

## Strategi Penentuan Harga dan Penjualan Hotel Pada *Online Travel Agency* (OTA)

Yusma Dewi Yulianti<sup>1)</sup>, Rahmat Nurcahyo<sup>2)</sup> dan M. Dachyar<sup>2)</sup>

<sup>1)</sup>Departemen Teknik Industri, Universitas Indonesia, Salemba-Jakarta

<sup>2)</sup>Departemen Teknik Industri, Universitas Indonesia, Depok

Email: yusmadewiy@gmail.com, rahmat@eng.ui.ac.id, mdachyar@yahoo.com

### ABSTRAK

OTA (*Online Travel Agency*) telah mengubah perilaku wisatawan yang sekarang bergantung pada Internet untuk mencari informasi, merencanakan perjalanan, membeli tiket dan akomodasi perjalanannya. Penelitian ini akan berfokus pada salah satu akomodasi yang ditawarkan oleh OTA yakni Hotel. Berbagai kemudahan membuat OTA menjadi lebih unggul dibanding oleh situs web yang disediakan oleh hotel-hotel dengan didapatnya review dan rating sehingga bisa menjadi rekomendasi bagi wisatawan lain. Hotel merasa terbantu dengan adanya OTA, karena OTA merupakan sarana tempat memilih hotel dengan menampilkan berbagai variasi harga, review, serta fasilitas yang diberikan oleh hotel. Namun disamping itu, hotel harus membayar komisi 18%-28% terhadap OTA. Permasalahan tersebut menjadikan tujuan bagi penelitian ini yaitu penetapan strategi hotel dalam meningkatkan penjualan dan menetapkan harga pada OTA. Metode klasifikasi menggunakan *text mining* dan k-means yang diperoleh serta wawancara dengan pihak hotel dalam penelitian ini diharapkan dapat menjadikan ilmu pengetahuan dan membantu pihak hotel dalam menentukan harga dan penjualan.

**Kata kunci:** K-means, Online Travel Agency (OTA), Strategi Harga dan Penjualan, *Text Mining*

### 1. Pendahuluan

Perkembangan Teknologi Informasi dan Komunikasi (TIK) dan khususnya Internet telah memiliki dampak besar pada era digital (Buhalis & Law, 2008; Kamarulzaman, 2007; amaro&duarte, 2004). OTA (*Online Travel Agency*) mencakup dua hal, pertama ialah transportasi (pesawat, kereta api, bus, kendaraan sewa) dan akomodasi untuk perjalanan (hotel, hostel, guest house, dan kos). Perkembangan pada era digital yang berdampak besar pada *travel agencies* telah mengubah perilaku wisatawan (Buhalis & Law, 2008; Hung, Yang, Yang, & Chuang, 2011) yang sekarang bergantung pada Internet untuk mencari informasi, merencanakan perjalanan, dan membeli tiket dan akomodasi perjalanannya (amaro & Duarte, 2008). Google (2014) menyatakan bahwa 65% wisatawan *leisure* dan 67% wisatawan bisnis, mencari informasi di internet terlebih dahulu untuk mencari segala informasi tentang suatu tempat yang akan dikunjungi, seperti tempat wisata, moda transportasi, tempat makan, tempat menginap, dan sebagainya (Paturusi & Tirtayasa, 2017). Menurut para wisatawan, membuat pengaturan perjalanan sendiri, membuat reservasi maskapai, dan hotel, melalui internet merupakan hal yang dapat menghemat biaya, karena dapat menghilangkan biaya perantara seperti pada *offline travel agencies* (Hamid, 2011). *Traditional travel agencies* biasanya menjual paket liburan, menawarkan konseling, serta nasihat perjalanan, juga memberikan keunggulan dalam penjualan, sementara OTA melayani pembeli yang lebih *independent* serta ditentukan oleh harga dan waktu yang sensitif. Nilai plus yang di berikan dibanding dengan *travel tour offline* adalah saluran non-stop yang dapat diakses oleh siapapun untuk pemesanan transportasi dan akomodasi, serta terciptanya hal yang praktis, efektif, dan efisien.

Penelitian ini akan berfokus pada salah satu akomodasi yang ditawarkan oleh OTA yakni Hotel. Dalam situs OTA, hotel yang berada dalam suatu lokasi dikelompokkan dalam satu kategori sehingga memudahkan dalam memilih, dimana didalamnya diberikan fasilitas untuk (1) melihat hotel pada daerah tujuan wisata yang dituju, (2) membandingkan harga dan fasilitas hotel-hotel yang terdaftar pada OTA serta melihat *review* atau ulasan dari tamu yang telah melakukan pemesanan, dan (3) memesan kamar pada situs web tersebut. Berbagai kemudahan tersebut membuat OTA menjadi lebih unggul dibanding situs web yang disediakan oleh hotel

(Tirtayasa & Paturusi, 2016). Hal tersebut terjadi karena situs web hotel hanya menyediakan informasi mengenai hotel-hotel tertentu.

Di negara maju, 61-75% dari customer lebih memilih OTA sebagai situs yang digunakan untuk memudahkan perjelesan mereka dalam memesan hotel dan akomodasi lainnya, sedangkan hanya 15 - 44% dari customer yang mengunjungi situs hotel yang akan diinapi oleh customer. Secara umum, 81% wisatawan nusantara sudah menggunakan saluran digital (*search engine*, situs destinasi wisata dan OTA) dalam merencanakan perjalanan mereka (Yosi Winosa, 2018).

OTA memiliki banyak hal yang menarik bagi wisatawan atau pebisnis dengan kumpulan informasi hotel, perbandingan harga, diskon, dan komentar ulasan (*nline review*). Sebaliknya, situs web hotel milik hotel dan hanya melayani milik hotel saja. Maka dari itu, OTA memiliki jaringan yang luar biasa dalam menjual hotel.

Dalam kasus harga, beberapa hotel *chain* ketika ditawarkan penurunan harga oleh OTA, mereka akan menyanggupi, karena OTA merupakan salah satu jaringan yang dapat meningkatkan revenue hotel, meskipun hotel masih harus membayar *fees* terhadap OTA. Permasalahan diatas menjadikan tujuan bagi penelitian ini yakni hotel harus mempunyai strategi dalam penentuan harga dan penjualan agar bisa menutupi kekurangan yang harus dibayarkan terhadap OTA.

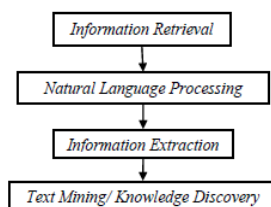
Ulasan merupakan salah satu hal yang dapat mempengaruhi penjualan, karena Konsumen mudah dipengaruhi oleh pendapat orang lain, dan kesediaan mereka untuk mencari pengalaman konsumen sebelumnya ketika akan membeli atau memesan suatu produk yang sama (Hou et al, 2019; García-Pablos, Cuadros, & Linaza, 2016; Guo et al., 2017; Hu, Chen, & Chou, 2017; Schuckert, Liu, & Law, 2015). Data industri menunjukkan bahwa sekitar 77% wisatawan akan "selalu" atau "biasanya" tidak membuat keputusan sampai mereka membaca ulasan online (Hou et al, 2019; Ye et al., 2009). Menurut Asosiasi Industri Perjalanan Amerika, bukti menunjukkan bahwa 64% pelancong menggunakan mesin pencari untuk perencanaan perjalanan mereka (Nilashi et al, 2018). Kualitas informasi ulasan online, mempengaruhi sikap konsumen terhadap adopsi sumber informasi dan membuat keputusan pariwisata (Hou et al, 2019; Filieri & Mcleay, 2014; Schuckert et al., 2015a). Selain untuk meningkatkan penjualan dan menjadikan evaluasi bagi hotel, ulasan online juga penting untuk meningkatkan manajemen layanan dan keunggulan kompetitif hotel dan perusahaan pariwisata (Hou et al, 2019; Fang et al., 2016; Lui et al., 2018; Miguéis & Nóvoa, 2017; Yang, Shin, Joun, & Koo, 2016).

Dampak ulasan dan *rating* online adalah intuitif, dan sangat didukung secara empiris (Maeyer, 2012). Sebagian besar, studi ulasan online berfokus pada *rating* (Hue, 2010), sedangkan, ulasan secara online apabila di terapkan dapat meningkatkan keuntungan serta mendapatkan keunggulan kompetitif yang dapat menjadikan strategi harga dan penjualan (Decker and Trusov, 2011; Archak et al., 2011). Dari penjabaran pendahuluan diatas, penelitian ini bertujuan untuk menentukan strategi harga dan penjualan hotel pada OTA dengan menggunakan ulasan *online*, dan menggunakan metode text mining, serta klasifikasi k-means untuk mengelompokkan ulasan *online*, yang dapat menjadikan strategi harga dan penjualan hotel pada OTA. OTA yang digunakan dalam penelitian ini adalah Traveloka, Pegipegi, dan Tiket.com. hotel yang akan dijadikan sampel pada penelitian adalah hotel yang terdapat pada lima kota besar di Indonesia.

## 2. Metode

Penelitian ini menggunakan metode *text mining*, dimana data yang didapat adalah ulasan online yang didapat dari situs web 3 OTA (Traveloka, Pegipegi, dan Tiket.com) menggunakan web crawling dengan bantuan *software* R. Tahap selanjutnya adalah, mengklasifikan ulasan menjadi suatu strategi penjualan dan harga yang dapat digunakan hotel dalam bekerja sama dengan OTA. *Text mining* adalah suatu proses mengekstraksi pola yang berguna dari sumber data tekstual (Kaur & Aggarwal, 2013). *Text mining* juga dapat dikatakan suatu proses mengambil data yang berupa teks dimana sumber datanya biasanya didapatkan dari dokumen

dengan tujuan untuk mencari kata-kata yang dapat mewakili isi dari suatu dokumen sehingga dapat dianalisis keterhubungan antar dokumen tersebut. *Text mining* merupakan penerapan konsep dan teknik data mining untuk mencari pola dalam suatu teks. Proses *data mining* untuk data dokumen atau teks memerlukan lebih banyak tahapan, hal ini dikarenakan data teks memiliki karakteristik yang lebih kompleks jika dibandingkan dengan data biasa. Menurut McDonald (2012), beberapa tahapan dalam *text mining* adalah sebagai berikut:



Gambar 1. Tahapan Text Mining

*Information retrieval* merupakan tahapan pertama dalam *text mining*. *Information retrieval* dilakukan untuk memperoleh dokumen yang sesuai dengan permintaan peneliti atau dapat disebut juga sebagai tahapan untuk mengumpulkan dokumen yang sesuai dengan permasalahan. Tahapan yang dilakukan selanjutnya adalah *Natural Language Processing*.

*Document pre-processing* merupakan tahapan yang paling penting dalam *text mining* karena pada tahap ini akan dilakukan proses mengubah dokumen sesuai dengan format yang digunakan dalam penelitian (*data transformation*) dan proses *cleaning data*. *Document pre-processing* dibagi menjadi tiga tahapan, yaitu:

1. *Tokenizing*

*Tokenizing* atau *parsing* adalah suatu proses penguraian string teks menjadi *term* atau kata. Tujuan dilakukannya *tokenizing* adalah untuk mendapatkan *term-term* yang nantinya akan diindeks menggunakan proses tertentu. Pengklasifikasian token dilakukan untuk teks yang dipisahkan dengan spasi atau enter dalam suatu dokumen. Beberapa kasus yang ditangani oleh *parsing* atau *tokenizing* antara lain adalah:

- a. *Handling Special Character*, untuk pengambilan polanya menggunakan *regular expression*.
- b. *Phrase* yaitu kata yang mengandung satu pengertian tetapi tidak memiliki subjek dan predikat seperti nama, tempat, dan kata sifat.
- c. *White Space* lebih mengarah pada karakter pemisah yang tidak diperlukan ini diabaikan oleh *tokenizing* dan dianggap sebagai pemisah.

2. *Filtering*

*Filtering* merupakan proses penyeleksian token-token yang dianggap tidak penting (*stopword*). *Stopword* merupakan kata yang sering muncul dan bisa diabaikan pada proses *filtering*. Kata-kata tersebut seperti “of”, “or”, “is”, “and”, “that”, dan yang lain nya pada dokumen berbahasa Inggris.

3. *Stemming*

*Stemming* adalah suatu proses yang bertujuan untuk mengambil kata dasar dari kata yang berimbuhan ataupun kata tunggal dari kata bentukan, tentunya tahapan ini dapat mengurangi jumlah *term* yang berbeda dalam koleksi. Berikut merupakan hasil dari *stemming*:

Setelah tahapan *document pre-processing* selesai dilakukan, maka diperoleh *Term Document Matrix*. Pada matriks ini, *term-term* yang terbentuk dijadikan sebagai kolom dan *document text* Twitter dijadikan sebagai baris. Sehingga akan terbentuk matriks  $m \times n$ , atau disebut juga dengan matriks **A**. Berikut adalah contoh *term document matrix*:

**Tabel 1.** Contoh term Document Matrix

Document Text	Indeks Frekuensi Term				
	Doc1	Doc2	Doc3	Doc4	Doc5
Term 1	1	0	0	0	0
Term 2	0	0	2	0	0
Term 3	0	0	1	1	0
Term 4	3	1	0	2	0
Term 5	0	1	0	0	1

Setelah diperoleh *term document matrix* maka selanjutnya dilakukan pembobotan terhadap *term* tersebut pada tahapan *treatment*. Tahapan ini merupakan tahapan penentuan seberapa jauh keterhubungan antar suatu kata atau term terhadap suatu dokumen atau kalimat dengan menghitung nilai/bobot keterhubungan. Algoritma *TF/IDF* digunakan dalam proses perhitungan bobot ( $W$ ) terminologi kata. Algoritma ini digunakan untuk menghitung bobot setiap kata yang paling umum digunakan pada *information retrieval*. Metode ini juga terkenal efisien, mudah dan memiliki hasil yang akurat.

Persamaan yang digunakan untuk menghitung bobot ( $W$ ) masing-masing dokumen terhadap kata kunci adalah:

$$W_{d,t} = tf_{d,t} * IDF \quad (1)$$

Dimana:

- $W$  = bobot dokumen ke-n
- $d$  = dokumen
- $t$  = kata kunci
- $tf$  = *terms frequency* (jumlah kemunculan kata)
- $IDF$  = *Inverse Document Frequency*

Nilai  $tf$  diperoleh:

$$tf_{d,t} = \frac{\text{Jumlah munculnya kata } t \text{ dalam dokumen}}{\text{Total jumlah seluruh kata dalam dokumen}} \quad (2)$$

Nilai  $IDF$  didapatkan dari:

$$IDF = \log_2 \frac{D}{df} \quad (3)$$

Dimana:

- $D$  = total dokumen, dalam hal ini total kalimat yang ada
- $df$  = jumlah dokumen yang mengandung kata kunci

### ***K-means***

Menurut Johnson (2002), *K-means* adalah metode non-hierarki yang paling sering digunakan. Pada metode ini sebuah set data diklasifikasikan ke dalam beberapa kluster yang sudah ditentukan, dimana kluster tersebut diasumsikan *fix*. *K-means* menggunakan *centroid* sebagai pusatnya klasternya, *centroid* tersebut biasanya berupa nilai rata-rata. *K-means* mempartisi objek ke dalam  $k$  kluster, kemudian menempatkan objek ke dalam kluster yang memiliki jarak *centroid* terdekat. Setelah itu dilakukan perhitungan kembali nilai *centroid* yang baru, dan dilakukan kembali penempatan objek ke dalam kluster yang memiliki jarak terdekat dengan nilai *centroid* yang baru. Proses tersebut terus dilakukan sampai tidak ada *centroid* yang berpindah.

Algoritma *K-means clustering* diperkenalkan oleh Macqueen pada tahun 1967. Algoritma *clustering* ini dirancang untuk mengelompokkan data dengan skala pengukuran minimal interval, dengan setiap kluster memiliki pusat yang disebut dengan *centroid* atau nilai rata-rata (*mean*). Dalam algoritma ini, jumlah kluster diasumsikan *fix*. Misalkan terdapat satu set data  $X$

dengan  $n$  objek dan  $p$  dimensi. dan  $C_1, C_2, \dots, C_k$  merupakan  $k$  kluster terpisah dari  $X$ . Proses pengelompokan pada algoritma ini bersifat *iterative*, proses iterasi akan berhenti saat diperoleh partisi yang meminimumkan *Error Sum Square* (ESS). *Error* dari fungsi dapat didefinisikan sebagai:

$$ESS = \sum_{i=1}^k \sum_{x \in C_i} d^2(x, \mu(C_i)) \quad (4)$$

dengan:

$(C_i)$  : *centroid* dari kluster  $C_i$

$d_2(x, \mu(C_i))$  : jarak antara  $x$  dan  $\mu(C_i)$

Ukuran jarak yang biasanya digunakan adalah *Euclidean Distance* atau *Standardized Euclidean Distance* yang biasa dikenal dengan *Mahalanobis Distance*.

### Studi Literatur

OTA menyediakan informasi berupa informasi penerbangan, menyediakan informasi tentang lokal tertentu; agen perjalanan online, yang memungkinkan pencarian dan pemesanan tiket hotel, transportasi, dan objek wisata; aplikasi fasilitator, panduan atraksi, yang memberikan tips dan informasi perjalanan tentang atraksi tertentu seperti taman hiburan, yang menyediakan hiburan selama perjalanan. Selain menyediakan informasi diatas, OTA juga memberikan informasi yang berbeda bagi wisatawan lain yang bertujuan untuk liburan, OTA memberikan peluang bagi wisatawan untuk menuliskan dan berbagi pengalaman perjalanan mereka, baik yang memuaskan maupun yang tidak memuaskan. Sehingga melalui web atau *platform* tersebut, wisatawan lain yang akan berkunjung ke wisata atau hotel tersebut dalam memilah mana tujuan yang sesuai atau tidak. (Gursoy, 2017)

OTA sangat berkontribusi pada pemesanan serta peningkatan penjualan kamar hotel. Hotel yang terdaftar pada OTA tentu nya dapat meningkatkan pemesanan langsung, Cornell (2009) menunjukkan bahwa hotel yang terdaftar di OTA melihat peningkatan pemesanan yang dilakukan melalui saluran pemesanan mereka sendiri. Studi ini menunjukkan hingga 26% peningkatan pendapatan yang berasal dari pemesanan langsung (Logt, 2017; Little Hotelier, 2017). Sebagian besar OTA telah meluncurkan aplikasi seluler mereka sendiri selain melalui situs web nya, informasi yang diberikan sama seperti halnya yang dilakukan oleh konsumen di situs web OTA. OTA juga menanamkan teknologi baru di seluruh pengalaman pelanggan yakni *online review* dimana harapannya adalah untuk memberikan penawaran layanan yang unik dapat ditingkatkan dapat dilihat sebagai pembeda. OTA merupakan saluran distribusi yang di gunakan oleh pelaku bisnis perhotelan untuk menjual kamar mereka melalui platform online dengan membayar komisi dari penjualan hotel melalui pesanan kamar hotel konsumen, biaya komisi yang dibayarkan kepada OTA cukup tinggi, sehingga hal ini merupakan salah satu persoalan yang cukup penting bagi Hotel (Gursoy, 2017; Gazzoli, Gon Kim dan Pallakurthi 2008). Untuk menghilangkan biaya komisi, hotel harus fokus untuk melakukan penjualan melalui *direct booking* (Toh et al. 2011; Tso dan Law 2005). Namun demikian, pelaku bisnis perhotelan tidak disarankan untuk berpaling dari penjualan yang dilakukan melalui saluran OTA sepenuhnya krena OTA merupakan saluran yang menyumbang *revenue* terhadap hotel (Gursoy, 2017; Ling et al. 2015).

#### **Online review pada OTA**

Ketika hotel berjuang untuk mempertahankan dan menumbuhkan loyalitas dari OTA distributor OTA lain, OTA menanamkan teknologi baru di seluruh pengalaman pelanggan. Harapannya adalah untuk memberikan penawaran layanan yang ditingkatkan dan layanan unik untuk tamu, yang dapat dilihat sebagai pembeda. Misalnya, banyak pengguna beralih ke ulasan perjalanan untuk menghasilkan ide dan membuat keputusan akhir. Sebagian besar pembaca menjadikan ulasan penting untuk keputusan akomodasi; dengan demikian, hotel dan agen perjalanan online biasanya mengirim email kepada wisatawan setelah mereka menginap,

meminta mereka untuk memberikan ulasan dan untuk sebagian yang beruntung akan mendapatkan voucher untuk menginap dilain waktu, strategi ini merupakan salah satu strategi yang diterapkan oleh OTA agar tetap mengikat pelanggan dan juga hotel (Gursoy, 2017).

### 3. Hasil dan Pembahasan

Pada bagian ini akan menjelaskan serangkaian proses dalam penelitian dimulai dari tahap pengumpulan hingga pengolahan data untuk mendapatkan hasil yang diharapkan. Peningkatan kualitas penjualan dan penetapan strategi bagi dunia perhotelan sangatlah diharuskan, namun OTA juga tidak kalah pentingnya dalam meningkatkan penjualan dan revenue perhotelan. Oleh karena nya diperlukan informasi yang terdapat pada OTA pada ulasan online yang terdapat pada ketiga website OTA yakni Traveloka, Pegipegi, dan Tiket.com dimana ketiga OTA tersebut memiliki *market share* yang tinggi dibandingkan dengan OTA yang lain nya. Penelitian ini menggunakan hotel sebagai objek. Hotel yang dijadikan objek pada penelitian ini adalah hotel yang berada di 5 lokasi hotel yang merupakan paling banyak peminatnya dan diharapkan dengan menggunakan 5 kota yang akan dijadikan perhitungan, dapat merepresentasikan hasil seperti yang diharapkan. 5 kota yang dijadikan penelitian adalah kota Bandung, Jojakarta, Jakarta, Bogor, dan Bali. Ulasan online yang terdapat di lima kota tersebut merupakan ulasan dan hotel terbanyak yang terdapat di ketiga OTA yang telah disebutkan sehingga dapat mewakili ulasan hotel-hotel online yang terdapat pada ketiga OTA. Metode text mining yang mampu menghasilkan pengetahuan dari sejumlah data yang sangat besar dan tidak berstruktur akan digunakan dalam penelitian ini. Secara garis besar, ulasan online dapat digunakan sebagai strategi dalam menentukan strategi penjualan dan harga hotel pada OTA dengan menggunakan text mining dengan teknik klasifikasi sehingga dapat diperoleh klasifikasi yang dapat ditingkatkan oleh hotel agar dapat meningkatkan revenue dan penjualan serta harga yang diberikan kepada OTA.

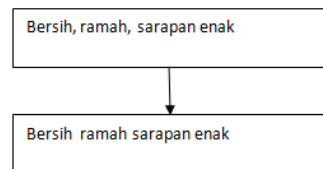
Tahap ini merupakan tahap kedua setelah mendefinisikan masalah dan tujuan penelitian yang sudah dilakukan di bagian sebelumnya. Selanjutnya, tahap yang dilakukan adalah dengan mengumpulkan data yang akan diolah, yang mana data yang digunakan adalah data yang terdapat pada ketiga OTA yang sudah dipilih yakni ([www.traveloka.com](http://www.traveloka.com), [www.pegipegi.com](http://www.pegipegi.com), [www.tiket.com](http://www.tiket.com)). OTA merupakan sebuah wadah untuk biro perjalanan yang berisikan akomodasi, tiket, serta hal-hal lain yang berhubungan travel. Teknik web scraping digunakan untuk memperoleh data yang akan digunakan dan dibutuhkan dalam penelitian. Web scraping adalah proses pengumpulan data melalui interaksi dengan API. Pada umumnya kegiatan ini dilakukan dengan membangun program otomatis yang melakukan permintaan dan biasanya berupa HTML dan komponen lain penyusun sebuah web, kemudian memecah data untuk mengekstrak informasi yang dibutuhkan (Mitchell, 2015). Web scraper beroperasi atas perintah script yang disusun untuk memanggil web dan mendapatkan respon berupa HTML atau halaman web serta informasi yang dibutuhkan. Script dalam penelitian ini disusun dengan menggunakan bahasa pemrograman R.

```
install.packages("httr")
library(httr)
# A simple GET request
r<- GET("https://www.traveloka.com/en-id/hotel/search?spec=06-04-2019.07-04-2019.1.1.HOTEL_GEO.103859.Bandung.2")
bandung2<- "https://www.traveloka.com/en/hotel/indonesia/horison-ultima-bandung-2000000081026"
bdg2<-read_html(bandung2)
review_html<-html_nodes(bdg2, '._1mI1m')
review<-html_text(review_html)
review<-gsub("\n", "", review)
nama_html<-html_nodes(bdg2, '.NkdVj')
nama<-html_text(nama_html)
nama<-gsub("\n", "", nama)
tableout<-data.frame(review, nama)
write.csv(tableout, "D:/bdg2.csv")
```

Gambar 2. Syntax R Web Crawling

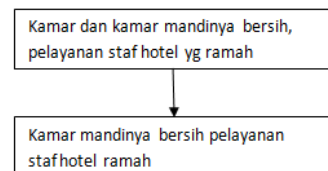
Program penggunaan web scrapping dioperasikan dengan menggunakan software R seperti yang terlihat pada gambar diatas program ini digunakan untuk mengunduh informasi yang digunakan yakni hotel-hotel yang berada pada 5 kota yang telah disebutkan pada bab sebelumnya.

Tahap selanjutnya adalah tokenization, tahap ini dilakukan yakni pemecahan kalimat kedalam potongan kata atau yang disebut dengan tokenization. Tahap ini dilakukan untuk mendapatkan token yang akan menjadi entitas bernilai dalam penyusunan matriks dokumen pada proses selanjutnya. Proses tokenization ditunjukkan pada gambar berikut:



**Gambar 3.** Tokenization

Tahap berikutnya adalah menghilangkan tanda baca atau kata yang tidak bermakna contohnya seperti emoticon atau symbol yang sering dijumpai pada tulisan bebas. Proses ini dikenal sebagai filtering yang dilakukan bersamaan dengan stop word removal yang ditunjukkan pada gambar diatas. Daftar stopwords yang digunakan adalah seperti on, of, that, dll.



**Gambar 4.** Stopwords Removal

Langkah berikutnya adalah mengoperasikan algoritma stemming untuk menormalkan token-token menjadi satu bentuk yang seragam yakni kata dasarnya. Proses stemming dilakukan dengan menghilangkan semua imbuhan, baik awalan, sisipan, dan akhiran.

### **Treatment**

Setelah diperoleh *term document matrix*, kemudian dilakukan pembobotan terhadap *term-term* yang muncul dengan menggunakan TF-IDF, pengolahan pembobotan *term* tersebut dilakukan dengan menggunakan *software R*. Pembobotan ini dilakukan untuk mendapatkan hasil yang digunakan untuk melakukan pengukuran jarak pada pengelompokan. Setelah dilakukan pembobotan dengan menggunakan *software R* dengan menggunakan *syntax*.

Selanjutnya menghitung inverse frekuensi dokumen yang mengandung kata kunci dengan menggunakan persamaan (3):

$$\begin{aligned} IDF &= \log_2 \frac{D}{df} \\ &= \log_2 \frac{10354}{2138} \\ &= 0.6815 \end{aligned}$$

### **Klasifikasi**

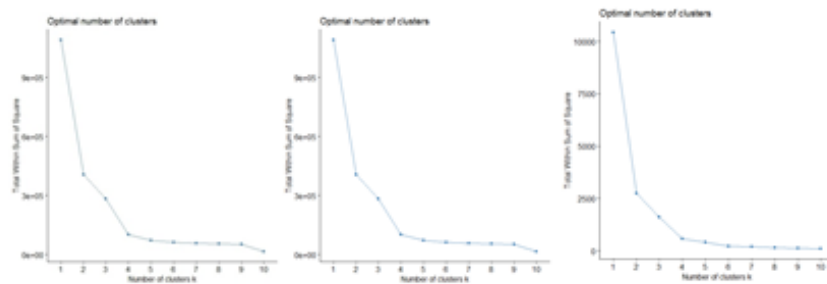
Tahapan ini bertujuan untuk menggolongkan kalimat ulasan kedalam 5 klasifikasi yang telah dibuat yang meliputi (Food & Beverage, Location, activities & entertainment, Service) berdasarkan ulasan online yang telah diambil. Pengklasifikasian ini digunakan pada penelitian





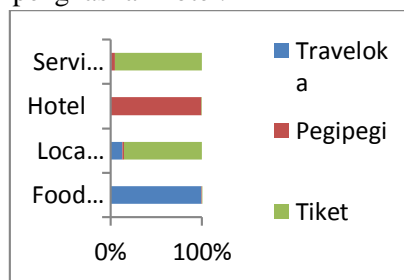
### Klasifikasi

Proses klasifikasi pada penelitian ini sudah dijelaskan pada bab sebelumnya, dimana ulasan online yang digunakan adaah selama tahun 2019. Kata-kata tematik yang dipilih dikategorikan oleh tiga ahlidalam e-commerce sesuai dengan kriteria bahwa kata-kata tematik dengan atribusi serupa dikelompokkan ke dalam kategori masing-masing dengan menganalisis keterkaitan mereka dan keteraturan logis. Sebagai hasilnya, kami mengidentifikasi empat topik: (Food & Beverage, Location, Hotel, Service). Dari ketiga Website menunjukkan 4 kluster merupakan pembentukan kluster yang optimum yang dapat dilihat dari gambar dibawah ini dengan menggunakan syntax *R*:



Gambar 8. Optimal Klustering K-means Traveloka, Pegipegi, tiket.com

Penulis mendefinisikan setiap topik untuk mengurangi bisa pemahaman pada topik berikutnya. Table dibawah menjelaskan bahwa ulasan yang diberikan oleh para pengguna di website OTA Traveloka lebih banyak yang bersifat puas terhadap makanan pada hotel, sedangkan pada pegipegi.com para pengguna lebih menyukai Hotel dibanding makanan, dan pada tiket.com para pengguna yang menggunakan ulasan lebih banyak yang puas dan mengomentari lokasi & service hotel tersebut. Hal ini dapat dijadikan patokan untuk para *hospitality* bahwa para pengguna pada masing-masing OTA memiliki kesukaan atau kelebihan kesukaan pada masing-masing OTA, para pemilik hotel dapat memanfaatkan hal ini untuk menunjang kompetitif advantage agar penjualan semakin banyak dan juga menjadikan sebagai strategi yang dapat meningkatkan penghasilan hotel.



Gambar 9. Klasifikasi Cluster pada setiap OTA

### Analisa Kualitatif

Setelah melakukan perhitungan, penelitian ini juga melakukan wawancara terhadap manajer hotel yang tentunya berhubungan dengan OTA serta dapa menunjang hasil yang sudah diperhitungkan. Manajer hotel mengatakn bahwa harga disediakan setiap hari oleh hotel dan diberikan kepada OTA, beberapa htel memberikan harga yang sama kepada OTA, yang membedakan hanyalaharga yang ditampilkan pada website setiap OTA, perbedaan yang terjadi adalah komisi atau fee yang diberikan OTA berbeda-beda ke setiap hote. Untuk beberapa *chain hotel*, ketika ditawarkan diskon oleh pihak OTA mereka mengiyakan, untuk menutupi revenue yang rendah, mereka menggunakan top session agar lebih dapat menghasilkan keuntungan. Para

hotel meningkatkan sales mereka dengan cara menelfon pihak perusahaan untuk menginap di hotel mereka atau saat sedang melakukan acara agar dapat dilakukan pada hotel mereka. Cara lain yang dilakukan untuk meningkatkan revenue dan menghasilkan keuntungan yang lebih adalah dengan cara melihat review yang terdapat pada OTA, mereka sangat membutuhkan ulasan online untuk dapat memperbaiki kinerja dan service mereka. Strategi harga yang diberikan oleh hotel kepada OTA merupakan harga yang sesungguhnya dan merupakan perintah bagi beberapa chain hotel. Strategi yang digunakan adalah dengan mengikuti diskon yang disediakan pada OTA dan pada top session serta menelfon perusahaan-perusahaan yang akan bekerja sama dengan hotel dan menggunakan ulasan online sebagai cara untuk memperbaiki system dan meningkatkan kinerja mereka atau sebagai masukan pada saat membuka cabang baru disuatu wilayah.

#### 4. Simpulan

Penelitian ini menemukan perbedaan yang dimiliki oleh pengguna dalam menggunakan OTA. Pengguna traveloka lebih focus pada hal *food and beverage*, sedangkan pegipegi lebih focus terhadap *location* hotel, dan tiket.com yang lebih focus kepada *service* pada hotel tersebut. Penelitian ini menunjukkan bahwa Temuan ini menunjukkan bahwa tingkat layanan pemandu wisata memiliki pengaruh besar pada kepuasan wisatawan pada satu platform. Dalam dunia pariwisata, hotel dan tour guide adalah kunci untuk memastikan kepuasan wisatawan. Makanan adalah focus utama pada wisatawan yang menggunakan traveloka, lokasi dan pelayanan hotel adalah focus utama wisatawan yang menggunakan pegipegi dan tiket.com. hal ini juga bisa dijadikan rekomendasi untuk hotel-hotel untuk dapat meningkatkan penjualan dengan cara mempertahankan dan memperbaiki focus yang dicari oleh para wisatawan.

Strategi harga yang digunakan terhadap OTA adalah dapat dilakukan dengan mengurangi harga seperti Rp. 398.000 agar dapat meningkatkan penjualan di setiap OTA, karena hal itu berpengaruh untuk para konsumen yang melihat. Atau dengan harga tetap namun meningkatkan pelayanan dan kualitas yang digunakan oleh hotel.

Lingkup Penelitian ini adalah 3 OTA (traveloka, Pegipegi, dan tiket.com) serta dengan 5 kota yang dipilih berdasarkan kota yang diminati untuk pekerjaan maupun liburan, dengan menggunakan metode k-means dan digunakan data ulasan online selama 2019. Penelitian selanjutnya disarankan dapat menggunakan data histori yang lebih banyak agar lebih akurat. Selain itu dapat juga menggunakan metode yang berbeda.

#### Daftar Pustaka

- Alizadeh, A., Akbari, E., Samad, S., Yadegaridehkordi, E., Ibrahim, O., Nilashi, M. (2018). Travelers decision making using online review in social network sites: A case on TripAdvisor. *Journal of Computational Science* 28 (2018) 168–179
- Amaro, S., Duarte, P. (2008). Online Travel Purchasing: A Literature Review. *Journal of Travel & Tourism Marketing* 30(8):755-785.
- Fuentes, M., A., & Mellinas, P., J. (2018). Hotels that Most Rely On Booking.com : OTA and Hotels Channel. *Tourism Review*
- Guo, X., Zheng, X., Ling, L., & Yang, C. (2014). Online Competition between Hotels and OTA : from the Perspective of Cash Back After Stay. *Tourism Management Perspectives* 12 (2014) 104–112
- Gursoy, D. (2017) *The Routledge Handbook of Hospitality Marketing*-Routledge. London.
- Practical Text Mining and Statistical principle
- Hamid, I. A. (2011). *Travel Agency Strategies for managing The Current Dynamic Environment*. Malaysia

- Iskandar. (2017, November 13). *Pasar Online Travel Indonesia Tumbuh Tercepat hingga 2020*. Diambil kembali dari Liputan6: <https://www.liputan6.com/teknoread/3160611/pasar-online-travel-indonesia-tumbuh-tercepat-hingga-2020>
- Maeyer, D., P., (2012). Impact of online consumer reviews on sales and price strategies: a review and directions for future research. *Journal of Product & Brand Management*, Vol. 21 Iss 2 pp. 132 – 139
- Online Travel Agency (2017, Agustus 28)
- Online Travel Agency di Akses pada (2018, November, 15)
- Permana, Y. (2016, Mei 29). *Mengenal Big Data*. Diambil kembali dari Code Politan: <https://www.codepolitan.com/mengenal-big-data>
- Tirtayasa, P. P. I., & Paturusi, P. A. (2016). Penggunaan Situs Web Hotel dan OTA sebagai Media Promosi dan Penjualan bagi Hotel Melati di Ubud . *JUMPA 3[1]:60-71*
- Villarinho, A., T. (2014). Pursuit of competitive Advantage for Online Travel Agencies: Driving From Price to Value. *Universidade Catolica Portuguesa*.
- Yu, C., Lian, T., Meng, Y., Cui, F., Hou, Z. (2019). Opinio Mining from Online Travel Reviews: A Comparative Analysis of Chinese major OTAs using semantic association analysis. *Management 74 (2019) 276–289, China*