk menyediakan rekomendasi berdasarkan pendapat pengguna lain yang memiliki behaviour yang mirip (Rizki DK, 2015). Pendapat tersebut bisa didapatkan secara eksplisit dari pengguna lain atau dengan melakukan pengukuran secara implisit.

Tujuan algoritma collaborative filtering adalah untuk menyarankan item baru atau untuk memprediksi item tertentu untuk seorang pengguna berdasarkan preferensi pengguna sebelumnya dan opini dari pengguna lain yang mirip. Dalam scenario collaborative filtering.

$$U = \{u_1, u_2, \dots, u_m\} \quad (1)$$

$$I = \{i_1, i_2, \dots, i_n\} \quad (2)$$

Keterangan :

U_m = User/Pengguna

I_m = Item

Setiap pengguna U_i memiliki sekumpulan item I_{ui} , yang merupakan opini dari pengguna. Opini tersebut bisa didapatkan secara eksplisit dari pengguna berdasarkan rating score yang diberikan, atau bisa juga didapatkan secara implisit dari riwayat pilihan, atau dari analisa timing logs, dan lain-lain. Ada pengguna yang disebut *active user*, yang merupakan landasan bagi algoritma collaborative filtering untuk mencari kesamaan item yang bisa saja menjadi dua bentuk.

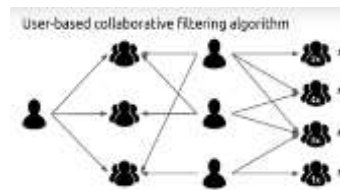
1) Prediksi

Prediksi merupakan nilai numerik, P_{aj} , yang merepresentasikan prediksi kesukaan dari $i_j \in I_{ua}$ untuk active user U_a . Nilai prediksi ini nilainya berupa skala yang sama (contoh dari skala 1 sampai 5) sebagai nilai opini yang diperoleh dari U_a .

2) Rekomendasi

Rekomendasi merupakan daftar dari N item, $I_r \subset I$, yang paling disukai oleh pengguna aktif (active user).

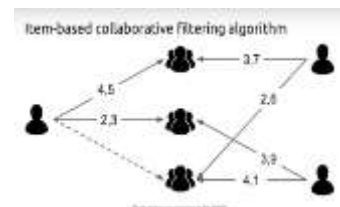
Algoritma collaborative filtering dibagi menjadi dua kategori utama, yaitu *user-based collaborative filtering*, dan *item-based collaborative filtering*.



Gambar 1. Algoritma User-Based Collaborative Filtering

Algoritma *user-based collaborative filtering* menggunakan seluruh database dari user-item untuk menghasilkan prediksi. Sistem tersebut menggunakan teknik statistik untuk mencari sekumpulan pengguna, yang disebut sebagai *neighbors*, yang memiliki riwayat pilihan yang mirip dengan target pengguna (contohnya, mereka pernah memilih produk yang sama). Teknik ini juga dikenal dengan sebutan *nearest-neighbor*.

Untuk mencari neighbors tersebut banyak cara yang ditawarkan, ada yang memanfaatkan teknik K-nearest neighbor (KNN) ada juga yang mencarinya secara manual dengan mencari irisan dari semua daftar item yang sudah di rating oleh pengguna aktif dengan semua daftar item yang sudah di rating oleh pengguna lain satu-persatu (Shinde, S. K. & A., 2011).



Gambar 2. Algoritma Item-Based Collaborative Filtering

Algoritma *item-based collaborative filtering* menyediakan rekomendasi item dengan lebih dulu membuat model dari *rating* pengguna. Algoritma pada kategori ini menggunakan pendekatan probabilitas dan mengibaratkan proses *collaborative filtering* sebagai proses menghitung nilai dugaan dari prediksi seorang pengguna. Pada algoritma *item-based* akan menghitung kesamaan di antara *item*, dilihat dari *rating* yang diberikan pengguna untuk item tersebut.

Untuk membuat sistem rekomendasi menggunakan metode *item-based* ada dua langkah yang harus dilakukan, yaitu seperti berikut (Hafizh, 2013).

a) Menghitung *Similarity*

Menghitung kemiripan antara satu item dengan item lainnya. Rumus yang biasa dipakai adalah *Cosine Similarity Measures*, yang menghitung kemiripan antara dua *item* dari sudut kosinus yang tersebut identik. Sedangkan jika nilainya 0, maka dua item tersebut sama selai tidak mempunyai kemiripan.

$$sim(i, j) = \frac{\sum_{u \in U} (r_{u,i} - r_{u,j})}{\sqrt{\sum_{u \in U} (r_{u,i} - r_{u,j})^2} \sqrt{\sum_{u \in U} (r_{u,j} - r_{u,i})^2}} \quad (3)$$

Jika kemiripan sudah dihitung, maka kita akan mengurutkan *item-item* berdasarkan *similarity* mereka, item-item yang mempunyai *similarity* besar akan berada pada urutan atas, dan sebaliknya.

b) Menghitung Prediksi *Rating*

Menghitung prediksi *rating* dari item-item tersebut dengan membandingkan rating yang pernah diberikan pengguna pada suatu item dengan kemiripan antara item tersebut dengan item lainnya. Metode yang digunakan merupakan metode *Weighted Sum*.

$$P(u, j) = \frac{\sum_{i \in I} (R_{u,i} * S_{i,j})}{\sum_{i \in I} |S_{i,j}|} \quad (4)$$

Jika prediksi *rating* sudah dihitung, maka rekomendasi item pun dapat dihasilkan.

b. *Sentiment analysis*

Sentiment analysis atau opinion mining mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi linguistik dan *text mining* yang bertujuan meng pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenaan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu (Liu, 2012).

Tugas dasar dalam analisis sentimen adalah mengelompokkan teks yang ada dalam sebuah kalimat atau dokumen kemudian menentukan pendapat yang dikemukakan dalam kalimat atau dokumen tersebut apakah bersifat *positif*, *negatif* atau netral. *Sentiment analysis* juga dapat menyatakan perasaan emosional sedih, gembira, atau marah.

Kita dapat mencari pendapat tentang produk-produk, tempat makan atau orang-orang dan menentukan apakah mereka dilihat *positif* atau *negatif* di web. Hal ini memungkinkan kita untuk mencari informasi tentang (Manalu, 2014):

- Deteksi *Flame (Rants Buruk)*
- Persepsi Produk Baru.
- Persepsi Merek.

d) Manajemen Reputasi.

Ekspresi atau sentiment mengacu pada fokus topik tertentu, pernyataan pada satu topik mungkin akan berbeda makna dengan pernyataan yang sama pada subject yang berbeda. Oleh karena itu pada beberapa penelitian, terutama pada review suatu tempat makan, pekerjaan didahului dengan menentukan elemen dari sebuah produk atau makanan yang sedang dibicarakan sebelum memulai proses *opinion mining*.

Berikut ini adalah langkah-langkah analisis sentimen pada dokumen :

1) *Crawling* dan *Input Data*

Pada tahap ini dilakukan pembacaan terhadap korpus, korpus adalah database besar yang menyimpan *text* yang akan dianalisis. Korpus didapat dengan cara mengumpulkan data dengan meng-*crawling*.

2) *Pre-processing*

Pada tahapan ini akan dilakukan, yaitu:

- Cleaning*, yaitu membersihkan dokumen dari kata yang tidak diperlukan sesuai kamus data yang telah ditentukan. Kata yang dihilangkan adalah HTML, simbol, ikon emosi, email, URL. Agar mengurangi *noise* saat menganalisis sentimen.
- Case Folding*, yaitu pengubahan bentuk huruf menjadi huruf kecil dan penghapusan tanda baca serta angka, sesuai dengan kamus data yang telah ditentukan.

3) Ekstraksi Fitur

Proses ekstraksi fitur bertujuan untuk meng-ekstrak kata-kata kunci dari korpus yang telah dibaca. Kata-kata kunci tersebut disebut dengan fitur atau term. Fitur inilah yang nantinya akan di proses dalam tahap analisis. Dalam proses ekstraksi fitur terdapat subproses, yaitu:

- Tokenization* adalah proses memecah text menjadi kata tunggal. Pada penelitian ini fitur yang digunakan dalam memecah text adalah unigram yaitu token yang terdiri hanya satu kata.
- Stopwords Removing* adalah proses menghilangkan kata tidak penting dalam text. Hal ini dilakukan untuk memperbesar akurasi dari pembobotan term. Untuk proses ini, diperlukan suatu kamus kata-kata yang bisa dihilangkan. Dalam Bahasa Indonesia, misalnya kata: dan, atau, mungkin, ini, itu, dll adalah kata-kata yang dapat dihilangkan.
- Stemming* adalah proses pemetaan variasi morfologikal kata dalam kata dasar atau kata umumnya (*stem*). Misalnya kata "perancangan" dan "merancang" akan diubah menjadi sebuah kata yang sama, yaitu "rancang". Proses *stemming* sangat tergantung kepada bahasa dari kata yang akan di-*stem*. Hal ini dikarenakan, dalam melakukan proses *stemming* harus mengaplikasikan aturan morfologikal dari suatu bahasa.

4) Pembobotan Term

Pembobotan term memberikan sebuah nilai untuk sebuah term berdasarkan tingkat kepentingan tersebut di dalam sekumpulan dokumen masukan. Pada penelitian ini akan digunakan metode TF-IDF sebagai proses pembobotan, yaitu dengan cara mencari representasi nilai dari tiap-tiap dokumen dari sekumpulan data training, dan akan dibentuk menjadi sebuah vektor. Dapat dirumuskan sebagai berikut : (D Manning et al, 2009).

$$w(t, d) = tf(t, d) * idf \quad (5)$$

$$idf = \log\left(\frac{N}{df}\right) \quad (6)$$

c. *K-Nearest Neighbor*

Algoritma K-Nearest Neighbor (K-NN) adalah sebuah metode melakukan klasifikasi terhadap objek berdasarkan data pembelajaran jaraknya paling dekat dengan objek tersebut. K-NN termasuk algoritma supervised learning dimana hasil dari query instance yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada K-NN, kemudian kelas yang paling banyak muncul yang akan menjadi kelas hasil klasifikasi.

Tujuan dari algoritma ini adalah mengklasifikasikan obyek berdasarkan atribut dan training sample. Classifier tidak menggunakan apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan titik query, akan ditemukan sejumlah k obyek atau (titik training) yang paling dekat dengan titik query. Klasifikasi menggunakan voting terbanyak diantara klasifikasi dari k obyek.

Algoritma K-Nearest Neighbor (K-NN) menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari query instance yang baru. Algoritma metode K-Nearest Neighbor (K-NN) sangat sederhana, bekerja berdasarkan jarak terpendek dari query instance ke training sampel untuk menentukan K-NN. Training sample diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi training sample. Sebuah titik pada ruang ini ditandai jika merupakan klasifikasi yang paling banyak ditemukan pada k buah tetangga terdekat dari titik tersebut (Guo et al, 2001)

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (7)$$

dimana matriks D(a,b) adalah jarak skalar dari kedua vektor a dan b dari matriks dengan ukuran d dimensi.

Satuan jarak yang digunakan Euclidian, jenis dari metode ini, jika dilihat dari nilai N nya ada dua macam yaitu (Darujati, 2010):

1. 1-NN

Pengklasifikasian dilakukan pada 1 label data terdekat, algoritmanya sebagai berikut :

- Menghitung jarak antara data baru ke setiap pelabelan data.
- Menentukan 1 pelabelan data yang mempunyai jarak paling minimal.
- Klasifikasi data baru ke dalam pelabelan data tersebut

2. k-NN

Pengklasifikasian dilakukan dengan menentukan nilai pada k label data terdekat, dengan syarat nilai $k > 1$, algoritmanya sebagai berikut :

- Menghitung jarak antara data baru ke setiap pelabelan data.
- Menentukan k pelabelan data yang mempunyai jarak paling minimal.
- Klasifikasi data baru ke dalam pelabelan data yang mayoritas.

Cosine similarity adalah penentuan kesesuaian dokumen dengan query untuk pengukuran antara vektor dokumen (D) dengan vektor query (Q). Persamaan yang digunakan (Darujati, 2010):

$$sim(Q, D) = \cos(Q, D) = \frac{Q \cdot D}{|Q||D|} \quad (2.8)$$

$$|D| = \sqrt{(\sum_{i=1}^n D_i^2)}$$

$$|Q| = \sqrt{(\sum_{i=1}^n Q_i^2)}$$

dimana Q adalah dokumen uji, D dokumen training, dan adalah nilai bobot yang telah diberikan pada setiap term pada dokumen.

Pada fase training, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi data training sample. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk testing data (yang klasifikasinya tidak diketahui). Jarak dari vektor baru yang ini terhadap seluruh vektor training sample dihitung dan sejumlah k buah yang paling dekat diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak dari titik-titik tersebut. Sebagai contoh, untuk mengestimasi $p(x)$ dari n training sample dapat memusatkan pada sebuah sel disekitar x dan membiarkannya tumbuh hingga meliputi k samples. Samples tersebut adalah K-NN dari x. Jika densitasnya tinggi di dekat x, maka sel akan berukuran relatif kecil yang berarti memiliki resolusi yang baik. Jika densitas rendah, sel akan tumbuh lebih besar, tetapi akan berhenti setelah memasuki wilayah yang memiliki densitas tinggi.

Nilai k yang terbaik untuk algoritma ini tergantung pada data. Secara umum, nilai k yang tinggi akan mengurangi efek noise pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi semakin kabur. Nilai k yang bagus dapat dipilih dengan optimasi parameter, misalnya dengan menggunakan cross-validation. Kasus khusus dimana klasifikasi diprediksikan berdasarkan training data yang paling dekat (dengan kata lain, $k = 1$) disebut algoritma nearest neighbor. Ketepatan algoritma K-NN sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi.

Berikut ini adalah langkah-langkah menghitung K-Nearest Neighbors pada dokumen :

1. Tentukan parameter k sebagai jumlah tetangga terdekat, dalam sistem ini digunakan $k=1$, sehingga jika ada tetangga terdekat, itu yang akan digunakan sebagai nilai prediksi.
2. Hitung jarak antara data yang masuk dan semua sampel latih yang sudah ada. Pada penelitian ini jenis jarak terdekat yang digunakan yaitu cosine similarity pada persamaan (2.8).
3. Mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak terkecil.
4. Mengumpulkan kategori Y (klasifikasi nearest neighbor).
5. Dengan menggunakan kategori mayoritas, maka dapat diprediksikan nilai query Insurance yang telah dihitung, kemudian tentukan jarak tetangga terdekat yang akan digunakan sebagai nilai prediksi dari data berikutnya. (Zee-Jing & Hsien-Wu, 2004)

Ketepatan algoritma k-NN ini sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi. Riset terhadap algoritma ini sebagian besar membahas bagaimana memilih dan memberi bobot terhadap fitur, agar performa klasifikasi menjadi lebih baik.

IMPLEMENTASI

Proses collaborative filtering ditanamkan pada salah satu laman dari sistem rekomendasi ini dan disimpan dalam basisdata guna mendapatkan hasil yang otomatis, kemudian hasil tersebut akan dipanggil dalam bentuk tabel matrix. Untuk dapat dilakukan perhitungan, dibutuhkan beberapa data sampel 5 objek wisata dan pengguna yang memberikan rating di objek wisata, seperti terlihat pada tabel 1:

Tabel 1. Rating pengguna

No	Objek wisata	Pengguna 1	Pengguna 2	Pengguna 3	Pengguna 4	Pengguna 5
1	Benteng marlborouh	5	4		5	4
2	Pantai panjang			4	4	4
3	Tabut		3		3	4
4	Tugu pers	4	2		4	
5	View tower			5	4	4

Algoritma collaborative filtering selanjutnya akan melakukan perhitungan tingkat kemiripan setiap objek wisata dengan menggunakan metode adjusted cosine similarity. Input data yang digunakan adalah rating dari pengguna. Output dari perhitungan tersebut berupa nilai kemiripan dari objek wisata pertama dan objek wisata kedua yang dibandingkan. Sehingga dihasilkan:

Tabel 2. Hasil perhitungan kemiripan

Pengguna	Objek 1	Objek 2	Kemiripan
Pengguna 1	Benteng	Pantai	-1,25
Pengguna 2	marlborough	panjang	
Pengguna 1	Benteng	Pantai	0,75
Pengguna 4	marlborough	panjang	
...

Setelah dilakukan proses kemiripan, metode weight sum dilakukan untuk mencari kemiripan dari objek wisata dan rating dari objek wisata tersebut. Output dari perhitungan tersebut akan menghasilkan prediksi nilai rating dari seorang pengguna terhadap objek wisata tertentu dengan menggunakan persamaan (2.4), dimana nilai prediksi tertinggi akan berada di posisi pertama, seperti yang didapat pada tabel.

Tabel 3 Prediksi rating

Pengguna	Objek	Prediksi
Pengguna 4	Pantai panjang	4
Pengguna 4	View tower	4
Pengguna 4	Tabut	3
Pengguna 2	Tabut	3

Tahapan selanjutnya adalah melakukan proses sentimen analisis untuk mengetahui nilai dari ulasan dari pengguna. Tahapan awal dari dalam sentiment analisis akan yang digunakan pada K-Nearest neighbors adalah text preprocessing, terdapat tiga tahapan preprocessing: tokenizing, filtering, dan stemming. Proses tokenizing merupakan proses pemotongan string input berdasarkan tiap kata yang menyusunnya. Selain itu, spasi digunakan untuk memisahkan antar kata tersebut, seperti ditunjukkan pada tabel 4.

Tabel 4. Proses sentiment analisis sebelum dan sesudah

No	Dokumen	Hasil text preprocessing
1	Tempatnya mantab, tetapi terlalu banyak pasir, saat tertiuip angin masuk kemata dan menyebabkan bnyak mata turis iritasi, shingga mrka jarang datang ksini	tempat mantab pasir tiup angin masuk mata sebab bnyak mata turis iritasi shingga mrka jarang ksini
2	Spotnya memang bagus. Tetapi pasir nya yg gak karuan dibawa angin bikin mata perih dan gak bisa nikmati spot	spotnya bagus pasir nya yg gak karuan bawa angin bikin mata perih gak nikmat spot eskrim
3	Sekarang Bengkulu sudah semakin menonjolkan salah satu asetnya. Semoga kita bisa menjaganya dengan lebih baik	bengkulu tonjol salah aset moga jaga
4	Pantainya bagus, dilengkapi dengan spot-spot foto yang bagus, akan tetapi kebersihan masih sangat kurang terjaga di lingkungan ini	pantai bagus lengkap spot foto bagus bersih jaga lingkungan
5	sebenarnya pantai panjangindah, dan berpotensi, tapi sayangnya pantai ini tidak terawat kebersihannya.	pantai bagus jgn kotor sampah jaga bersih

Dari tabel tersebut, terdapat proses tokenizing dengan menghilangkan karakter selain 'a' sampai dengan 'z' yang dianggap sebagai delimitter dan mengubah kalimat tersebut menjadi potongan string / kata penyusunnya. Selanjutnya proses filtering dijalankan untuk mengeliminasi kata-kata yang kurang penting seperti kata 'untuk' pada ulasan nomor 1, kata 'disini', 'selain', 'itu', 'agak', pada ulasan lainnya. Dan dilanjutkan dengan proses stemming guna mentransformasi kata-kata yang terdapat dalam suatu dokumen ke kata-kata akarnya (dasar) dengan menggunakan aturan-aturan tertentu, seperti menghilangkan imbuhan pada kata 'tempatny' menjadi 'tempat', 'menyebabkan' menjadi 'sebab' dan lain sebagainya.

KESIMPULAN

Proses perhitungan rekomendasi objek pariwisata menggunakan metode collaborative filtering berhasil dilaksanakan dan diimplementasikan pada objek wisata Kota Bengkulu. Berdasarkan hasil perhitungan rating, setiap pengguna dapat melihat rekomendasi objek wisata yang terbaik dan layak untuk dikunjungi terlebih dahulu. Begitu pula dengan melihat hasil ulasan menggunakan sentimen analisis, sistem rekomendasi memberikan hasil yang terbaik dari perhitungan sentiment analisis menggunakan k-nearest neighbor.



Kedua metode ini akan mendapatkan hasil yang lebih baik, jika objek wisata yang dijadikan sebagai variabel diperbanyak sebanyak mungkin dan setiap objek wisata mendapatkan rating-rating dari setiap pengguna. Begitu juga dengan prinsip sentiment analisis, makin banyak ulasan yang diolah maka sistem akan merekomendasikan objek wisata terbaik.

DAFTAR PUSTAKA

- Badan Pusat Statistik. (2012, Februari) BPS. [Online]. Available: <http://www.bps.go.id/index.php/istilah/357>.
- Collaborative Filtering: Creating the Best Teams Ever. (2017, Februari) Maurits van der Goes [Online]. Available: <https://neo4j.com/blog/collaborative-filtering-creating-teams/>
- e-SCHEDULING INFORMATION SYSTEMS MAINTENANCE OF IT ASSETS (e-SISMITAS) DI PT. KARYA PUTRA SUKSES (2015, Juni) Abdul Rohman. [Online]. Available: <https://widuri.raharja.info/index.php/SII114469852>
- Fakhrurezi, 2016, Perancangan Aplikasi Virtual Tour Universitas Bengkulu Dengan Pemodelan Objek 360.
- Ian, H, Witten. Frank, Eibe & Mark A, Hall. 2011. Data mining: Practical Machine Learning Tools and Techniques 3rd Edition. Elsevier
- BRIAN, B. Sentiment Analysis And Opinion Mining. Chicago: Morgan & Claypool Publishers, May 2012
- Manning, D, Christopher. Raghavan, Prabhakar & Schütze, Hinrich. An Introduction do Information Retrieval. Cambridge: Cambridge University Press. 2009
- Rizki DK. Sistem Rekomendasi Komunitas Pemuda Di Kota Semarang Berbasis Item Based Collaborative Filtering Dengan Metode Adjusted Cosine Similarity., Universitas Dian Nuswantoro. 2015.
- Y. Setiawan, B.Susilo, D. Puspitaningrum and A. Erlansari, "Design Dan Implementasi Sistem Informasi Pariwisata Berbasis Konten sebagai Startup Lokal Bengkulu,"in Seminar Nasional Teknologi Informasi (SNTI), Jakarta, 2017