

# Penentuan Aspek Implisit dengan Ekstraksi *Knowledge* Berbasis *Rule* pada Ulasan Bahasa Indonesia

## (*Determination of Implicit Aspects with Rule Based Knowledge Extraction in Indonesian Reviews*)

Yuliana Setiowati<sup>1</sup>, Fitri Setyorini<sup>2</sup>, Afrida Helen<sup>3</sup>

**Abstract**—Determination of implicit aspects is one of the important things in opinion sentences. This study proposes a new approach for developing rule-based knowledge by forming a relation between opinion words with aspect categories. The relationship is obtained from the combination of rules, based on Opinion Word Similarity (OWS). Evaluation for rule-based knowledge extraction is in the form of threshold values of frequency and confidence to produce the best precision, recall, and f-measure values. The knowledge extraction consists of two phases: training phase and testing phase. The training phase is described as the process to extract rule-based knowledge. The testing phase is described as the process to obtain the implicit aspects of opinion sentences by referring to rule-based knowledge. To extract rule-based knowledge on user reviews, it is necessary to identify opinion sentences with explicit aspects and get pairs of aspects and words of opinion with rules generated from regular expressions. The evaluation result of rule-based knowledge with confidence using OWS showed better results compared to rule-based knowledge without using OWS. By using OWS, precision value increased by 0.25%, recall value increased by 1.15%, and precision value increased by 0.83%.

**Intisari**—Penentuan aspek implisit adalah salah satu hal penting dalam kalimat opini. Makalah ini mengajukan pendekatan dengan mengembangkan *knowledge* berbasis *rule* berupa relasi antara kata opini dengan kategori aspek, dengan melakukan penggabungan *rule* berdasarkan *Opinion Word Similarity* (OWS). Evaluasi untuk ekstraksi *knowledge* berbasis *rule* berupa nilai *threshold* dari frekuensi dan *confidence* untuk menghasilkan nilai *precision*, *recall*, dan *f-measure* terbaik. Terdapat dua fase, yaitu proses pelatihan atau proses untuk mengekstrak *knowledge* berbasis *rule*, dan fase pengujian, yaitu fase untuk mendapatkan aspek implisit pada kalimat opini dengan mengacu ke *knowledge* berbasis *rule*. Untuk mengekstrak *knowledge* berbasis *rule* pada ulasan pengguna, perlu diidentifikasi kalimat opini dengan aspek eksplisit dan mendapatkan pasangan aspek dan kata opini dengan *rule* yang dibangkitkan dari *regular expression*. Hasil evaluasi *knowledge* berbasis *rule* dengan frekuensi atau *confidence* menggunakan OWS menunjukkan hasil yang lebih bagus dibandingkan dengan *knowledge* tanpa menggunakan OWS, yaitu nilai *precision*

meningkat sebesar 0,25%, *recall* meningkat sebesar 1,15%, dan *precision* meningkat sebesar 0,83%.

**Kata Kunci**—Aspek Implisit, *Knowledge* Berbasis *Rule*, *Opinion Word Similarity*, *Confidence*, Frekuensi.

### I. PENDAHULUAN

Penelitian di bidang Sentimen Analisis berbasis Aspek (*Aspect Based Sentiment Analysis/ABSA*) mengalami perkembangan yang sangat pesat pada saat ini, termasuk di Indonesia. Manfaat penelitian ABSA di domain hotel di antaranya untuk mengetahui respons pengguna hotel terhadap kinerja departemen pada sebuah hotel [1], [2]; mengetahui ulasan pengguna restoran dalam bentuk *rating* dalam hal makanan, layanan, harga, dan kenyamanan tempat [3]; dan untuk mengetahui keunggulan produk dengan produk lainnya [4].

Dalam penelitian ABSA, mendapatkan pasangan aspek dan kata opini adalah hal penting. Aspek merupakan fitur atau layanan yang sedang dibahas dalam kalimat opini pengguna, sedangkan kata opini adalah penilaian pengguna mengenai aspek tersebut. Terdapat dua tipe untuk mendapatkan pasangan aspek dan kata opini, yaitu tipe 1, yang mendapatkan aspek terlebih dahulu, selanjutnya mendapatkan kata opini dengan jarak tertentu [5], [6]; dan tipe 2, yang mendapatkan pasangan secara langsung menggunakan *rule* [7]. Kelebihan tipe 1 adalah setiap aspek pasti mendapatkan kata opini terdekat, tetapi kelemahannya memungkinkan bukan pasangan dari kata aspek tersebut. Kelebihan dari tipe 2 yaitu mendapatkan pasangan aspek dan kata opini secara langsung, tetapi memungkinkan beberapa aspek tidak mendapatkan pasangan kata opini. Beberapa metode untuk mendapatkan aspek adalah *supervised learning* [3], ekstraksi aspek berdasarkan frekuensi minimum [8], *sequential patterns-based rules*, dan *association mining* [9]. Penilaian pengguna dapat berupa penilaian dengan sentimen positif atau negatif [5], atau berupa *rating* dengan *range* tertentu [6]. Sebagai contoh, pada kalimat opini “lokasi hotel strategis”, aspek yang sedang dibahas adalah “lokasi hotel”, sedangkan kata opini dalam “strategis”. Kalimat opini tersebut mempunyai nilai sentimen positif, karena pengguna hotel merasa nyaman dengan lokasi hotel yang strategis.

Dalam penelitian ABSA terdapat dua aspek, yaitu aspek eksplisit dan aspek implisit [10]. Aspek eksplisit adalah aspek yang disebutkan secara langsung dalam kalimat opini. Sedangkan aspek implisit adalah aspek yang tidak disebutkan secara langsung dalam kalimat opini, tetapi hanya menyebutkan kata opini saja, sehingga harus dicari kata aspek

<sup>1,2</sup> Departemen Informatika dan Komputer, Politeknik Elektronika Negeri Surabaya, Jl. Raya ITS, Keputih, Sukolilo Surabaya 60111, Jawa Timur, INDONESIA (tlp: 031-5947280; fax: 031-5946114; email: yuliana@pens.ac.id, fitri@pens.ac.id)

<sup>3</sup> Departemen Ilmu Komputer, Fakultas Matematika dan IPA, Universitas Padjadjaran, Jalan Raya Bandung-Sumedang Km. 21 Jatinangor Sumedang 45363, Jawa Barat, INDONESIA (tlp: 022-7797712; fax: 022-7794545; email: helen@unpad.ac.id)

yang sesuai dengan kata opini tersebut. Sebagai contoh adalah kalimat opini yang mengandung aspek implisit “Nyaman, sarapan pagi lezat”. Kalimat opini tersebut mempunyai aspek eksplisit dan aspek implisit. Kalimat opini yang mengandung aspek eksplisit adalah “Sarapan pagi lezat” dan kalimat opini yang mengandung aspek implisit adalah “nyaman”. Kalimat opini dengan aspek implisit hanya berisi kata opini “nyaman”, tetapi tidak menyebutkan aspek yang dibahas. Berkaitan dengan kata opini “nyaman”, pada umumnya pengguna membahas mengenai aspek “hotel”.

Pada penelitian mengenai aspek implisit pada umumnya terdapat dua fase, yaitu fase pelatihan untuk mengetahui relasi antar kata opini dengan aspek dan fase pengujian untuk mencari aspek implisit yang tepat untuk kata opini pada kalimat opini. Pada fase pelatihan dalam penelitian aspek implisit sebelumnya, dilakukan proses pelatihan pada kalimat opini dengan aspek eksplisit dengan memberikan label pada aspek dan kata opini [11]. Penelitian lain dengan menemukan aspek implisit yang tepat tanpa *pre-knowledge* dengan aspek implisit yang potensial diberikan skor berdasarkan frekuensi kemunculannya, bersama dengan kata-kata dari sebuah kalimat. Skor dengan *threshold* tertentu merupakan aspek implisit [12]. Penelitian ini perlu melakukan pelatihan dengan memberikan label secara manual pada data. Proses ini membutuhkan waktu. Pada penelitian lainnya, dengan membangun *knowledge* berisi relasi antara kata opini dan aspek dengan *co-occurrence association rule mining*, *knowledge* tersebut yang digunakan untuk mendapatkan aspek implisit [13]. Ekstraksi *knowledge* berbasis *rule* dengan mendapatkan pasangan aspek dan kata opini menggunakan *rule*, sehingga proses mendapatkan *knowledge* dilakukan secara otomatis dari *dataset*. Penelitian ini menggunakan metode *co-occurrence association rule mining* untuk menentukan aspek implisit.

Pada aspek implisit di bahasa Indonesia, sebuah kata opini dapat digunakan untuk beberapa aspek yang berbeda. Permasalahannya adalah cara menentukan aspek yang tepat untuk kata opini tersebut. Sebagai contoh, kata opini “memuaskan” pada kalimat opini “pelayanan sangat memuaskan” dengan aspek “pelayanan”. Dapat juga pada kalimat opini “sarapan pagi memuaskan”, dengan aspek “sarapan pagi”. Jika terdapat kata opini “memuaskan”, maka aspek yang digunakan bisa “pelayanan” atau “sarapan pagi”. Contoh lainnya yaitu kata opini “nyaman” pada kalimat opini “hotel nyaman” dengan aspek “hotel”, atau dapat juga pada kalimat opini “kamar nyaman” dengan aspek “kamar”. Jika terdapat kata opini “nyaman”, maka aspek yang digunakan bisa “hotel” atau “nyaman”. Selain itu, sebuah aspek dapat mempunyai beberapa kata opini yang berbeda. Contohnya aspek “menu makanan” dapat mempunyai kata opini “lengkap”, “lezat”, atau “bervariasi”. Penentuan aspek implisit yang tepat dilakukan dengan mencari kata opini yang mempunyai relasi kuat terhadap aspek tersebut. Atau dengan kata lain, kata opini yang sering digunakan dengan aspek tersebut. Sebagai contoh, kata opini “lezat” mempunyai relasi kuat dengan aspek “menu makanan”, sedangkan kata opini “memuaskan” mempunyai relasi kuat dengan aspek

“pelayanan”. Dari ulasan pengguna, dapat diketahui *rule* berupa kata opini B sering muncul bersamaan dengan aspek A, sehingga mempunyai frekuensi tinggi. Dapat disimpulkan bahwa kata opini B dengan aspek A mempunyai relasi yang kuat karena mempunyai frekuensi yang tinggi. Namun, terdapat kata opini C yang tidak sering dipakai. Akan tetapi, kata opini B dan C mempunyai makna yang sama.

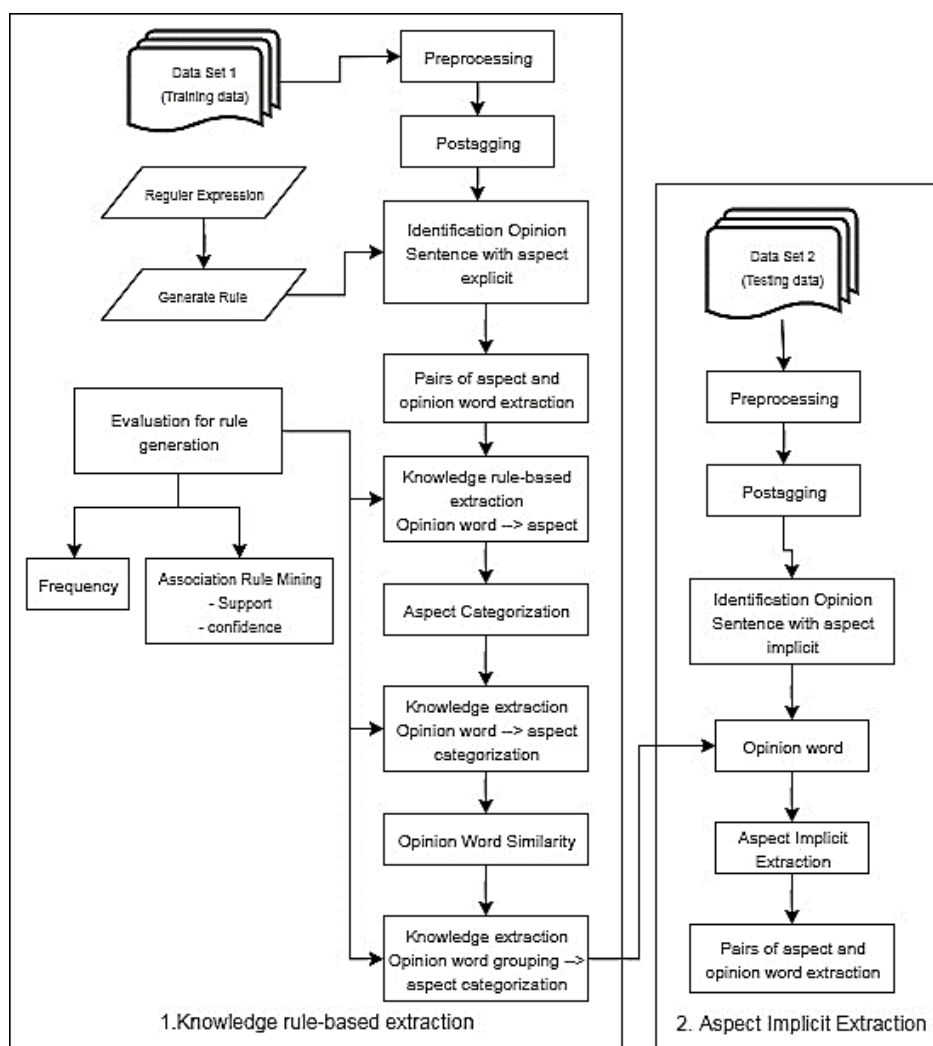
Jika dilakukan pemangkasan pada *rule* tersebut, maka kata opini C akan dipangkas karena mempunyai frekuensi rendah. Padahal jika kata opini B dan C digabungkan akan dihasilkan frekuensi yang lebih tinggi dibandingkan *rule* dengan kata opini B saja. Sebagai contoh, kata opini “ramah” sering muncul bersama dengan aspek “pelayanan”, sehingga mempunyai frekuensi tinggi dan membentuk *rule* 1. Sedangkan kata opini “bersahabat” jarang muncul bersama dengan aspek “pelayanan”, sehingga mempunyai frekuensi rendah, dan membentuk *rule* 2. Jika dilakukan pemangkasan *rule* dengan frekuensi minimum tertentu, maka *rule* 2 akan dipangkas dari *rule*, karena tidak memenuhi frekuensi minimum. Padahal kata opini “bersahabat” mempunyai kedekatan makna dengan kata opini “ramah” atau mempunyai *Opinion Word Similarity* (OWS) tinggi, sehingga *rule* 1 dapat digabungkan dengan *rule* 2, untuk memperkuat relasi *rule* tersebut guna mendapatkan aspek implisit.

Pada penelitian sebelumnya, untuk mencari aspek implisit yang tepat dari kata opini, kata opini tersebut akan dicari pada *knowledge* berbasis *rule*, pada kategori aspek yang sesuai. *Rule* yang dijalankan adalah *rule* dengan kata opini yang mempunyai kategori aspek dengan frekuensi yang paling tinggi. Namun, jika tidak ada *rule* dengan kata opini tersebut, maka tidak ada kategori aspek yang ditemukan atau dengan kata lain tidak ada aspek implisit yang ditemukan [13]. Penelitian ini mengajukan pengembangan metode yaitu dengan mencari kata opini B yang berdekatan makna dengan kata opini A dan mempunyai *rule* dengan kata opini B pada *knowledge*. Maka, *rule* dengan kata opini B dapat ditentukan aspek implisit yang tepat.

Pada makalah ini diajukan pendekatan untuk menentukan aspek implisit pada kalimat opini. Pendekatan ini mengembangkan *knowledge* berbasis *rule* berupa relasi antara kata opini dengan kategori aspek, dengan melakukan penggabungan *rule* berdasarkan kedekatan makna kata opini atau OWS. Evaluasi untuk ekstraksi *knowledge* berbasis *rule* berupa nilai *threshold* dari frekuensi dan *confidence* untuk menghasilkan nilai *precision*, *recall*, dan *f-measure* terbaik.

## II. USULAN METODE

Pada makalah ini diajukan pendekatan untuk menentukan aspek implisit pada kalimat opini menggunakan *knowledge* berbasis *rule* berupa relasi antara kata opini dengan kategori aspek. Terdapat dua fase, yaitu fase pelatihan atau proses untuk mengekstraksi *knowledge* berbasis *rule* pada ulasan pengguna, dengan mengidentifikasi kalimat opini dengan aspek eksplisit dan mendapatkan pasangan aspek dan kata opini, dan fase pengujian, yaitu fase untuk mendapatkan aspek implisit pada kalimat opini dengan mengacu ke *knowledge* berbasis *rule*, seperti ditunjukkan pada Gbr. 1.



Gbr. 1 Perancangan sistem.

**A. Identifikasi Kalimat Opini dengan Aspek Eksplisit dan Mendapatkan Pasangan Aspek dan Kata Opini**

Sebuah ulasan pengguna terdiri atas satu atau lebih kalimat opini yang dipisahkan dengan tanda “.”. Kalimat opini merupakan kalimat opini dengan aspek eksplisit dan aspek implisit. Untuk mengekstraksi *knowledge* berbasis *rule* diperlukan data ulasan pengguna berisi kalimat opini dengan aspek eksplisit, yang berisi pasangan aspek dan kata opini. Untuk menentukan sebuah kalimat opini merupakan kalimat opini dengan aspek eksplisit, *rule* yang dibangkitkan dari *regular expression* perlu dipenuhi. *Rule* bertujuan untuk mendapatkan pasangan aspek dan kata opini. Terdapat dua tipe pasangan, yaitu tipe 1, yang diawali dengan satu atau lebih aspek dan diikuti dengan satu atau lebih kata opini, dan tipe 2, yang diawali dengan satu atau lebih kata opini dan diakhiri dengan satu atau lebih aspek. Jika *rule* tersebut terpenuhi, maka kalimat termasuk kalimat opini dengan aspek eksplisit. Proses berikutnya yaitu menentukan aspek yang valid, yang didapatkan dengan melakukan kategori aspek. Jika dapat dilakukan kategori aspek, maka aspek tersebut

merupakan aspek valid, sedangkan jika tidak dapat dilakukan kategori aspek, maka aspek tersebut bukan aspek yang valid.

**B. Kategori Aspek**

Kategori aspek digunakan untuk mengelompokkan aspek berdasarkan kategori aspek. Pada makalah ini, digunakan tujuh kategori aspek, yaitu F&B, *room*, *facility*, *surrounding*, *location*, *service*, dan *guest perspective* [1], [14]. Pada setiap kategori aspek terdapat *initial seed aspect* yang ditentukan secara manual sebagai indikator yang membedakan antar kategori aspek. Sebagai contoh, berikut adalah *initial seed aspect* untuk (1) F&B: kuliner, makanan, minuman, sarapan, dan masakan; (2) *room*: kamar, kunci, toilet, TV, debu, handuk, karpet, kamar mandi, tempat tidur, dan lantai; (3) *facility*: kolam renang, travel, lobi, *lift*, parkir, restoran, Wi-Fi, dan internet; (4) *surrounding*: wisata, malioboro, stasiun, beringharjo, liburan, kuliner, perbelanjaan, dan keraton; (5) *location*: akses, letak, lokasi, hotel, tempat, posisi, dan kawasan; (6) *service*: staf, pelayanan, servis, penerima tamu, karyawan, teknisi, komplain, dan satpam; dan (7) *guest*

*perspective*: harga, perabot, suasana, *view*, model, kondisi, bangunan, gaya, manajemen, nuansa, dan kualitas.

Setelah membuat *initial seed aspect* untuk kategori aspek, selanjutnya setiap kata diperluas makna katanya menggunakan *word vector similarity* [5]. Terdapat kategori aspek  $ac_1, ac_2, ac_3, \dots, ac_n$ . Kategori aspek  $ac_1$  mempunyai *initial seed aspect*  $a_{11}, a_{12}, \dots, a_{1n}$ . Setiap *initial seed aspect* diperluas makna katanya sebanyak lima kata, dengan mencari kedekatan kata tersebut dengan kata lainnya menggunakan Word2vec dengan data wikipedia.  $TopNSim(a_{ij}, n, V)$  adalah fungsi untuk mendapatkan satu set  $N$  kata yang sama dengan  $a_{ij}$ .  $V = \{v_1, v_2, v_3, \dots, v|V|\}$  adalah daftar *vocabulary*.

Kandidat aspek dicocokkan dengan kata-kata (*words*) di kategori aspek. Jika kategori aspek terpenuhi, maka kandidat aspek tersebut merupakan aspek yang valid. Jika tidak, maka kandidat aspek tersebut bukan merupakan aspek yang valid.

C. Ekstraksi Knowledge dengan Metode Co-occurrence Berbasis Frekuensi minimum

*Knowledge* adalah pengetahuan yang dapat digunakan sebagai acuan dalam menyelesaikan suatu permasalahan. Dalam makalah ini, *knowledge* digunakan untuk mendapatkan aspek implisit dari kata opini pada kalimat opini. Untuk membangun *knowledge* terdapat dua fase. Fase pertama adalah membangun *knowledge* berbasis *rule* kata opini  $\rightarrow$  aspek, dengan melakukan *training* pada *dataset* kalimat opini dengan aspek eksplisit menggunakan matriks *co-occurrence*. Evaluasi ekstraksi *knowledge* dilakukan menggunakan frekuensi dan *support* dan *confidence* pada *association rule mining*. Pada fase kedua dilakukan pengelompokan aspek untuk menghasilkan *rule* yang lebih bagus untuk setiap kata opini. Bentuk *rule*-nya adalah kata opini  $\rightarrow$  kategori aspek.

Pada fase pertama, diperlukan ulasan pengguna yang berisi pasangan aspek dan kata opini sebagai *input* matriks *co-occurrence*. Matriks *co-occurrence* adalah frekuensi kejadian sebuah kata opini muncul bersama-sama dengan kata aspek dalam sebuah kalimat opini. Sebuah baris pada matriks *co-occurrence* menyatakan sebuah opini dan sebuah kolom menyatakan sebuah aspek. Sebagai contoh, pada kalimat opini “menu makanan lezat”, maka pada baris opini “lezat” dan kolom aspek “menu makanan” diberikan nilai 1. Proses ini dilakukan berdasarkan jumlah data kalimat opini pada data pelatihan. Evaluasi *knowledge* yang dibahas adalah frekuensi, sehingga *rule* pada *knowledge* harus memenuhi frekuensi minimum yang ditentukan. Contoh *rule* berbasis frekuensi yang dihasilkan berupa “lezat  $\rightarrow$  menu makanan,  $f=10$ ”. Maksud dari *rule* tersebut adalah dalam kalimat opini terdapat sepuluh pasangan aspek “menu makanan” dan kata opini “lezat”. Semakin tinggi frekuensi dari suatu *rule* menunjukkan bahwa kata opini mempunyai relasi kuat dengan aspek tersebut, dan sebaliknya, sehingga menjadi prioritas ketika dilakukan pencarian aspek implisit yang tepat berdasarkan kata opini.

Pada fase kedua, *rule* dengan aspek yang mempunyai kategori aspek yang sama digabungkan dan hasil frekuensi dari kedua *rule* tersebut digabungkan. Proses ini bertujuan untuk menghasilkan *rule* yang lebih bagus. Contohnya, pada

TABEL I  
PEMBANGUNAN *KNOWLEDGE* KATA OPINI  $\rightarrow$  KATEGORI ASPEK

Rule Kata Opini $\rightarrow$ Aspek	Kategori Aspek	Rule Opini $\rightarrow$ Kategori Aspek	Pengembangan Rule
$o_1 \rightarrow a_{11}, f=10$	$ac_1 = \{a_{11}, a_{12}\}$	$o_1 \rightarrow ac_1, f=10$	$o_1 \rightarrow ac_1, f=10$
$o_2 \rightarrow a_{21}, f=8$	$ac_2 = \{a_{21}\}$	$o_2 \rightarrow ac_2, f=8,$ $ac_3, f=2$	$o_2 \rightarrow ac_2, f=8, ac_3$ $f=2$
$o_2 \rightarrow a_{31}, f=2$	$ac_3 = \{a_{31}, a_{32}\}$	$o_3 \rightarrow ac_1, f=4,$ $ac_3, f=10$	$o_3 \rightarrow ac_1$ $f=4, ac_3, f=10$
$o_3 \rightarrow a_{12}, f=4$	$ac_4 = \{a_{41}, a_{42}\}$	$o_4 \rightarrow ac_4, f=6$	$\{o_4, o_5\} \rightarrow ac_4$ $f=9$
$o_3 \rightarrow a_{32}, f=10$		$o_5 \rightarrow ac_4, f=3$	
$o_4 \rightarrow a_{41}, f=6$			
$o_5 \rightarrow a_{42}, f=3$			

aspek “sarapan pagi” dan “menu makanan” dilakukan kategorisasi aspek menjadi “*food and beverage*”, sehingga *robust rule* yang dihasilkan berupa “lezat”  $\rightarrow$  “*food and beverage*”.

Dalam makalah ini terdapat tujuh kategorisasi aspek  $ac_1, ac_2, ac_3, \dots, ac_7$ . Kategorisasi aspek  $ac_1$  mempunyai aspek  $a_{11}, a_{12}, \dots, a_{1n}$  dan kategorisasi aspek  $ac_2$  mempunyai aspek  $a_{21}, a_{22}, \dots, a_{2m}$ . Terdapat daftar kata opini  $O$ , yaitu  $o_1, o_2, o_3, \dots, o_n$ . Sebagai contoh, terdapat *rule* kata opini  $\rightarrow$  kata aspek dengan frekuensi tertentu, seperti ditunjukkan pada Tabel I. Sebuah opini dimungkinkan mempunyai relasi dengan beberapa kategori aspek yang berbeda. Sebagai contoh, *rule*  $o_2 \rightarrow a_{21}$  dengan frekuensi 8 dan *rule*  $o_2 \rightarrow a_{31}$  dengan frekuensi 2. Berdasarkan kategori aspek, maka  $a_{21}$  merupakan kategori aspek  $ac_2$ , sedangkan  $a_{31}$  merupakan kategori aspek  $ac_3$ . Kata opini  $o_2$  digunakan pada kategori aspek  $ac_2$ , dengan frekuensi 8 dan kategori aspek  $ac_3$  dengan frekuensi 2. Sebuah kata opini dapat mempunyai relasi dengan beberapa kategori aspek yang berbeda.

$$AC = \{ac_1, ac_2, ac_3, \dots, ac_r\}$$

$$ac_1 = \{a_{11}, a_{12}, a_{13}, \dots, a_{1n}\}$$

$$ac_2 = \{a_{21}, a_{22}, a_{23}, \dots, a_{2m}\}$$

$$O = \{o_1, o_2, o_3, \dots, o_k\}$$

Untuk menghasilkan *knowledge* berbasis yang bagus, *rule* harus memenuhi frekuensi minimum. *Rule* yang tidak memenuhi frekuensi harus dipangkas. Proses pemangkasan *rule* dapat menemukan kata opini yang mempunyai relasi rendah terhadap kategori aspek, menunjukkan penggunaan kata opini yang tidak tepat atau jarang digunakan pada aspek tersebut. Sebagai contoh adalah kalimat opini pengguna “Petugas ramah dan bersih”. Kata opini “ramah” sering digunakan dengan aspek “petugas”, *rule* yang dibentuk berupa “ramah  $\rightarrow$  petugas”, sehingga kemungkinan mempunyai frekuensi tinggi. Sedangkan kata opini “bersih” jarang digunakan dengan aspek “petugas”, *rule* yang dibentuk berupa “bersih  $\rightarrow$  petugas”, sehingga kemungkinan mempunyai

frekuensi rendah dan tidak memenuhi frekuensi minimum. Maka, perlu dilakukan pemangkasan pada *rule* tersebut.

*D. Ekstraksi Knowledge dengan Metode Co-occurrence Association Rule Mining*

Sebuah penelitian mengenai ekstraksi *knowledge* dengan metode *Co-occurrence Association Rule Mining* (coAR), mencari relasi antara kata opini dengan aspek dengan melakukan evaluasi berdasarkan *minimum support* dan *confidence* [13]. Persamaan (1) digunakan untuk menghitung nilai *support*, sedangkan nilai *confidence* dihitung menggunakan (2).

$$support_{ij} = \frac{\#(o_i \rightarrow ac_j)}{\#o_i} \tag{1}$$

$$confidence_{ij} = \frac{\#(o_i \rightarrow ac_j)}{N} \tag{2}$$

dengan  $\#(o_i \rightarrow ac_j)$  merupakan jumlah *co-occurrence* relasi antara kata opini  $o_i$  dengan kategori aspek  $ac_j$ .  $o_i$  adalah jumlah kata opini pada kalimat opini dan  $N$  menyatakan jumlah ulasan pengguna untuk membangun *knowledge* berbasis *rule*.

*E. Pengembangan Knowledge Berbasis Rule dengan Menggabungkan Rule Berdasarkan Opinion Word Similarity*

Pada *knowledge* berbasis *rule* berupa kata opini  $\rightarrow$  kategori aspek, dimungkinkan kata opini mempunyai kedekatan dengan kata opini lainnya. Untuk mengetahui tingkat kedekatan antara dua buah kata opini, digunakan Word2vec dengan data dari Wikipedia dan Kamus Besar Bahasa Indonesia (KBBI) Kemendikbud. Tingkat kedekatan dua kata opini dengan menghitung nilai similaritas disebut OWS menggunakan *cosine similarity* antar dua kata tersebut. Semakin tinggi nilai similaritas mengandung arti dua buah kata mempunyai makna yang berdekatan.  $TopNSim(a_{ij}, n, V)$  adalah fungsi untuk mendapatkan himpunan  $N$  kata yang sama dengan  $a_{ij}$ .  $V = \{v_1, v_2, v_3, \dots, v|V|\}$  adalah himpunan semua kata opini.

Pada Tabel I terlihat *rule*  $o_4 \rightarrow ac_4$  dengan frekuensi 6 dan  $o_5 \rightarrow ac_4$  dengan frekuensi 3, kata opini  $o_4$  dan  $o_5$ , mempunyai nilai similaritas yang tinggi, sehingga dapat digabungkan menjadi satu *rule* dan menjadi opini  $\{o_4, o_5\} \rightarrow ac_4$ , dengan frekuensi 9. Nilai frekuensi diperoleh dengan menggabungkan frekuensi kedua *rule*, yaitu 6 dan 3, dijumlahkan menjadi 9.

*F. Identifikasi Kalimat Opini dengan Aspek Implisit*

Aspek adalah topik yang sedang dibahas dalam kalimat opini. Aspek harus dinyatakan dengan kata yang bersifat spesifik daripada kata yang bersifat umum pada domain tertentu. Aspek dalam kalimat opini adalah kata benda atau frasa kata benda. Menurut KBBI, frasa kata benda merupakan gabungan dua atau lebih kata benda. Kalimat opini dengan aspek implisit adalah kalimat opini yang hanya berisi satu atau lebih kata opini saja, tanpa mempunyai aspek. Kata opini dalam bahasa Indonesia dapat berupa kata kerja (*verb*/VB) atau kata sifat (*adjective*/JJ).

TABEL II  
DISTRIBUSI DATASET 1 UNTUK MENGEKSTRAKSI KNOWLEDGE

No	Kategori aspek	Aspek	Kategori aspek		
			Negatif	Positif	Total
1	Location	16	18	223	241(24,47%)
2	Service	36	20	201	221(22,44%)
3	Room	52	92	118	210(21,32%)
4	F & B	25	20	130	150(15,23%)
5	Facility	24	17	45	62(6,3%)
6	Guest perspective	20	13	28	41(4,16%)
7	Surrounding	28	5	55	60(6,09%)
	Jumlah		185 201	800 (81,22%)	985 (100%)

Beberapa permasalahan dalam mengidentifikasi kalimat opini terjadi jika kalimat opini mengandung kalimat opini dengan aspek eksplisit maupun dengan aspek implisit. Sebagai contoh adalah kalimat opini “Bersih, dekat dengan jalan Malioboro”. Pada kalimat opini tersebut terdapat kalimat opini dengan aspek eksplisit “dekat dengan jalan Malioboro” dan kalimat opini dengan aspek implisit “bersih”. Akan tetapi, dengan *regular expression* dianggap terdapat satu kalimat opini dengan aspek eksplisit, yaitu aspek “jalan Malioboro” dengan kata opini “nyaman” dan “bersih”. Contoh lain adalah kalimat opini “Menu lengkap dan pasti lezat, mantap sekali”. Dengan *regular expression* kalimat tersebut diidentifikasi hanya sebagai kalimat opini dengan aspek eksplisit, padahal kalimat opini tersebut mengandung kalimat opini dengan aspek eksplisit “Menu lengkap dan pasti lezat” dan kalimat opini dengan aspek implisit “mantap sekali”. Oleh karena itu, untuk kalimat opini dibuat menjadi dua kalimat opini, yaitu kalimat opini dengan aspek eksplisit atau implisit, sehingga menjadi “Menu lengkap dan pasti lezat. Mantap sekali”.

*G. Menentukan Aspek Implisit dari Kalimat Opini*

Penentuan aspek implisit dari kalimat opini dilakukan dengan mencari *rule* yang tepat, dengan frekuensi terbesar, sebagai perwakilan dari kategori aspek dari kata opini tersebut. Sebagai contoh, pada kalimat opini terdapat kata opini  $o_3$ , dicari aspek implisit yang tepat untuk kata opini tersebut. Kata opini  $o_3$  mempunyai relasi dengan  $ac_1$  dengan frekuensi 4, juga mempunyai relasi dengan  $ac_3$  dengan frekuensi 10. Bentuk *rule*-nya yaitu  $o_3 \rightarrow ac_1 f = 4$ ,  $o_3 \rightarrow ac_3 f = 10$ . Frekuensi yang paling besar dimiliki oleh kategori aspek  $ac_3$ , sehingga kata opini  $o_3$  mempunyai relasi dengan kategori aspek  $ac_3$  yang mempunyai aspek  $a_{31}, a_{32}$ .

III. HASIL DAN ANALISIS

Metode yang diusulkan untuk ekstraksi aspek selanjutnya diuji coba dan dievaluasi.

*A. Persiapan Dataset*

*Dataset 1* pada Tabel II berupa ulasan pengguna diambil dari hotel di Yogyakarta dari *website* [www.traveloka.com](http://www.traveloka.com), berjumlah 307 buah. Ulasan pengguna diambil pada tahun

TABEL III  
DISTRIBUSI *DATASET 2* UNTUK MENGEKSTRAKSI ASPEK IMPLISIT PADA KALIMAT OPINI

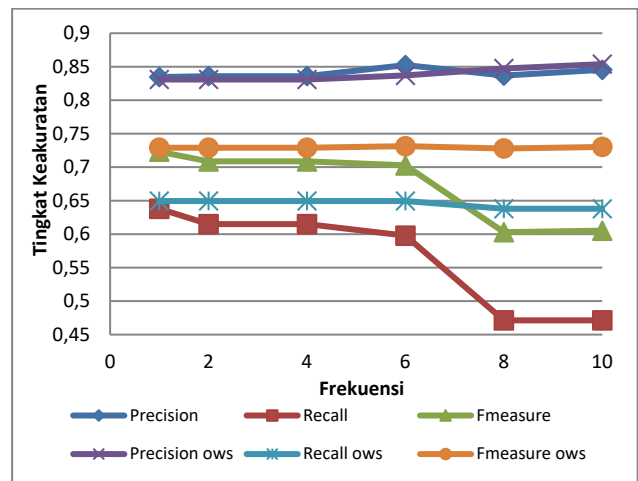
No	Kategori aspek	Aspek	Kata Opini	Kategori aspek		
				Negatif	Positif	Total
1	Location	3	16	1	120	121 (69,54%)
2	Service	2	9	0	28	28 (16,09%)
3	Room	1	6	0	22	22 (12,64%)
4	F & B	1	1	0	2	2 (1,15%)
5	Guest perspective	1	1	0	1	1 (0,57%)
	Jumlah	8	33	1 (0,57%)	173 (99,43%)	174 (100%)

2018 dan tahun 2019. Untuk membangun *knowledge* berisi 307 ulasan pengguna, selanjutnya data dipisahkan menjadi kalimat-kalimat opini berdasarkan tanda baca “.”, sehingga terdapat 627 kalimat opini. Dalam *dataset* terdapat 443 kalimat opini atau 70,65% yang mengandung pasangan aspek dan kata opini.

Setelah membangun *knowledge* dari *dataset 1*, *knowledge* digunakan untuk menentukan aspek implisit dari kalimat opini menggunakan *dataset 2* pada Tabel III. *Dataset 2* berisi kalimat opini dengan aspek implisit. Dari 1.092 ulasan pengguna, hanya diperoleh 118 ulasan pengguna dengan aspek implisit, yang selanjutnya dipisahkan menjadi kalimat-kalimat opini berdasarkan tanda baca “.”. Terdapat 174 kalimat opini dengan aspek implisit. Kata opini yang terdapat pada kalimat opini dengan aspek implisit tidak terlalu bervariasi pada bahasa Indonesia. Contohnya, pada kategorisasi aspek “*location*”, kata opini yang digunakan hanya sebanyak enam belas buah, yaitu kata opini bagus, bersih, nyaman, strategis, baik, enak, mantap, memuaskan, menyenangkan, lega, oke, pas, tenang, lumayan, puas, dan tua. Sedangkan pada kategorisasi aspek “*service*”, kata opini yang digunakan hanya sebanyak sembilan buah, yaitu kata opini puas, bagus, baik, ramah, keren, memuaskan, mantap, menyenangkan, dan oke.

### B. Evaluasi Knowledge berbasis Rule

Dari *dataset* pada Tabel II, dihasilkan *knowledge* berupa *rule* sebanyak 202 buah, sedangkan jika dilakukan penggabungan *rule* berdasarkan OWS, *rule* menjadi 137 buah. Jumlah *rule* pada *knowledge* dengan OWS berkurang, sehingga menjadi 67,82%. Pada makalah ini dikembangkan *knowledge* berbasis *rule* berupa relasi antara kata opini dengan kategori aspek, dengan melakukan penggabungan *rule* berdasarkan kedekatan makna kata opini atau OWS. Evaluasi untuk ekstraksi *knowledge* berbasis *rule* berupa nilai *threshold* dari frekuensi dan *confidence* untuk menghasilkan nilai *precision*, *recall*, dan *f-measure* terbaik. Pada penelitian dibandingkan metode ekstraksi *knowledge* dengan OWS dengan evaluasi frekuensi dan *confidence* dengan metode ekstraksi *knowledge* tanpa OWS. Metode perbandingan



Gbr. 2 Perbandingan hasil evaluasi *knowledge* berbasis rule tanpa dan menggunakan OWS berdasarkan frekuensi minimum.

ekstraksi *knowledge* tanpa OWS dengan evaluasi *confidence* merupakan metode coAR [13].

Hasil *rule* ditunjukkan pada Tabel IV dan Tabel V, berupa lima *rule* kata opini → kategori aspek dengan nilai frekuensi tertinggi untuk setiap kategori aspek. Sebagai contoh, pada Tabel IV, lima *rule* tertinggi pada kategori “*location*” adalah strategis → *location* dengan frekuensi 85, nyaman → *location* dengan frekuensi 27, dekat → *location* dengan frekuensi 19, bersih → *location* dengan frekuensi 18, dan bagus → *location* dengan frekuensi 14. Dari *rule* tersebut, kata opini strategis sering dipakai untuk kategori aspek “*location*”.

Pada Tabel IV terdapat *rule* ramah → *service* dengan frekuensi 97, pada Tabel V *rule* menjadi [ramah = 97, bersahabat = 1, sopan = 3, ramah-ramah = 2, santun = 1] → *service* dengan frekuensi 104. Kata opini ramah mempunyai kedekatan dengan kata opini bersahabat, sopan, ramah-ramah dan santun, sehingga *rule* dengan kata opini tersebut digabungkan. Pada ekstraksi *knowledge* dengan frekuensi, jika frekuensi minimum sebesar 2, maka *rule* yang mempunyai frekuensi 1 dipangkas dari *knowledge*, sehingga dimungkinkan kehilangan banyak *rule* dan hal ini berpengaruh terhadap identifikasi aspek implisit pada kalimat opini. Pada penggabungan *rule* berdasarkan OWS, *rule* dengan kata opini di bawah frekuensi minimum tidak jadi dipangkas, karena dapat digabungkan dengan *rule* yang lain dan memperkuat *rule* tersebut.

Hasil evaluasi *knowledge* berbasis *rule* berdasarkan frekuensi minimum tanpa OWS ditunjukkan pada Tabel VI. *Rule* tersebut selanjutnya digunakan untuk menentukan aspek implisit dalam kalimat opini. Dari hasil uji coba didapatkan nilai *threshold* dengan frekuensi 1, dengan *precision* 0,83, *recall* 0,64, dan *f-measure* 0,72. *Rule* dengan frekuensi 2, 4, 6, 8, dan 10 mengalami penurunan nilai *recall* dan *f-measure*. Hasil evaluasi *knowledge* berbasis *rule* menggunakan OWS berdasarkan frekuensi minimum. Dari hasil uji coba didapatkan nilai *threshold* dengan frekuensi 6, dengan *precision* 0,84, *recall* 0,65, dan *f-measure* 0,73. *Rule* dengan frekuensi 8 dan 10 mengalami penurunan nilai *recall* dan *f-measure*. Gbr. 2 menunjukkan perbandingan evaluasi *knowledge* tanpa dan

TABEL IV  
LIMA RULE KATA OPINI → KATEGORI ASPEK TANPA OWS

Kategori Aspek						
Location	Service	Room	F&B	Facility	Guest Perspective	Surrounding
strategis, $f = 85$	ramah, $f = 97$	bersih, $f = 36$	enak, $f = 32$	ada, $f = 8$	lama, $f = 7$	dekat, $f = 17$
nyaman = 27	baik, $f = 21$	luas, $f = 25$	banyak, $f = 21$	bersih, $f = 6$	tua, $f = 7$	ramah, menyenangkan, enak $f = 6$
dekat = 19	memuaskan, $f = 17$	nyaman, $f = 14$	bervariasi, $f = 19$	luas, $f = 6$	nyaman, $f = 4$	banyak, $f = 5$
bersih = 18	bagus, $f = 16$	dingin, $f = 11$	lezat, $f = 12$	bagus, $f = 4$	terjangkau, menyenangkan, memuaskan, $f = 3$	mudah, $f = 4$
bagus = 14	cepat, $f = 10$	ada, $f = 9$	lengkap, $f = 10$	nyaman, oke, besar, representatif, lengkap, $f = 3$		puas, luas, strategis, $f = 3$

TABEL V  
LIMA RULE KATA OPINI → KATEGORI ASPEK DENGAN OWS

Kategori Aspek						
Location	Service	Room	F&B	Facility	Guest Perspective	Surrounding
strategis = 85, $f = 85$	ramah = 97, bersahabat = 1, sopan = 3, ramah-ramah = 2, santun = 1, $f = 104$	bersih = 36, rapi = 2, $f = 38$	lengkap=10, banyak=21, beragam=3, bervariasi=19, variatif=3, $f=59$	besar = 3, luas = 6, $f = 9$	tua = 7, lama = 7, $f = 14$	dekat = 17, $f = 17$
ramah = 6, menyenangkan = 8, tenang = 1, nyaman = 27, $f = 42$	keren = 1, bagus = 16, baik=21, $f = 38$	besar = 4, luas = 25, $f = 29$	enak-enak=1, lezat=12, enak=32, berasa=1, $f=46$	ada = 8, $f = 8$	berkesan = 1, menyenangkan = 3, tenang = 1, nyaman = 4, $f = 9$	ramah = 6, menyenangkan = 6, $f = 12$
nyaman = 27, $f = 27$	memuaskan = 17, puas = 1, $f = 18$	tenang = 2, nyaman = 14, $f = 16$	memuaskan=5, $f=5$	teratur = 2, bersih = 6, $f = 8$	memuaskan = 3, bagus = 2, $f = 5$	layak = 1, pas = 2, cocok = 2, sesuai = 1, nyaman = 2, $f = 8$
terbaik = 3, memuaskan = 3, bagus = 14, mengesankan = 1, baik = 3, $f = 24$	cepat = 10, mudah = 1, tanggap = 2, responsif = 1, $f = 14$	bagus = 7, baik = 5, $f = 12$	mantap=1, oke=4, $f=5$	terbaik = 1, bagus = 4, baik = 1, $f = 6$	mahal = 1, terjangkau = 3, $f = 4$	enak = 6, $f = 6$
tepat = 6, layak = 2, pas = 6, cocok = 5, sesuai = 1, $f = 20$	mantap = 2, oke = 5, $f = 7$	dingin = 11, $f = 11$	cocok=2, sesuai=1, $f=3$	ramah = 1, menyenangkan = 1, nyaman = 3, $f = 5$	bersih = 2, $f = 2$	banyak = 5, $f = 5$

menggunakan OWS berdasarkan frekuensi minimum. Dari grafik hasil evaluasi *knowledge* berbasis *rule* dengan frekuensi menggunakan OWS, terlihat hasil yang lebih bagus dibandingkan dengan *knowledge* tanpa menggunakan OWS, nilai *precision* meningkat sebesar 0,25%, *recall* meningkat sebesar 1,15%, dan *precision* meningkat sebesar 0,83%.

Hasil evaluasi *knowledge* berbasis *rule* berdasarkan *confidence* minimum ditunjukkan pada Tabel VII. Dari hasil uji coba didapatkan nilai *threshold* dengan nilai *confidence* 0,01, dengan *precision* 0,83, *recall* 0,64, dan *f-measure* 0,72. *Rule* dengan *confidence* dengan range 0,02 sampai dengan 0,1

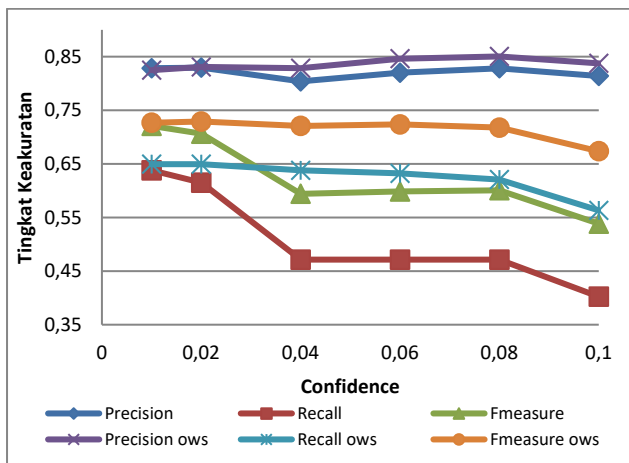
mengalami penurunan nilai *recall* dan *f-measure*. Pada evaluasi *knowledge* berbasis *rule* berdasarkan *confidence* minimum, dari hasil uji coba didapatkan nilai *threshold* dengan *confidence* 0,02, dengan *precision* 0,83, *recall* 0,65, dan *f-measure* 0,73. *Rule* dengan *confidence* dengan range 0,03 sampai dengan 0,1 mengalami penurunan nilai *precision*, *recall*, dan *f-measure*. Gbr. 3 memperlihatkan perbandingan evaluasi *knowledge* tanpa dan menggunakan OWS berdasarkan *confidence* minimum. Grafik hasil evaluasi *knowledge* berbasis *rule* menggunakan OWS menunjukkan hasil yang lebih bagus dibandingkan dengan *knowledge* tanpa

TABEL VI  
EVALUASI *KNOWLEDGE* BERBASIS *RULE* BERDASARKAN FREKUENSI TANPA OWS

Frekuensi	Precision	Recall	F-measure	Precision OWS	Recall OWS	F-measure OWS	Aspek Implisit Tidak Teridentifikasi dengan Rule	Aspek Implisit Tidak Teridentifikasi dengan Rule OWS	Jumlah Rule Setelah Pemangkas	Jumlah Rule OWS Setelah Pemangkas
1	<b>0,83</b>	<b>0,64</b>	<b>0,72</b>	0,83	0,65	0,73	<b>2,30%</b>	2,30%	<b>54,46%</b>	54,74%
2	0,84	0,61	0,71	0,83	0,65	0,73	4,02%	2,30%	38,12%	45,99%
4	0,84	0,61	0,71	0,83	0,65	0,73	4,02%	2,30%	23,27%	31,39%
6	0,85	0,60	0,70	<b>0,84</b>	<b>0,65</b>	<b>0,73</b>	12,07%	<b>2,30%</b>	13,86%	<b>23,36%</b>
8	0,84	0,47	0,60	0,85	0,64	0,73	16,67%	10,34%	10,40%	19,71%
10	0,85	0,47	0,61	0,85	0,64	0,73	17,24%	10,92%	8,91%	16,06%

TABEL VII  
EVALUASI *KNOWLEDGE* BERBASIS *RULE* BERDASARKAN *CONFIDENCE* DENGAN OWS

Confidence	Precision	Recall	F-measure	Precision OWS	Recall OWS	F-measure OWS	Aspek Implisit Tidak Teridentifikasi - Rule	Aspek Implisit Tidak Teridentifikasi - Rule OWS	Jumlah Rule Setelah Pemangkas	Jumlah Rule OWS Setelah Pemangkas
0,01	<b>0,83</b>	<b>0,64</b>	<b>0,72</b>	0,82	0,65	0,73	<b>1,72%</b>	1,72%	<b>68,32%</b>	67,88%
0,02	0,83	0,61	0,71	<b>0,83</b>	<b>0,65</b>	<b>0,73</b>	4,02%	<b>2,30%</b>	44,06%	<b>49,64%</b>
0,04	0,80	0,47	0,59	0,83	0,64	0,72	4,02%	2,30%	22,28%	32,85%
0,06	0,82	0,47	0,60	0,85	0,63	0,72	13,79%	10,34%	15,35%	22,63%
0,08	0,83	0,47	0,60	0,85	0,62	0,72	14,94%	10,34%	9,41%	16,79%
0,1	0,81	0,40	0,54	0,84	0,56	0,67	29,31%	10,34%	7,43%	13,14%



Gbr. 3 Perbandingan hasil evaluasi *knowledge* berbasis *rule* tanpa dan menggunakan OWS berdasarkan *confidence* minimum.

menggunakan OWS, yaitu terjadi peningkatan nilai *precision* sebesar 0,25%, nilai *recall* sebesar 1,15% dan nilai *f-measure* sebesar 0,83%.

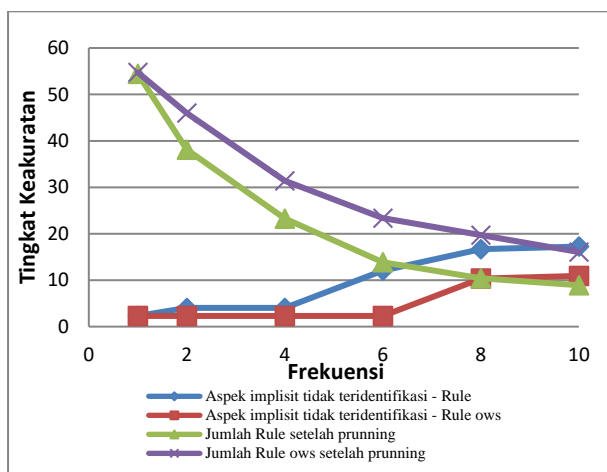
C. Evaluasi Identifikasi Aspek Implisit pada Kalimat Opini

*Knowledge* berbasis *rule* yang telah dihasilkan dari proses pelatihan selanjutnya digunakan untuk mengidentifikasi aspek implisit pada kalimat opini. Pada *knowledge* berbasis *rule*

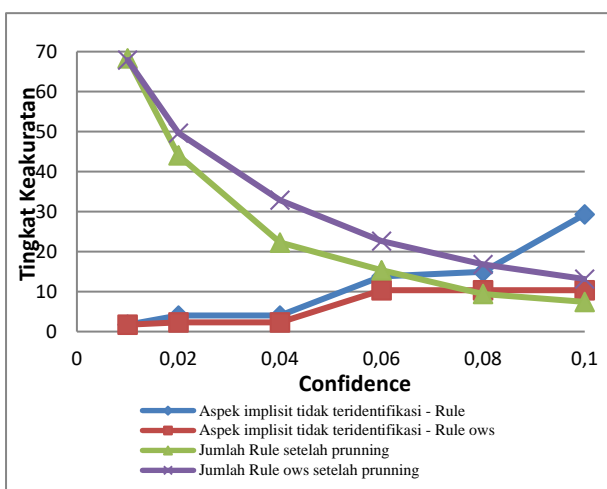
dengan *threshold* frekuensi 1 tanpa OWS, aspek implisit pada kalimat opini yang tidak terdeteksi sebanyak empat buah atau 2,3%. Sedangkan rata-rata jumlah aspek implisit pada kalimat opini yang tidak terdeteksi sebesar 9,39%. Jumlah *rule* setelah dilakukan pemangkas *rule* sebesar 54,46%. Pada *knowledge* berbasis *rule* dengan *threshold* frekuensi 6 dengan OWS, aspek implisit pada kalimat opini yang tidak terdeteksi sebanyak empat buah atau sebesar 2,3%, sedangkan rata-rata jumlah aspek implisit pada kalimat opini yang tidak terdeteksi sebesar 5,08%. Jumlah *rule* setelah dilakukan pemangkas adalah sebesar 23,36%. Gbr. 4 merupakan perbandingan hasil identifikasi aspek implisit pada kalimat opini tanpa dan dengan OWS dengan *knowledge* berbasis *rule* dengan frekuensi. Hasilnya menunjukkan bahwa aspek implisit tidak teridentifikasi dengan *rule* OWS lebih kecil dengan selisih rata-rata sebesar 4,31% dibandingkan dengan *rule* tanpa OWS. Sedangkan jumlah *rule* setelah dilakukan pemangkas *rule* dengan mengacu pada *knowledge* berbasis *rule* dengan frekuensi menggunakan OWS lebih besar, dengan selisih rata-rata sebesar 7,04%, dibandingkan dengan *rule* tanpa OWS.

Pada *knowledge* berbasis *rule* dengan *threshold confidence* 0,01 tanpa OWS, aspek implisit pada kalimat opini yang tidak terdeteksi sebanyak tiga atau sebesar 1,72%. Sedangkan rata-rata jumlah aspek implisit pada kalimat opini yang tidak terdeteksi sebesar 11,3%. Jumlah *rule* setelah dilakukan pemangkas *rule* adalah sebesar 68,32%.





Gbr. 4 Perbandingan hasil identifikasi aspek implisit pada kalimat opini dan jumlah rule setelah dilakukan pemangkasan rule dengan mengacu pada knowledge berbasis rule dengan frekuensi.



Gbr. 5 Perbandingan hasil identifikasi aspek implisit pada kalimat opini dan jumlah rule setelah dilakukan pemangkasan rule dengan mengacu pada knowledge berbasis rule dengan confidence.

Pada *knowledge* berbasis *rule* dengan *confidence* 0,02 dengan OWS, aspek implisit pada kalimat opini yang tidak terdeteksi sebanyak empat buah atau sebesar 2,3%, sedangkan rata-rata jumlah aspek implisit pada kalimat opini yang tidak terdeteksi sebesar 6,23%. Jumlah *rule* setelah dilakukan pemangkasan *rule* sebesar 49,64%. Gbr. 5 memperlihatkan perbandingan hasil identifikasi aspek implisit pada kalimat opini tanpa dan dengan OWS dengan *knowledge* berbasis *rule* dengan *confidence*. Hasilnya menunjukkan bahwa aspek implisit tidak teridentifikasi dengan *rule* OWS lebih kecil dengan rata-rata sebesar 5,08% dibandingkan dengan *rule* tanpa OWS. Sedangkan jumlah *rule* setelah dilakukan pemangkasan *rule* dengan mengacu pada *knowledge* berbasis *rule* dengan *confidence* dengan OWS lebih besar, dengan rata-rata sebesar 6,16%, dibandingkan dengan *rule* tanpa OWS.

#### IV. KESIMPULAN

Pada makalah ini diajukan pendekatan untuk menentukan aspek implisit pada kalimat opini. Pendekatan ini

dikembangkan dengan *knowledge* berbasis *rule* berupa relasi antara kata opini dengan kategori aspek, dengan melakukan penggabungan *rule* berdasarkan OWS. Hasil evaluasi *knowledge* berbasis *rule* dengan frekuensi atau *confidence* menggunakan OWS menunjukkan hasil yang lebih bagus dibandingkan dengan *knowledge* tanpa menggunakan OWS. Nilai *precision* meningkat sebesar 0,25%, *recall* meningkat sebesar 1,15%, dan *precision* meningkat sebesar 0,83%.

Hasil pengujian menunjukkan bahwa jumlah aspek implisit tidak teridentifikasi dengan *rule* OWS dengan evaluasi frekuensi lebih kecil dengan selisih rata-rata sebesar 4,31% dibandingkan dengan *rule* tanpa OWS. Sedangkan jumlah *rule* setelah dilakukan pemangkasan *rule* dengan mengacu pada *knowledge* berbasis *rule* dengan frekuensi dengan OWS lebih besar dengan selisih rata-rata sebesar 7,04% dibandingkan dengan *rule* tanpa OWS. Hasil menunjukkan bahwa jumlah aspek implisit tidak teridentifikasi dengan *rule* OWS dengan evaluasi *confidence* lebih kecil dengan rata-rata sebesar 5,08% dibandingkan dengan *rule* tanpa OWS. Sedangkan jumlah *rule* setelah dilakukan pemangkasan *rule* dengan mengacu ke *knowledge* berbasis *rule* dengan *confidence* dengan OWS lebih besar dengan rata-rata sebesar 6,16% dibandingkan dengan *rule* tanpa OWS.

#### UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Politeknik Elektronika Negeri Surabaya yang telah memberikan dana Penelitian Lokal untuk penelitian ini.

#### REFERENSI

- [1] N. Akhtar, N. Zubair, A. Kumar, dan T. Ahmad, "Aspect based Sentiment Oriented Summarization of Hotel Reviews," *Procedia Comput. Sci.*, Vol. 115, hal. 563–571, 2017.
- [2] Y. Setiowati dan F. Setyorini, "Service Extraction and Sentiment Analysis to Indicate Hotel Service Quality in Yogyakarta based on User Opinion," *International Seminar on Research of Information Technology and Intelligent Systems*, 2018, hal. 427–432.
- [3] D. Ekawati dan M.L. Khodra, "Aspect-based Sentiment Analysis for Indonesian Restaurant Reviews," *Proc. - 2017 Int. Conf. Adv. Informatics Concepts, Theory Appl. ICAICTA 2017*, 2017, hal. 1-6.
- [4] L. Zhang, B. Liu, S.H. Lim, dan E. O'Brien-Strain, "Extracting and Ranking Product Features in Opinion Documents," *The 23rd International Conference on Computational Linguistics: Posters*, 2009, hal. 1462–1470.
- [5] D.H. Sasmita, A.F. Wicaksono, S. Louvan, dan M. Adriani, "Unsupervised Aspect-based Sentiment Analysis on Indonesian Restaurant Reviews," *Proc. 2017 Int. Conf. Asian Lang. Process. IALP 2017*, 2017, hal. 383–386.
- [6] S. Gojali dan M.L. Khodra, "Aspect Based Sentiment Analysis for Review Rating Prediction," *Int. Conf. Adv. Informatics Concepts, Theory Appl. ICAICTA 2016*, 2016, hal. 1-6.
- [7] Y. Kang dan L. Zhou, "Rube: Rule-based Methods for Extracting Product Features from Online Consumer Reviews," *Inf. Manag.*, Vol. 54, No. 2, hal. 166–176, 2017.
- [8] E. Marrese-Taylor, J.D. Velásquez, dan F. Bravo-Marquez, "A Novel Deterministic Approach for Aspect-based Opinion Mining in Tourism Products Reviews," *Expert Syst. Appl.*, Vol. 41, No. 17, hal. 7764–7775, 2017.
- [9] T.A. Rana dan Y.N. Cheah, "A Two-fold Rule-based Model for Aspect Extraction," *Expert Syst. Appl.*, Vol. 89, hal. 273–285, 2017.
- [10] F. Lazhar dan T.G. Yamina, "Mining Explicit and Implicit Opinions from Reviews," *Int. J. Data Mining, Model. Manag.*, Vol. 8, No. 1, hal.

- 75, 2016.
- [11] K. Schouten dan F. Frasincar, "Implicit Feature Extraction for Sentiment Analysis in Consumer Reviews," *Int. Conf. on Appl. of Natural Language to Data Bases/Information Systems*, 2014, hal. 228–231.
- [12] K. Schouten., F. Frasincar, dan Erasmus, "Finding Implicit Features in Consumer Reviews for Sentiment Analysis," *Int. Conf. Web Eng. ICWE 2014 Web Eng.*, 2014, hal. 130–144.
- [13] Z. Hai, K. Chang, dan J.J. Kim, "Implicit Feature Identification via Co-occurrence Association Rule Mining," in *Int. Conf. on Intelligent Text Processing and Computational Linguistics*, 2011, hal. 393–404.
- [14] P. Prameswari, I. Surjandari, dan E. Laoh, "Opinion Mining from Online Reviews in Bali Tourist Area," *Proc. - 2017 3rd Int. Conf. Sci. Inf. Technol. Theory Appl. IT Educ. Ind. Soc. Big Data Era, ICSITech 2017*, 2017, hal. 226–230.