

PENERAPAN METODE ASSOCIATION RULE MINING UNTUK ASOSIASI ULASAN TERHADAP ASPEK TEMPAT WISATA JAWA TIMUR PARK 3

Aisyatul Maulidah^{*1}, Fitra A. Bachtiar²

^{1,2}Sistem Informasi, Fakultas Ilmu Komputer, Universitas Brawijaya Malang
Email: ¹ais.isydah@gmail.com, ²fitra.bachtiar@ub.ac.id

*Penulis Korespondensi

(Naskah masuk: 27 November 2020, diterima untuk diterbitkan: 19 Oktober 2021)

Abstrak

Google Review pada salah satu fitur Google Maps dapat menjadi salah satu media untuk mengukur tingkat kepuasan pengunjung Jawa Timur Park 3 (Jatim Park 3). Akan tetapi jumlah ulasan yang mencapai ribuan dan belum tersedianya media pengelola data ulasan dapat mempersulit manajemen Jatim Park 3 dalam mengeksplorasi dan menganalisis masukan pengunjung secara mendetail. Penelitian ini memanfaatkan teknik *Association Rule Mining* (ARM) dalam mengelola data ulasan sehingga dapat menemukan hubungan kata yang sering muncul pada ulasan. Teknik ini paling populer untuk menemukan hubungan tersembunyi antar variabel. Algoritma yang digunakan dalam mengimplementasikannya adalah algoritma Apriori karena dianggap paling efisien. Pada penelitian ini menggunakan data ulasan sebanyak 1067 ulasan dalam Bahasa Indonesia dari bulan Januari sampai bulan April tahun 2019. Berdasarkan wawancara, data tersebut digolongkan menjadi 8 aspek berdasarkan kata kunci yang sudah ditentukan sebelumnya. Aspek tersebut antara lain akses jalan, biaya, kebersihan, kepuasan, keramaian, pelayanan, keamanan, dan teknologi. Pengujian dilakukan untuk mengetahui pengaruh *minimum support* dan *minimum confidence* terhadap *rule* yang terbentuk. Keseluruhan aspek mampu menghasilkan asosiasi kata dengan algoritma Apriori. Selain itu, Keseluruhan *rule* yang terbentuk menghasilkan rata-rata *lift ratio* di atas 1 dimana *rule* dengan nilai *lift ratio* diatas 1 tersebut merupakan *rule* yang unik diantara *rule*-*rule* lain yang terbentuk dari asosiasi tersebut. Pada penelitian ini, *rule* yang terbentuk divisualisasikan untuk menampilkan keterkaitan antara kata kunci dengan aspek pada data ulasan pengunjung Jatim Park 3. Penelitian ini mencoba menggali informasi mengenai pemetaan layanan mana saja yang mendapatkan perhatian pengunjung di Jatim Park 3.

Kata kunci: *Aspek, Text Preprocessing, Association Rule Mining, Visualisasi Data*

APPLICATION OF ASSOCIATION RULE MINING METHOD FOR ASSOCIATION OF REVIEWS ON ASPECT IN TOURIST ATTRACTION JAWA TIMUR PARK 3

Abstract

Google Review, which is one of the features of Google Maps can be a medium to measure the satisfaction rate visitors of Jawa Timur Park 3 (Jatim Park 3). the number of reviews that reached thousands and media of review data manager is not available yet complicate the management of Jatim Park to explore and analyze visitor feedback in detail. The Association Rule Mining (ARM) technique is a text mining method that can support the knowledge discovery process in large document collections. ARM is able to link keywords to comments to find words that appear frequently. This technique is most popular for finding hidden relationships between variables. The algorithm used in this study is apriori algorithm because it is the most efficient. In this study, there are 1067 reviews of the visitors in Indonesian from January to April 2019 as the data. The data is classified into 8 aspects based on predetermined keywords. These aspects include road access, cost, cleanliness, satisfaction, hustle, service, security, and technology. Testing was conducted to determine the minimum support and minimum confidence impact of the established rules. The whole aspects is capable of generating word associations with an Apriori algorithm. In addition, the overall rules that are formed produce an average lift ratio above 1 where the rule with that value is a unique rule among other rules formed from the association. In this study, the rules that are formed are visualized to show the relationship between keywords and aspects of visitor reviews of Jatim Park 3. This research tries to dig up information about mapping which services get the attention of visitors in Jatim Park 3.

Keywords: *Aspects, Text Preprocessing, Association Rule Mining, Data Visualization*

1. PENDAHULUAN

Pihak manajemen Jawa Timur Park 3 (Jatim Park 3) telah berfokus pada peningkatan kualitas pelayanan untuk meningkatkan jumlah pengunjung. Berdasarkan hasil wawancara dengan pihak manajemen pada Jatim Park 3, salah satu upaya yang dilakukan oleh pihak manajemen dalam mengukur kepuasan pengunjung adalah dengan menyediakan beberapa media penyampaian keluhan dan saran seperti brosur, *call center*, email, maupun *web customer service*. Namun pengunjung saat ini lebih memilih media sosial seperti instagram, Twitter, Facebook, TripAdvisor, dan Google Review.

Fenomena tersebut belum diimbangi dengan tersedianya media pengolahan ulasan untuk menghasilkan sebuah informasi baru. Banyaknya ulasan yang masuk dan belum terkategori dapat mempersulit dalam mengeksplorasi atau menganalisis masukan pengunjung secara mendetail. Sedangkan membaca ulasan maupun komentar satu persatu tidaklah efektif untuk mendapatkan informasi yang memadai. Sebagai langkah awal, pihak marketing ingin mengetahui kesan yang disampaikan oleh para pengunjung melalui media sosial.

Dari permasalahan yang dijelaskan pada Jatim Park 3 maka perlu untuk melakukan pengolahan terhadap masukan yang berupa ulasan maupun komentar melalui *text mining*. *Text mining* dapat mendukung proses *knowledge discovery* pada koleksi dokumen yang berukuran besar dan membantu memecahkan masalah *information overload* dengan menggunakan teknik-teknik dari bidang ilmu yang terkait. Salah satu teknik *text mining* yang digunakan dalam penelitian ini adalah teknik *Association Rule Mining* (ARM). ARM dapat memecahkan sejumlah masalah, khususnya mencari hubungan antar kata kunci pada ulasan. Teknik ARM juga dapat dimanfaatkan untuk menemukan hubungan tersembunyi antara variabel dalam *database* dari konsep *data mining* (Wu dkk., 2019)

Pada penelitian sebelumnya yang dilakukan oleh Hong dkk. (2020) memanfaatkan metode ARM untuk mengidentifikasi hubungan tersembunyi antara faktor-faktor penyebab terjadinya kecelakaan yang melibatkan kendaraan HAZMAT di jalan tol. Isanta dkk. (2016) juga telah memanfaatkan proses ARM dalam proses asosiasi target dan kata kunci *tweet* untuk menghasilkan visualisasi tren produk brand menggunakan ekstraksi kata kunci dari metadada *twitter*. Rekik dkk. (2018) juga memanfaatkan ARM untuk menemukan aturan asosiasi antara kriteria dan kategori situs web guna mempelajari interdependensi dalam kriteria yang muncul dari kumpulan literatur yang diperoleh dari publikasi ilmiah. Sedangkan Cozza dkk. (2018) menerapkan teknik *class Association Rule* untuk mendapatkan hubungan di antara properti hotel, karakteristik pemberi ulasan, dan skor ulasan.

Berdasarkan uraian sebelumnya, dalam penelitian ini juga akan memanfaatkan teknik ARM dalam pengolahan data ulasan pengunjung pada Jatim Park 3. Informasi yang akan digali berupa asosiasi kata kunci dengan layanan yang ada pada Jatim Park 3. Hasil asosiasi akan ditampilkan dalam bentuk visualisasi yang diharapkan dapat mempermudah bagian pengelola dalam memahami setiap *feedback* pada ulasan. Pada penelitian ini juga mencoba menggali informasi mengenai pemetaan layanan mana saja yang mendapatkan perhatian pengunjung di Jatim Park 3.

2. METODE PENELITIAN

Teknik *Association Rule mining* (ARM) merupakan pendekatan *machine learning* non-parametrik yang paling populer untuk menemukan hubungan tersembunyi antara variabel dalam *database* dari konsep *data mining* (Wu dkk., 2019). ARM juga digunakan untuk menemukan aturan asosiatif antara suatu kombinasi item atau hubungan suatu data yang digunakan untuk memprediksi suatu pola (Olson dan Shi, 2008). Han dkk. (2012) mengungkapkan bahwa *Association Rule* terdiri dari *frequent itemset* yang sering ditemukan, dimana *Association Rule* yang kuat dalam bentuk $A \rightarrow B$ dihasilkan. Aturan ini juga memenuhi *minimum confidence threshold*. Asosiasi dapat dianalisis lebih lanjut untuk mengungkap aturan korelasi statistik antara *itemset* A dan B. ARM dianggap sebagai metode *data mining* yang tidak bergantung pada hipotesis apa pun dan dapat menemukan koneksi bermakna yang tersembunyi dalam kumpulan data besar (Yu dkk., (2019). Berdasarkan beberapa keunggulan tersebut, teknik ARM dimanfaatkan pada penelitian ini untuk menemukan hubungan setiap kata pada ulasan.

Agar mendapatkan hasil penelitian yang sesuai, maka diperlukan beberapa tahapan penelitian yang ditunjukkan pada Gambar 1.



Gambar 1. Alur Metodologi Penelitian

Tahap pertama pengumpulan data melalui penggalian informasi dari Jatim Park 3 dengan

teknik wawancara terhadap bagian marketing dan HRD Jatim Park Group. Data selanjutnya yang dibutuhkan merupakan data ulasan pada Google Review yang diperoleh melalui proses ekstraksi data dengan menggunakan teknik *web scraping*. *Web Scraping* dilakukan dengan memanfaatkan *tools* yang bernama WebHarvy. WebHarvy dapat dimanfaatkan untuk mendapatkan berbagai macam data pada suatu situs web. Data yang diperoleh kemudian disimpan dalam format file *Comma Separated Values* (CSV). Periode data ulasan yang digunakan yakni pada bulan Januari sampai bulan April tahun 2019 dimana data yang diperoleh sebanyak 1067 ulasan.

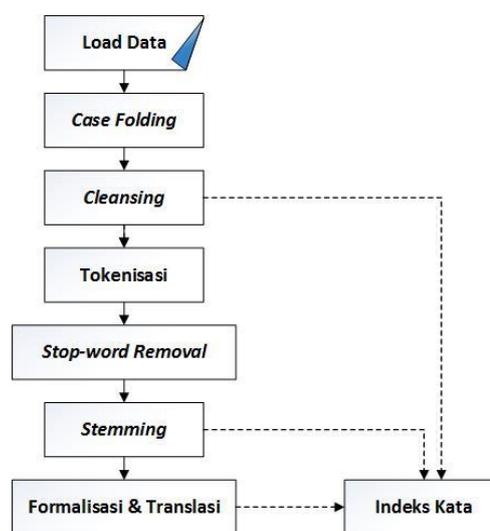
Pada penelitian ini, kata kunci aspek digunakan untuk mengasosiasikan layanan yang dimiliki Jatim Park 3 dengan data ulasan pengunjung. Proses penentuan kata kunci aspek berdasarkan hasil observasi serta persetujuan dari pihak Jatim Park 3 terhadap data ulasan yang telah diperoleh sebelumnya. Terdapat 8 aspek yang dihasilkan, yaitu akses jalan, biaya, kebersihan, kepuasan, keramaian, pelayanan, keamanan, dan teknologi. Keseluruhan aspek tersebut menjadi fokus layanan yang ingin dapat didalami oleh manajemen Jatim Park 3.

Tabel 1 Contoh Pelabelan Data untuk Setiap Aspek

Aspek	Kata Kunci	Contoh
Akses jalan	Strategis, lokasi, rute, jalan.	Letaknya berada di sebelah jalan raya sehingga memudahkan akses menuju ke lokasi wisata ini.
Biaya	Harga, biaya, mahal, murah, terjangkau, gratis.	Tiketnya mahal tapi sangat seru dan lengkap. enggak cukup dijelajahi kalau 1 hari saja.
Kebersihan	Bersih, kotor, bau, sampah.	Cocok banget untuk liburan bersama keluarga, fasilitas: parkir luas, toilet bersih, musholla juga bersih. Suasana disini juga nyaman.
Kepuasan	Puas, kecewa, kurang, menyesal, rekomendasi.	Tempat edukasi keluarga yang rekomen banget Dengan suasana yang nyaman Dan tempat yang asyik untuk berpetualang Sekaligus mengeksplorasi dunia luar tanpa ribet Ehehehehe
Keramaian	Sepi, padat, ramai, antri, macet.	Tempat yang sangat menyenangkan. atraksi dino park yang sangat penuh antri panjang hanya untuk atraksi dinosaurus lima zaman yang hanya berlangsung 5menit dan tidak beroperasi ketika hujan.
Pelayanan	Layanan, Petugas, staff, pegawai, karyawan.	Tempatnya dan arena yang ada keren, parkirnya luas, petugasnya sangat ramah dan dengan senang hati menawarkan untuk di fotokan, toilet nyapun bersih.. mantap jiwa deh..
Teknologi	Teknologi, canggih, digital.	Ke sana main ke funtech untuk bermain dengan teknologi. Di sini kita bakal main game seru yang semuanya mengandalkan teknologi mulai dari VR dan masih banyak lagi.
Keamanan	Aman, bahaya, selamat.	Wahananya pas dan aman untuk anak , tidak terlalu menantang dan menguji adrenalin.

Kemudian *dataset* dilabelkan ke beberapa aspek yang sudah ditentukan. Proses pelabelan dilakukan secara manual pada Excel dengan menyesuaikan kata kunci pada setiap aspek yang dijelaskan pada Tabel 1.

Setelah itu, dilakukan *text pre-processing* untuk mendapatkan data yang terbebas dari *noise* serta menyeragamkan bentuk kata dan mengurangi volume kata pada data ulasan. Feldman dan Sanger (2007) menyatakan bahwa tahap *text preprocessing* bertanggung jawab dalam mentransformasikan data tidak terstruktur yang tersimpan dalam kumpulan dokumen ke format intermediate yang lebih eksplisit. Tahapan ini terdiri dari beberapa proses seperti yang ditunjukkan pada Gambar 2.

Gambar 2. Proses *Text Pre-processing*

Case folding dilakukan untuk mengubah semua karakter huruf menjadi huruf kecil, *cleansing* untuk memisah dan menghilangkan tanda baca. Tokenisasi untuk memisahkan kata pada kalimat menjadi sebuah string. *Stopword removal* untuk menghilangkan kata-kata yang dianggap tidak penting yang masuk ke dalam *stop-list*. *Stemming* untuk mengubah kata berimbuhan menjadi kata dasar. Sedangkan Formalisasi dan translasi untuk mengubah kata slang dan bahasa asing ke bentuk baku sesuai Kamus Besar Bahasa Indonesia (KBBI).

Tahap selanjutnya adalah proses asosiasi. Asosiasi dengan menerapkan metode ARM dan memanfaatkan algoritma Apriori. Metode generasi yang kuat dengan teknik pemangkasan baru yang digunakan membuat algoritma Apriori lebih efisien dibandingkan dengan algoritma lainnya. Algoritma Apriori juga memiliki kekuatan pada kemampuannya untuk menghindari pemborosan pada proses penghitungan kandidat *itemset* yang jarang (Hong dkk., 2020)

Dalam proses ini akan dilakukan proses pemindaian data hasil *preprocessing* untuk menemukan set *item* yang sering muncul pada setiap aspek. Selain itu, nilai *support* dan *confidence*

digunakan untuk mencari keterkaitan antar kata. *Support* merupakan probabilitas kemunculan X dan Y secara bersama-sama dalam kumpulan *dataset* (Lee dkk. 2019). Sedangkan *confidence* menunjukkan probabilitas kemunculan item Y karena item X muncul (Hong dkk., 2020). Nilai *support* diperoleh dari persamaan (1) dan nilai *confidence* diperoleh dari persamaan (2) sebagai berikut.

$$Support(X \rightarrow Y) = \frac{|XUY|}{|D|} \quad (1)$$

$$Confidence(X \rightarrow Y) = \frac{|XUY|}{|X|} = \frac{Support(XUY)}{Support(X)} \quad (2)$$

Dimana $|X \cup Y|$ menunjukkan total itemset X dan Y muncul bersamaan, sedangkan $|D|$ mewakili jumlah item secara keseluruhan pada *dataset*. Sedangkan $|X|$ menunjukkan jumlah kemunculan hanya *itemset* X.

Apriori menggunakan pendekatan iteratif dimana *k-itemset* digunakan untuk mengeksplorasi ($k+1$) *itemset* (Han dkk., 2012). Fauzy dkk. (2016) mengungkapkan bahwa terdapat dua tahap pada setiap proses iterasi, yaitu tahap pembangkitan kandidat dan tahap pembangkitan rule. Pada tahap pembangkitan kandidat (*candidate generation*) menggunakan himpunan semua frequent($k-1$)-itemset F_{k-1} yang ditemukan pada pass ke- $(k-1)$ untuk membangkitkan kandidat *itemset* C_k . Kemudian data dipindai dalam tahap *support counting*. Pada akhir pass C_k diperiksa untuk menemukan kandidat mana yang sering muncul, menghasilkan F_k . Perhitungan *support* berakhir ketika F_k atau C_{k-1} kosong.

Sedangkan pada tahap pembangkitan *rule*, *candidate rule* akan dibangkitkan terlebih dahulu. *Candidate rule* berisi semua kemungkinan *rule* yang memiliki *support* lebih besar dari *minimum support*. Kemudian *candidate rule* akan digabung dengan Tabel F untuk menemukan *support antecedent*. *Candidate rule* dihitung dengan cara membandingkan *support rule* dengan *support antecedent rule*. Hanya *rule* yang memiliki nilai *confidence* lebih besar dari *minimum confidence* yang disimpan dalam Tabel *rule*.

Secara umum, teknik di atas menggunakan kerangka *support-confidence* sebagai indikator untuk memangkas *itemset* yang jarang muncul pada *dataset* (Wu dkk., 2009). Namun, indikator tersebut memiliki kelemahan karena masih menghasilkan sejumlah besar aturan asosiasi yang tidak menarik (Weng dkk., 2016). Untuk mengatasi masalah ini, *lift* diusulkan sebagai indikator tambahan karena *lift* dapat mempertimbangkan seberapa besar kemungkinan terjadinya perubahan item Y mengingat item X telah terjadi (Cunjin dkk., 2015). *lift* dapat diketahui melalui perbandingan antara nilai *confidence* dengan nilai *benchmark confidence*. *Benchmark confidence* merupakan nilai perbandingan antara jumlah semua *item* yang

menjadi *consequent* terhadap total jumlah transaksi (Santosa, 2007). Nilai *lift* diperoleh dari persamaan (3) berikut.

$$Lift(X \rightarrow Y) = \frac{Conf(XUY)}{Support(Y)} = \frac{Support(XUY)}{Support(X) \cdot Support(Y)} \quad (3)$$

Menurut Brijs dkk. (2003), jika nilai *lift* < 1 , maka X dan Y dianggap lebih jarang muncul secara bersamaan atau dengan kata lain saling bergantung secara negatif. Jika *lift* = 1, maka X dan Y tidak bergantung satu sama lain. Jika *lift* > 1 , maka X dan Y dianggap saling bergantung secara positif. Pada penelitian ini, nilai *lift* digunakan untuk memangkas hasil asosiasi untuk menemukan *rule* yang unik atau saling bergantung diantara *rule* yang lain. Sehingga dapat menghasilkan sejumlah aturan asosiasi yang menarik.

Selanjutnya dilakukan analisis pada hasil asosiasi untuk mencari aturan asosiasi yang paling berguna. Selain itu dilakukan proses pengujian dengan memanfaatkan nilai *support*, *confidence*, dan *lift* sebagai acuan. Nilai *support* dan *confidence* digunakan untuk mengetahui pengaruh kedua nilai tersebut terhadap jumlah *rule* yang terbentuk. Selain itu kedua nilai tersebut juga digunakan untuk mengetahui nilai rata-rata nilai *lift ratio* yang dihasilkan. Semakin tinggi nilai *lift* yang diperoleh, maka aturan asosiasinya akan semakin kuat, begitupun sebaliknya.

Setelah diperoleh hasil asosiasi, maka hasil tersebut akan divisualisasikan dalam bentuk *Dashboard* menggunakan software Power BI. Visualisasi yang dihasilkan berupa 4 antarmuka *Dashboard*, yaitu *Dashboard* data ulasan, *Dashboard* frekuensi wahana setiap aspek, *Dashboard* frekuensi kata ulasan setiap aspek, dan *Dashboard* hasil asosiasi ulasan Jatim Park 3. *Dashboard* tersebut juga akan diuji dengan metode *System Usability Scale* (SUS). Menurut Sauro (2011), metode ini dapat digunakan pada sampel yang jumlahnya sedikit. SUS merupakan kuesioner yang memiliki 10 pertanyaan. Setiap pertanyaan memiliki 5 jawaban berupa skala 0 sampai 4. Kemudian nilai SUS dirata-rata. Berdasarkan hasil rata-rata, nilainya dapat dikelompokkan ke dalam 3 kategori dengan rentang yang berbeda-beda yaitu :

1. *Not Acceptable* : 0-50.
2. *Marginal* : 50-70.
3. *Acceptable* : 70-100

3. PENELITIAN TERKAIT

Metode yang digunakan pada penelitian ini merujuk pada beberapa penelitian sebelumnya. Sejumlah besar penelitian sudah menggunakan teknik ARM untuk menemukan hubungan suatu item. Pada penelitian Hong dkk. (2020) memanfaatkan metode ARM untuk mengidentifikasi hubungan tersembunyi antara faktor-faktor penyebab terjadinya kecelakaan yang melibatkan kendaraan HAZMAT di jalan tol. Pada penelitian ini

tidak menggunakan aturan khusus untuk menetapkan ambang batas. Metode *trial and error* digunakan untuk menemukan ambang batas yang sesuai untuk menemukan aturan yang menarik.

Isanta dkk. (2016) juga menggunakan ARM untuk mengetahui seberapa kuat relasi antara target *tweet* dengan kata kunci melalui nilai *confidence*. Pada penelitian ini menghasilkan tren suatu brand berdasarkan kata dengan jumlah kemunculan terbesar serta memanfaatkan fitur pada *twitter* seperti *favourite* dan *retweet*. Penelitian ini menghasilkan kesimpulan bahwa semakin tinggi nilai *confidence* maka relasi antara target *tweet* dan kata kunci akan semakin kuat.

Sedangkan pada penelitian yang dilakukan oleh Rekik dkk. (2018) memanfaatkan teknik ARM menggunakan algoritma Apriori untuk menemukan aturan asosiasi antara kriteria dan kategori situs web guna mempelajari interdependensi dalam kriteria yang muncul dari kumpulan literatur yang diperoleh dari publikasi ilmiah dengan lingkup topik layanan web. Ambang batas yang ditetapkan pada penelitian ini antara lain *minimum support* 5% dan *minimum confidence* 25% serta nilai *lift ratio* lebih besar dari 1. Hasil dari penelitian ini divisualisasikan dalam bentuk *network graph* dan *dependency wheel graph* pada setiap kategori yang dihasilkan.

Penelitian lainnya dari Cozza dkk. (2018) yang menerapkan *class Association Rule* untuk mendapatkan hubungan diantara properti hotel, karakteristik pemberi ulasan, dan skor ulasan, dimana setiap item dibagi sesuai dengan kelas yang berbeda. Pada penelitian ini juga dilakukan ekstraksi *hotel-centric* dan *reviewer-centric features* untuk meningkatkan kumpulan data ulasan dengan fitur tambahan berdasarkan karakteristik peninjau dan hotel. Pendekatan pada penelitian ini dapat membantu penyedia layanan hotel untuk mencapai kesadaran yang lebih baik tentang cara membaca ulasan serta cara meningkatkan layanan tanpa bergantung pada wawancara.

Algoritma yang digunakan dalam mengimplementasikannya adalah algoritma Apriori. Algoritma ini dianggap lebih efisien dibandingkan teknik lainnya karena memiliki kemampuan untuk menghindari pemborosan upaya menghitung *itemset* kandidat yang jarang (Hong dkk., 2020). Selain itu, algoritma ini juga dapat melakukan iterasi dengan cara memindai database untuk menemukan pola frekuensi tinggi.

4. HASIL DAN PEMBAHASAN

4.1. Pengujian Hasil Asosiasi

a. Pengujian Terhadap Jumlah Rule yang Terbentuk

Pengujian dilakukan dengan menggunakan nilai *minimum support* mulai dari 5% sampai dengan 40% dan nilai *minimum confidence*

mulai dari 5% sampai dengan 100%. Batas nilai *minimum support* dan *minimum confidence* diambil berdasarkan dengan batas dimana *rule* tidak lagi dihasilkan. Khusus untuk aspek keamanan dan teknologi, pengujian dilakukan dengan menggunakan rentang nilai *minimum support* dari 10% sampai dengan 80%. Berikut merupakan Tabel yang menunjukkan jumlah *rule* yang terbentuk pada masing-masing aspek.

Tabel 2. Pengujian Terhadap Jumlah Rule pada Aspek Akses Jalan

Minimum confidence (%)	Jumlah rule yang terbentuk							
	Minimum support (%)							
	5	10	15	20	25	30	35	40
5	8089	424	87	30	8	2	0	0
10	8089	424	87	30	8	2	0	0
20	8010	424	87	30	8	2	0	0
30	7788	397	87	30	8	2	0	0
40	7373	354	74	30	8	2	0	0
50	7087	289	61	21	8	2	0	0
60	6579	206	27	12	5	2	0	0
70	5822	153	27	8	4	2	0	0
80	4419	106	13	3	1	1	0	0
90	4184	52	6	2	1	1	0	0
100	4184	52	6	2	1	1	0	0

Tabel 3. Pengujian Terhadap Jumlah Rule pada Aspek Biaya

Minimum confidence (%)	Jumlah rule yang terbentuk							
	Minimum support (%)							
	5	10	15	20	25	30	35	40
5	819	77	22	4	2	2	2	2
10	800	77	22	4	2	2	2	2
20	733	69	22	4	2	2	2	2
30	619	56	19	4	2	2	2	2
40	472	47	14	3	2	2	2	2
50	389	45	14	3	2	2	2	2
60	286	38	12	3	2	2	2	2
70	202	22	8	1	1	1	1	1
80	108	5	1	0	0	0	0	0
90	31	0	0	0	0	0	0	0
100	23	0	0	0	0	0	0	0

Tabel 4. Pengujian Terhadap Jumlah Rule pada Aspek Kebersihan

Minimum confidence (%)	Jumlah rule yang terbentuk							
	Minimum support (%)							
	5	10	15	20	25	30	35	40
5	9912	288	87	20	12	6	6	0
10	9875	288	87	20	12	6	6	0
20	9754	275	82	20	12	6	6	0
30	9438	255	74	14	10	5	6	0
40	8857	227	67	12	8	5	5	0
50	8563	165	57	10	6	3	3	0
60	7959	135	44	10	6	3	3	0
70	7158	106	34	10	6	3	3	0
80	5438	75	30	10	6	3	3	0
90	5411	70	28	10	6	3	3	0
100	5408	67	25	8	4	1	1	0

Tabel 5. Pengujian Terhadap Jumlah Rule pada Aspek Kepuasan

Minimum confidence (%)	Jumlah rule yang terbentuk							
	Minimum support (%)							
	5	10	15	20	25	30	35	40
5	309	20	0	0	0	0	0	0

Jumlah rule yang terbentuk								
10	309	20	0	0	0	0	0	0
20	294	20	0	0	0	0	0	0
30	206	20	0	0	0	0	0	0
40	146	17	0	0	0	0	0	0
50	86	8	0	0	0	0	0	0
60	56	5	0	0	0	0	0	0
70	25	2	0	0	0	0	0	0
80	15	1	0	0	0	0	0	0
90	8	0	0	0	0	0	0	0
100	2	0	0	0	0	0	0	0

Jumlah rule yang terbentuk								
10	-	-	91	91	16	9	9	2
20	-	-	91	91	16	9	9	2
30	-	-	91	91	16	9	9	2
40	-	-	91	91	16	9	9	2
50	-	-	91	91	16	9	9	2
60	-	-	76	76	16	9	9	2
70	-	-	69	69	13	9	9	2
80	-	-	63	63	13	9	9	2
90	-	-	60	60	10	6	6	2
100	-	-	60	60	10	6	6	2

Tabel 6. Pengujian Terhadap Jumlah Rule pada Aspek Keramaian

Jumlah rule yang terbentuk								
Minimum m confidence (%)	Minimum support (%)							
	5	10	15	20	25	30	35	40
5	2416	112	22	8	2	0	0	0
10	2390	112	22	8	2	0	0	0
20	2289	106	22	8	2	0	0	0
30	1993	91	19	8	2	0	0	0
40	1621	77	14	6	2	0	0	0
50	1349	60	12	4	1	0	0	0
60	1029	44	11	4	1	0	0	0
70	804	32	9	3	1	0	0	0
80	538	18	4	2	1	0	0	0
90	282	11	3	2	1	0	0	0
100	269	2	0	0	0	0	0	0

Tabel 7. Pengujian Terhadap Jumlah Rule pada Aspek Pelayanan

Jumlah rule yang terbentuk								
Minimum m confidence (%)	Minimum support (%)							
	5	10	15	20	25	30	35	40
5	-	892	123	25	12	4	0	0
10	-	892	123	25	12	4	0	0
20	-	866	123	25	12	4	0	0
30	-	815	107	21	12	4	0	0
40	-	707	106	20	11	4	0	0
50	-	625	83	15	6	2	0	0
60	-	468	56	14	6	2	0	0
70	-	369	45	13	6	2	0	0
80	-	337	31	6	3	1	0	0
90	-	182	20	4	3	1	0	0
100	-	179	17	1	0	0	0	0

Tabel 8. Pengujian Terhadap Jumlah Rule pada Aspek Keamanan

Jumlah rule yang terbentuk								
Minimum m confidence (%)	Minimum support (%)							
	10	20	30	40	50	60	70	80
5	-	-	-	24	24	9	2	2
10	-	-	-	24	24	9	2	2
20	-	-	-	24	24	9	2	2
30	-	-	-	24	24	9	2	2
40	-	-	-	24	24	9	2	2
50	-	-	-	24	24	9	2	2
60	-	-	-	19	19	9	2	2
70	-	-	-	16	16	8	2	2
80	-	-	-	14	14	6	2	2
90	-	-	-	13	13	5	1	1
100	-	-	-	13	13	5	1	1

Tabel 9. Pengujian Terhadap Jumlah Rule pada Aspek Teknologi

Jumlah rule yang terbentuk								
Minimum m confidence (%)	Minimum support (%)							
	10	20	30	40	50	60	70	80
5	-	-	91	91	16	9	9	2

b. Pengujian Kinerja Rule yang Dihasilkan

Pengujian ini dilakukan untuk mengetahui rata-rata nilai lift ratio yang dihasilkan dengan rentang nilai minimum support mulai dari 5% sampai dengan 40% dan nilai minimum confidence mulai dari 5% sampai dengan 100%. Khusus untuk aspek keamanan dan teknologi, pengujian dilakukan dengan menggunakan rentang nilai minimum support dari 10% sampai dengan 80%. Keseluruhan aspek menghasilkan nilai rata-rata lift ratio di atas 1. Tabel 10 menunjukkan nilai rata-rata lift ratio secara keseluruhan pada masing-masing aspek.

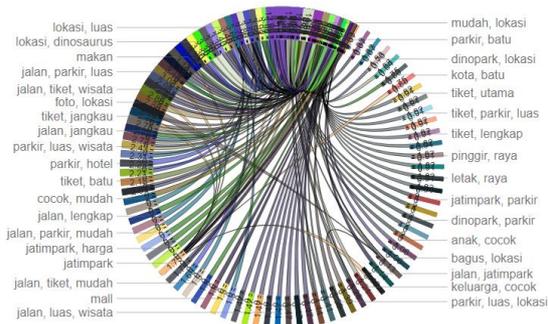
Tabel 10. Nilai Rata-rata Lift ratio pada Setiap Aspek

Aspek	Nilai Rata-rata Lift ratio
Akses Jalan	2,36
Biaya	1,43
Kebersihan	1,57
Kepuasan	2,76
Keramaian	1,75
Pelayanan	1,56
Keamanan	1,05
Teknologi	1,05

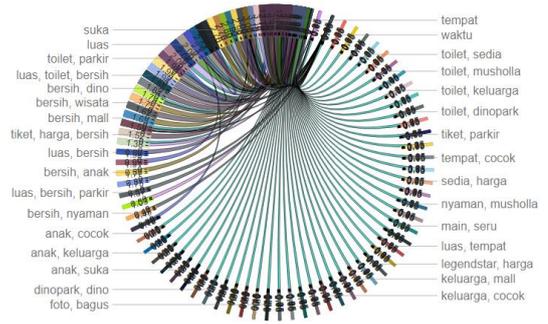
4.2 Hasil Asosiasi

a. Aspek Akses Jalan

Aspek akses jalan menggambarkan kondisi mengenai lokasi beserta kemudahan akses wisata Jatim Park 3. Ulasan yang tergolong dalam aspek ini terdapat sebanyak 42 ulasan. Rule yang terbentuk sebanyak 289 rule yang memenuhi ketentuan minimum support 10%, minimum confidence 50%, dan nilai lift samadengan atau lebih dari 1. Gambar 3 menunjukkan presentasi hubungan antar itemset yang membentuk rule pada aspek akses jalan. Dari Gambar tersebut terlihat cukup banyak rule yang terbentuk. Kata ‘lokasi’, ‘mudah’ dan ‘jalan’ banyak terlihat pada Gambar ini, hal tersebut dapat diartikan bahwa akses jalan menuju lokasi Jatim Park 3 terbilang mudah karena pada nyatanya Jatim Park 3 terletak di pinggir jalan raya sehingga mudah terlihat dan diakses.



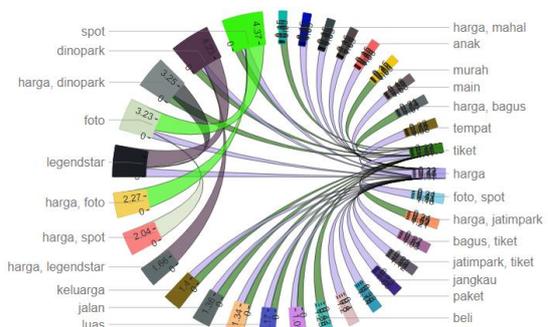
Gambar 3. Hubungan *Rule* Hasil Asosiasi pada Aspek Akses Jalan



Gambar 5. Hubungan *Rule* Hasil Asosiasi pada Aspek Kebersihan

b. Aspek Biaya

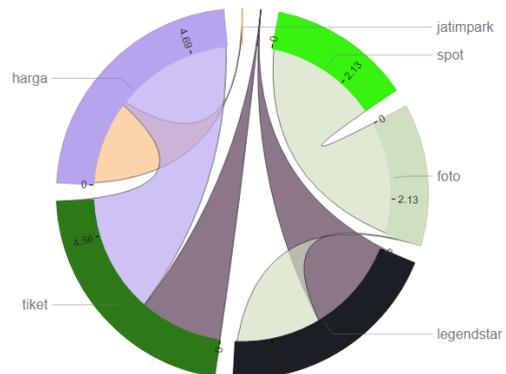
Aspek biaya menggambarkan kondisi mengenai tingkat biaya yang dirasakan pengunjung untuk menikmati Jatim Park 3. Ulasan yang tergolong dalam aspek ini terdapat sebanyak 132 ulasan. *Rule* yang terbentuk sebanyak 44 *rule* yang memenuhi ketentuan *minimum support* 10%, *minimum confidence* 50%, dan nilai *lift* diatas atau samadengan 1. Gambar 4 menunjukkan presentasi hubungan antar itemset yang membentuk *rule* pada aspek biaya. Pada Gambar ini terlihat bahwa *itemset* 'spot' memiliki nilai *lift* paling tinggi. Sedangkan *itemset* 'tiket' dan 'harga' memiliki hubungan paling banyak dengan *itemset* lain.



Gambar 4. Hubungan *Rule* Hasil Asosiasi pada Aspek Biaya

d. Aspek Kepuasan

Aspek kepuasan menggambarkan kondisi mengenai tingkat kepuasan pengunjung terhadap Jatim Park 3. Ulasan yang tergolong dalam aspek ini terdapat sebanyak 172 ulasan. *Rule* yang terbentuk sebanyak 8 *rule* yang memenuhi ketentuan *minimum support* 10%, *minimum confidence* 50%, dan nilai *lift* diatas atau samadengan 1. Gambar 6 menunjukkan presentasi hubungan antar itemset yang membentuk *rule* pada aspek kepuasan. Pada Gambar ini terlihat bahwa tidak banyak *itemset* yang memenuhi serta tidak banyak *rule* yang terbentuk. Pada Gambar ini juga terlihat bahwa *itemset* 'tiket' memiliki nilai *lift* paling tinggi.



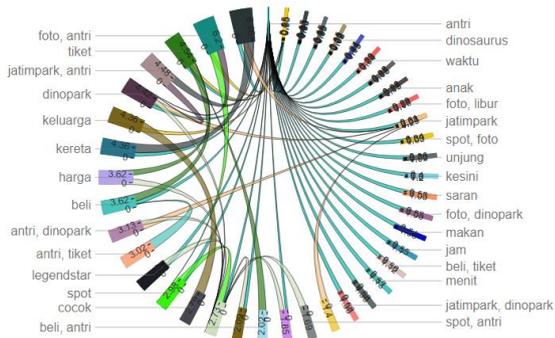
Gambar 6. Hubungan *Rule* Hasil Asosiasi pada Aspek Kepuasan

c. Aspek Kebersihan

Aspek kebersihan menggambarkan kondisi mengenai kondisi kebersihan pada Jatim Park 3. Ulasan yang tergolong dalam aspek ini terdapat sebanyak 53 ulasan. *Rule* yang terbentuk sebanyak 165 *rule* yang memenuhi ketentuan *minimum support* 10%, *minimum confidence* 50%, dan nilai *lift* diatas atau samadengan 1. Gambar 5 menunjukkan presentasi hubungan antar itemset yang membentuk *rule* pada aspek kebersihan. Pada Gambar ini terlihat bahwa aspek kebersihan menghasilkan *rule* yang cukup banyak. Kata 'bersih' dan 'nyaman' cukup terlihat pada Gambar ini, hal tersebut dapat diartikan bahwa tingkat kebersihan pada wisata Jatim Park 3 tergolong bersih dan nyaman.

e. Aspek Keramaian

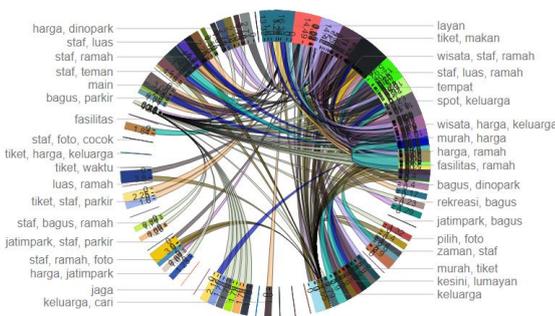
Aspek keramaian menggambarkan kondisi mengenai tingkat keramaian terhadap Jatim Park 3. Ulasan yang tergolong dalam aspek ini terdapat sebanyak 94 ulasan. *Rule* yang terbentuk sebanyak 60 *rule* yang memenuhi ketentuan *minimum support* 10%, *minimum confidence* 50%, dan nilai *lift* diatas atau samadengan 1. Gambar 7 menunjukkan presentasi hubungan antar itemset yang membentuk *rule* pada aspek keramaian. Gambar tersebut juga menunjukkan *itemset* 'antri' memiliki hubungan paling banyak dengan *itemset* lain. Sehingga dapat diartikan masih terdapat titik-titik antrian pada wisata Jatim Park 3.



Gambar 7. Hubungan Rule Hasil Asosiasi pada Aspek Keramaian

f. Aspek Pelayanan

Aspek pelayanan menggambarkan kondisi mengenai tingkat pelayanan yang diberikan oleh Jatim Park 3. Ulasan yang tergolong dalam aspek ini terdapat sebanyak 39 ulasan. Rule yang terbentuk sebanyak 625 rule yang memenuhi ketentuan *minimum support* 10%, *minimum confidence* 50%, dan nilai *lift* diatas atau samadengan 1. Gambar 8 menunjukkan presentasi hubungan antar itemset yang membentuk rule pada aspek pelayanan. Pada Gambar ini terlihat bahwa cukup banyak rule yang terbentuk pada aspek ini. kata 'ramah' cukup banyak terlihat pada Gambar ini, hal tersebut dapat diartikan bahwa tingkat pelayanan dari staf pada wisata Jatim Park tergolong bagus karena dianggap ramah.

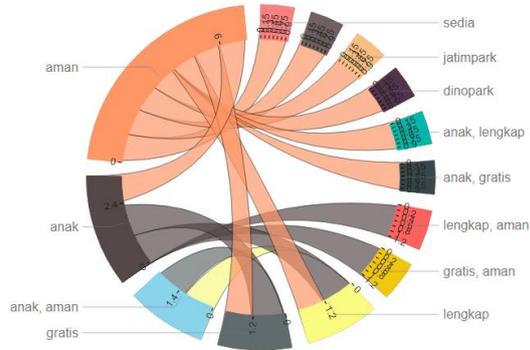


Gambar 8. Hubungan Rule Hasil Asosiasi pada Aspek Pelayanan

g. Aspek Keamanan

Aspek keamanan menggambarkan kondisi mengenai tingkat keamanan pada Jatim Park 3. Ulasan yang tergolong dalam aspek ini terdapat sebanyak 6 ulasan. Rule yang terbentuk sebanyak 24 rule yang memenuhi ketentuan *minimum support* 40%, *minimum confidence* 50%, dan nilai *lift* diatas atau samadengan 1. Nilai *minimum support* pada aspek ini berbeda dengan aspek sebelumnya. Jika nilai *minimum support* 10%, maka sistem akan mengalami runtime out. Gambar 9 menunjukkan presentasi hubungan antar itemset yang membentuk rule pada aspek keamanan. Pada Gambar ini terlihat

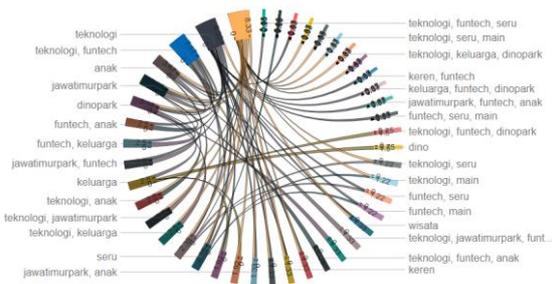
bahwa itemset 'aman' memiliki hubungan paling banyak dengan itemset lain. Hal tersebut menunjukkan bahwa berdasarkan ulasan, Jatim Park 3 tergolong aman.



Gambar 9. Hubungan Rule Hasil Asosiasi pada Aspek Keamanan

h. Aspek Teknologi

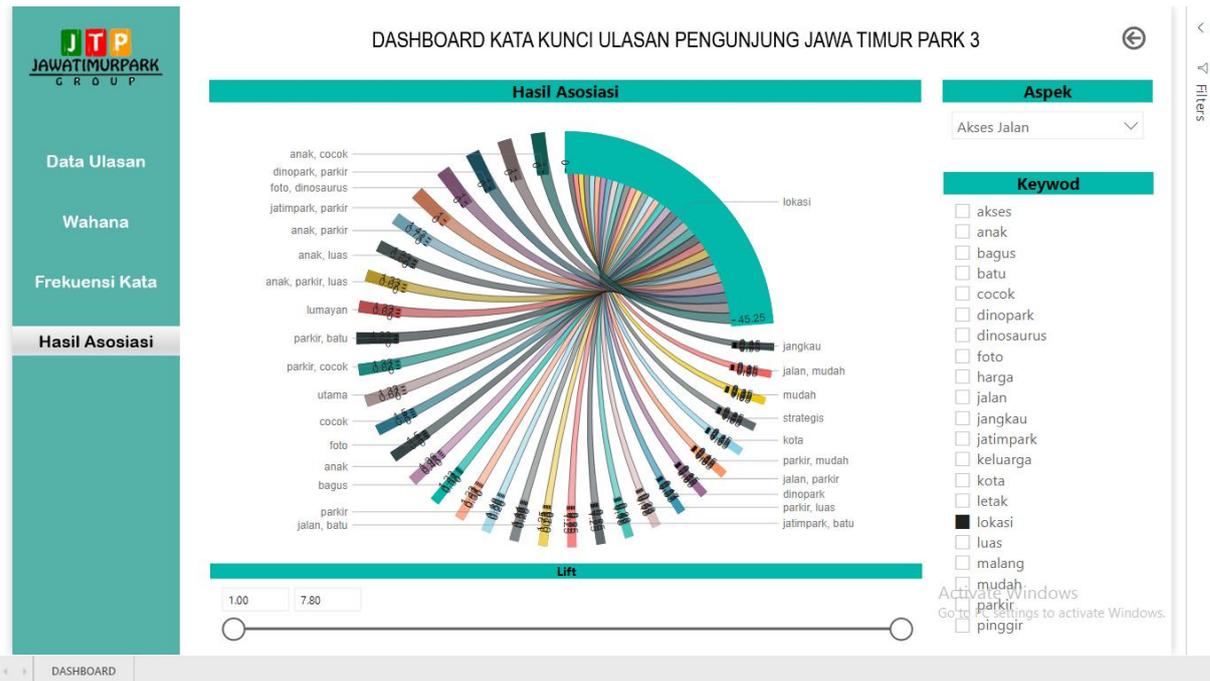
Aspek teknologi menggambarkan kondisi mengenai tingkat penerapan teknologi yang terdapat pada Jatim Park 3. Ulasan yang tergolong dalam aspek ini terdapat sebanyak 7 ulasan. Rule yang terbentuk sebanyak 91 rule yang memenuhi ketentuan *minimum support* 40%, *minimum confidence* 50%, dan nilai *lift* diatas atau samadengan 1. Nilai *minimum support* pada aspek ini sama dengan aspek keamanan karena memiliki kondisi yang sama. Gambar 10 menunjukkan presentasi hubungan antar itemset yang membentuk rule pada aspek teknologi. Pada Gambar ini terlihat bahwa kata 'funtech' dan 'teknologi' memiliki hubungan lebih banyak dengan kata lainnya.



Gambar 10. Hubungan Rule Hasil Asosiasi pada Aspek Teknologi

4.3. Visualisasi Dashboard

Visualisasi hasil asosiasi ditampilkan dalam bentuk Dashboard yang ditunjukkan pada Gambar 11. Data hasil asosiasi divisualisasikan menggunakan bentuk chord chart. Data masukan yang digunakan adalah data antecedent dan concequent beserta nilai lift nya

Gambar 11. *Dashboard Hasil Asosiasi*

4.4 Hasil Usability Testing

Metode yang digunakan dalam *Usability Testing* adalah *System Usability Scale* (SUS). Berdasarkan hasil SUS yang dilakukan oleh HRD Holding Jatim Park Group, maka *SUS Score* yang diperoleh untuk *Dashboard* adalah 62,5. *Dashboard* tergolong ke dalam kategori Marginal (high) dan dalam rating “Good”. Sehingga dapat disimpulkan bahwa *Dashboard* memiliki peringkat yang bagus namun hanya memiliki tingkat penerimaan dalam kategori marjinal atau belum memenuhi kategori acceptable oleh pihak Jatim Park 3.

5. KESIMPULAN

Pada penelitian ini dihasilkan 8 aspek yang digunakan berdasarkan observasi data ulasan dan persetujuan pihak Jatim Park 3. Aspek tersebut antara lain akses jalan, biaya, kebersihan, kepuasan, keramaian, pelayanan, keamanan, dan teknologi. Dari proses pemberian label pada data ulasan dihasilkan kelompok data pada masing-masing aspek, antara lain aspek akses jalan memiliki 42 ulasan, aspek biaya memiliki 132 ulasan, aspek kebersihan memiliki 53 ulasan, aspek kepuasan memiliki 172 ulasan, aspek keramaian memiliki 94 ulasan, aspek pelayanan memiliki 39 ulasan, aspek keamanan memiliki 6 ulasan, dan aspek teknologi memiliki 7 ulasan.

Implementasi *Association Rule Mining* menggunakan algoritma Apriori membentuk *rule* dengan jumlah yang berbeda pada setiap aspeknya. Aspek akses jalan menghasilkan *rule* sebanyak 289 *rule*, aspek biaya menghasilkan *rule* sebanyak 44 *rule*, aspek kebersihan menghasilkan *rule* sebanyak 165 *rule*, aspek kepuasan menghasilkan *rule*

sebanyak 8 *rule*, aspek keramaian menghasilkan *rule* sebanyak 59 *rule*, aspek pelayanan menghasilkan *rule* sebanyak 625 *rule*, aspek keamanan menghasilkan *rule* sebanyak 24 *rule*, sedangkan aspek teknologi menghasilkan *rule* sebanyak 91 *rule*. Keseluruhan *rule* yang terbentuk menghasilkan rata-rata *lift ratio* di atas 1 sehingga dimana *rule* dengan nilai *lift ratio* di atas 1 tersebut merupakan *rule* yang unik diantara *rule*-rule lain yang terbentuk dari asosiasi tersebut.

Berdasarkan hasil tersebut, Algoritma Apriori bersifat fleksibel sehingga dapat menguntungkan karena tidak memiliki fungsi yang ditentukan dan tidak memerlukan variabel dependen. Hal tersebut membuat peneliti dapat menyesuaikan dengan kebutuhan topik penelitian. Namun Apriori juga tidak memiliki aturan khusus untuk menentukan *threshold* sehingga terdapat perbedaan nilai *threshold* pada setiap aspek untuk menemukan aturan asosiasi yang menarik

Proses *usability testing* pada *Dashboard* menggunakan *System Usability Scale* (SUS) telah menghasilkan nilai yang baik yakni 62.5 yang berarti *Dashboard* tersebut tergolong ke dalam kategori Marginal (high). Proses visualisasi menghasilkan 4 *Dashboard* yakni *Dashboard* data ulasan, *Dashboard* wahana, *Dashboard* frekuensi kata, dan *Dashboard* hasil asosiasi.

Pada penelitian berikutnya dapat menambahkan sumber data ulasan pengunjung dari sumber lain seperti, TripAdvisor, dan platform lain sejenisnya yang memuat data ulasan pengunjung wisata Jatim Park 3. Sehingga penerapan kerja asosiasi wisata dapat menghasilkan asosiasi kata yang lebih beragam dan informatif. Selain itu, lingkup penelitian juga dapat dikembangkan dengan

memfokuskan pengelompokan data ulasan per wahana yang terdapat pada Jatim Park 3.

Penerapan *stoplist* pada *stop-word removal* dapat membuat baru dengan menggunakan metode hukum Zipf. Hal tersebut diharapkan dapat menghasilkan *stoplist* yang sesuai dengan topik pariwisata. Pemilihan aspek juga dapat menggunakan pendekatan ekstraksi aspek dengan metode *Part-Of-Speech Tagger* untuk mendapatkan aspek yang memang benar-benar ada dalam dataset.

UCAPAN TERIMA KASIH

Penulis berterimakasih kepada Manajemen Jatim Park Group terutama manajemen Jatim Park 3 karena sudah mengizinkan penelitian serta membantu selama penelitian.

DAFTAR PUSTAKA

- BRIJS, T., VANHOOF, K. dan WETS, G., 2003. Defining Interestness for Association Rules. *Institute of Information Theories and Applications FOI ITHEA*, 10(4), pp. 370-376.
- COZZA, V., PETROCCHI, M. dan SPOGNARDI, A., 2018. Mining Implicit Data Association from Tripadvisor Hotel Reviews. *CEUR Workshop Proceedings*. 2083, (2018), pp. 56-61.
- CUNJIN, X., WANJIAO, S., LIJUAN, Q., QING, D. & XIAOYANG, W., 2015. A Mutual-information-based Mining method for Marine Abnormal Association Rules *Comput. Geosci*, 76, pp. 121-129.
- FAUZY, M., SALEH, K.R. dan ASROR, I., 2016. Penerapan Metode Association Rule Menggunakan Algoritma Apriori Pada Simulasi Prediksi Hujan Wilayah Kota Bandung. *Jurnal Ilmiah Teknologi Informasi Terapan*, Vol. II No.2.
- FELDMAN, R. dan SANGER, J., 2007. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. New York: Cambridge University Press.
- HAN, J., KAMBER, M. dan PEI, J., 2012. *Data Mining Concepts and Techniques*. 3th ed. USA: Morgan Kaufmann Publishers.
- HONG, J., TAMAKLOE, R. dan PARK, D., 2020. Application of Association Rules Mining Algorithm for Hazardous Materials Transportation Crashes on Expressway. *Accident Analysis and Prevention*, Vol 142.
- ISANTA, S.A., FATICHAH, C. dan PURWITASARI, D., 2016. Ekstraksi Kata Kunci Metadata Twitter Berbahasa Indonesia dengan Pendekatan Grammatical Tagging untuk Visualisasi Trend Produk Brand. S1. Institut Teknologi Sepuluh Nopember.
- LEE, S., CHA, Y., HAN, S. & HYUN, C., 2019. Application of Association Rule Mining and Social Network Analysis for Understanding Causality of Construction Defects. *Sustainability* 11, 3.
- OLSON, D. dan SHI, Y., 2008. *Pengantar Ilmu Panggilan Data Bisnis*. Jakarta. Mc GrawHill.
- REKIK, R., KALLEL, I., CASILLAS, J., & ALIMI, A.M., 2018. Assessing Web Sites Quality: A Systematic Literature Review By Text and Association Rule Mining. *International Journal of Information Management*, 38, pp. 201-216.
- SANTOSA, B., 2007. *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- SAURO, J., 2011. *Measuring Usability With The System Usability Scale (SUS)*. Tersedia di: <<https://measuringu.com/sus/>> [Diakses 02 Februari 2020].
- WENG, J., ZHU, J.Z., YAN, X. & LIU, Z., 2016. Investigation of Work Zone Crash Casualty Patterns Using Association Rules. *Accident Analysis & Prevention*, 92, pp.43-52.
- WU, H., LU, Z., PAN, L. & XU, R., 2009. An Improved Apriori-based Algorithm for Association Rules Mining. *Sixth International Conference on Fuzzy Systems and Knowledge Discovery*, pp. 51-55.
- WU, B., ZHANG, J.H., YAN, X.P. & YIP, T.L., 2019. Use of Association Rules for Cause-effects Relationships Analysis of Collision Accidents in The Yangtze River. In: Weintrit, A., Neumann, T. (Eds.), *Advances in Marine Navigation and Safety of Sea Transportation*. CRC Press/Balkema, Leiden, pp. 65.
- YU, S., JIA, Y. dan SUN, D., 2019. Identifying factors that influence the patterns of road crashes using association rules: a case study from Wisconsin, United States. *Sustainability*, 11 (7), 1-14.