

***COMPANY PROFILE SEBUAH HOTEL MELALUI  
ONLINE REVIEW MENGGUNAKAN TEXT  
SUMMARIZATION***

**Tugas Akhir**

**Diajukan untuk Memenuhi Salah Satu Persyaratan Mencapai Derajat  
Sarjana Informatika**



Dibuat Oleh:

**DIONISIUS YUDHA SETIAWAN**

**160708950**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNOLOGI INDUSTRI  
UNIVERSITAS ATMA JAYA YOGYAKARTA  
2020**

# LEMBAR PENGESAHAN

## HALAMAN PENGESAHAN

Tugas Akhir Berjudul

COMPANY PROFILE SEBUAH HOTEL MELALUI ONLINE REVIEW MENGGUNAKAN TEXT SUMMARIZATION

yang disusun oleh

DIONISIUS YUDHA SETIAWAN

160708950

dinyatakan telah memenuhi syarat pada tanggal 03 Juli 2020

		Keterangan
Dosen Pembimbing 1	: Prof. Ir. A. Djoko Budiyanto, M.Eng., Ph.D.	Telah menyetujui
Dosen Pembimbing 2	: Ike Devi Sulistyningtyas, S.Sos., M. Si.	Telah menyetujui
Tim Penguji		
Penguji 1	: Prof. Ir. A. Djoko Budiyanto, M.Eng., Ph.D.	Telah menyetujui
Penguji 2	: Dr. Pranowo, S.T., M.T.	Telah menyetujui
Penguji 3	: Yulius Harjoseputro, ST., MT.	Telah menyetujui

Yogyakarta, 03 Juli 2020

Universitas Atma Jaya Yogyakarta

Fakultas Teknologi Industri

Dekan

ttd

Dr. A. Teguh Siswanto, M.Sc



# PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Dionisius Yudha Setiawan  
NPM : 160708950  
Program Studi : Informatika  
Fakultas : Teknologi Industri  
Judul Penelitian : *Company Profile* Sebuah Hotel Melalui *Online Review* Menggunakan *Text Summarization*

Menyatakan dengan ini:

1. Tugas Akhir ini adalah benar tidak merupakan salinan sebagian atau keseluruhan dari karya penelitian lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta atas penelitian ini, berupa Hak untuk menyimpan, mengelola, mendistribusikan, dan menampilkan hasil penelitian selama tetap mencantumkan nama penulis.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum atas pelanggaran Hak Cipta dalam pembuatan Tugas Akhir ini.

Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 17 Juni 2020

Yang menyatakan,

Dionisius Yudha Setiawan

160708950

## HALAMAN PERSEMBAHAN

**“Pendidikan merupakan perlengkapan paling baik untuk hari tua”  
(Aristoteles)**



## KATA PENGANTAR

Puji dan syukur penulis haturkan kepada Tuhan Yang Maha Esa karena berkat rahmat dan karunia-Nya penulis dapat menyelesaikan pembuatan tugas akhir “*Company Profile Sebuah Hotel Melalui Online Review Menggunakan Text Summarization*” ini dengan baik.

Penulisan tugas akhir ini bertujuan untuk memenuhi salah satu syarat untuk mencapai derajat sarjana Informatika dari Program Studi Informatika, Fakultas Teknologi Industri di Universitas Atma Jaya Yogyakarta.

Penulis menyadari bahwa dalam pembuatan tugas akhir ini penulis telah mendapatkan bantuan, bimbingan, dan dorongan dari banyak pihak. Untuk itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Tuhan Yesus Kristus yang selalu membimbing dalam iman-Nya, memberikan berkat-Nya, dan menyertai penulis selalu.
2. Bapak Dr. A. Teguh Siswanto, M.Sc., selaku Dekan Fakultas Teknologi Industri, Universitas Atma Jaya Yogyakarta.
3. Bapak Martinus Maslim, S.T., M.T., selaku Kepala Program Studi Informatika, Universitas Atma Jaya Yogyakarta.
4. Bapak Prof. Ir. A. Djoko Budiyanto SHR, M.Eng., Ph.D., selaku dosen pembimbing I yang telah membimbing dan memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.
5. Ibu Ike Devi Sulistyningtyas, S.Sos., M.Si., selaku dosen pembimbing II yang telah membimbing dan memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.
6. Kedua orang tua yang selalu mendoakan dan memberi dukungan secara moral maupun material kepada penulis untuk dapat selalu berusaha dalam mencapai hasil yang terbaik.
7. Hastungkoro yang selalu membantu dan menyemangati penulis dalam proses

pengerjaan maupun penulisan.

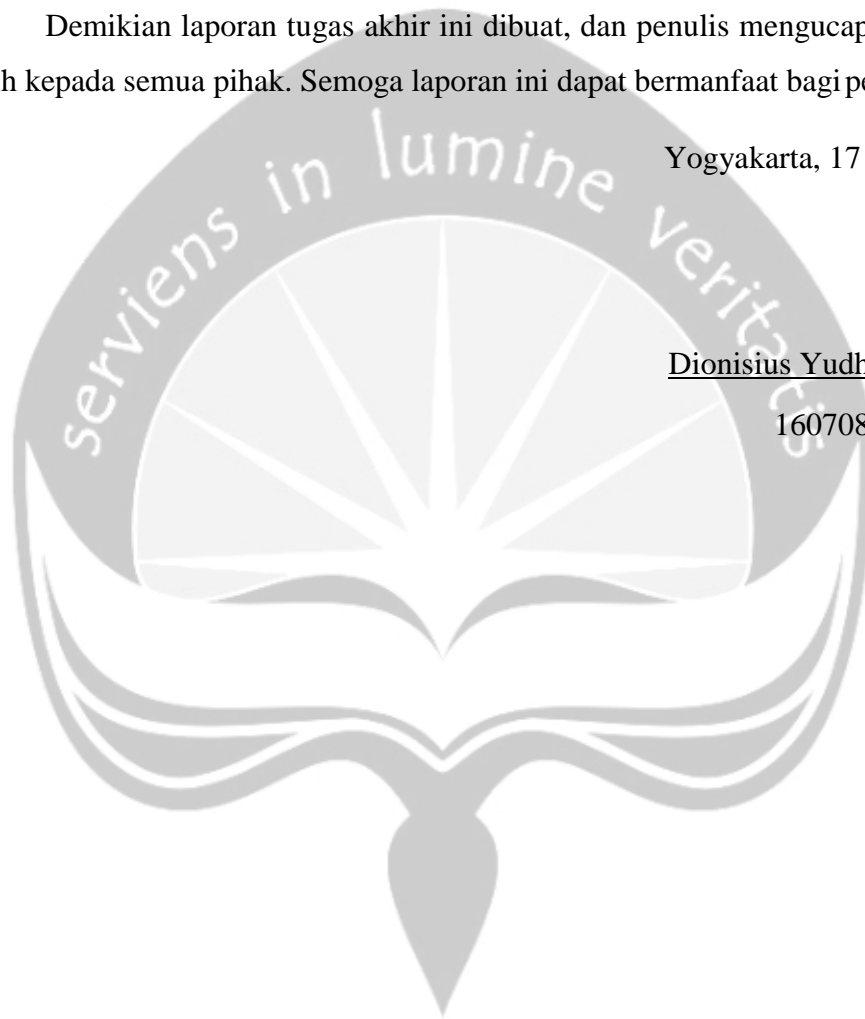
8. Marcellinus Hari Prasetyo yang memberikan dukungan kepada penulis selama proses pengerjaan.
9. Kepada semua pihak yang telah berkenan memberikan bantuan dan dukungan serta kerja sama yang baik sehingga laporan ini dapat diselesaikan dengan lancar.

Demikian laporan tugas akhir ini dibuat, dan penulis mengucapkan terima kasih kepada semua pihak. Semoga laporan ini dapat bermanfaat bagi pembaca.

Yogyakarta, 17 Juni 2020

Dionisius Yudha Setiawan

160708950



# DAFTAR ISI

JUDUL .....	i
LEMBAR PENGESAHAN .....	ii
PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH.....	iii
HALAMAN PERSEMBAHAN .....	iv
KATA PENGANTAR .....	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR .....	viii
DAFTAR TABEL.....	ix
INTISARI.....	x
BAB I. PENDAHULUAN .....	1
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	2
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian .....	3
1.5. Metode Penelitian.....	3
1.6. Sistematika Penulisan .....	4
BAB II. TINJAUAN PUSTAKA.....	6
BAB III. LANDASAN TEORI.....	16
3.1 <i>Text Summarization</i> .....	16
3.2 <i>Company Profile</i> .....	19
3.3 Hotel.....	19
3.4 <i>Online Review</i> .....	20
BAB IV. ANALISIS DAN PERANCANGAN EKSPERIMEN .....	22
4.1. Deskripsi Problem.....	22
4.2. Analisis Kebutuhan Eksperimen.....	22
4.3. Perancangan Eksperimen .....	24
BAB V. HASIL EKSPERIMEN DAN PEMBAHASAN .....	38
5.1. Deskripsi Eksperimen .....	38
5.2. Hasil Eksperimen .....	38
5.3. Pembahasan Eksperimen.....	106
BAB VI. PENUTUP .....	112
6.1. Kesimpulan .....	112
6.2. Saran.....	112
DAFTAR PUSTAKA .....	114

## DAFTAR GAMBAR

Gambar 1. Mengambil <i>Review</i> Hotel Menggunakan <i>ScrapeStorm</i> .....	23
Gambar 2. Hasil Pengambilan <i>Review</i> Hotel dari Media <i>TripAdvisor</i> .....	23
Gambar 3. Membaca Data yang Akan Dilakukan <i>Preprocessing</i> .....	24
Gambar 4. Mengubah Huruf Besar Menjadi Huruf Kecil.....	24
Gambar 5. <i>Remove Emoticon</i> .....	25
Gambar 6. <i>Remove Tanda Baca</i> .....	25
Gambar 7. Normalisasi Singkatan dan Tanda Baca Titik-Titik.....	26
Gambar 8. Daftar Kata Normalisasi Singkatan dan Tanda Baca Titik-Titik .....	26
Gambar 9. Menambah Spasi Setelah Tanda Titik (.).....	27
Gambar 10. <i>Remove Angka</i> .....	27
Gambar 11. <i>Stopword Removal</i> .....	27
Gambar 12. Daftar <i>Stopword</i> .....	28
Gambar 13. Menghapus Spasi Berlebih.....	28
Gambar 14. Memanggil Fungsi untuk Melakukan <i>Preprocessing</i> dan Menyimpan Hasil ke dalam file .txt .....	29
Gambar 15. Membaca File yang Akan Dilakukan <i>Text Summarization</i> .....	30
Gambar 16. <i>Tokenize the Sentence</i> .....	31
Gambar 17. Menghitung Frekuensi Kemunculan Kata.....	31
Gambar 18. Menghitung <i>Term Frequency(TF)</i> .....	32
Gambar 19. Menghitung Frekuensi Kemunculan Kata (t) dalam Satu Dokumen (DK) .....	32
Gambar 20. Menghitung <i>Inverse Document Frequency (IDF)</i> .....	33
Gambar 21. Menghitung <i>TF-IDF</i> dan Menghasilkan Matrik .....	33
Gambar 22. Menghitung <i>TF-IDF</i> untuk Memberi Bobot Setiap Kalimat (D).....	34
Gambar 23. Menghitung Skor Rata-Rata Kalimat .....	34
Gambar 24. Menentukan Hasil <i>Summarization</i> .....	35
Gambar 25. Memanggil Fungsi untuk Melakukan <i>Text Summarization</i> dan Menyimpan Hasil ke dalam File .txt.....	35
Gambar 26. Hasil <i>Text Summarization</i> .....	36



## DAFTAR TABEL

Tabel 1. Perbandingan Penelitian.....	13
Tabel 2. Hasil <i>Preprocessing</i> .....	29
Tabel 3. Perbandingan Hasil <i>Text Summarization</i> Dua dan Tiga Kali Hotel Grand Tjokro.....	50
Tabel 4. Perbandingan Hasil <i>Text Summarization</i> Dua dan Tiga Kali Hotel Novotel .....	64
Tabel 5. Perbandingan Hasil <i>Text Summarization</i> Dua dan Tiga Kali Hotel Royal Ambarukmo .....	78
Tabel 6. Perbandingan Hasil <i>Text Summarization</i> Dua dan Tiga Kali Hotel Tentrem .....	92
Tabel 7. Perbandingan Hasil <i>Text Summarization</i> Dua dan Tiga Kali Hotel THE 101 .....	103



# INTISARI

## *Company Profile Sebuah Hotel Melalui Online Review Menggunakan Text Summarization*

Intisari

Dionisius Yudha Setiawan

160708950

*Online review* memiliki banyak pengaruh terhadap proses pembelian sebuah produk atau pemilihan jasa. Ulasan konsumen merupakan salah satu informasi yang lebih diandalkan oleh konsumen pemula dibandingkan dengan melihat ulasan atau saran yang diberikan oleh pihak penjual. Penelitian ini bertujuan untuk mendapatkan sebuah informasi mengenai kepuasan konsumen terhadap pelayanan yang dilakukan oleh suatu hotel karena itu diperlukan suatu metode untuk memahami dan menyajikan seluruh informasi yang dibutuhkan secara jelas dan ringkas dalam *summarization*. Sehingga ini memungkinkan para pembaca dapat menghemat waktu dan energi.

Untuk mendapatkan sebuah informasi dan poin paling penting mengenai sebuah hotel dari data yang ada, penelitian ini akan dilakukan pencarian informasi menggunakan *Extractive Text Summarization*. Kemudian, menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)* untuk melakukan *text summarization* pada *review* hotel dalam Bahasa Indonesia.

Dalam penelitian ini menggunakan *review* 5 hotel dengan total data yang didapatkan berjumlah 3418 data. Hasil penelitian menunjukkan bahwa nilai rata-rata yang didapatkan dalam perhitungan setiap hotel dapat mengambil kalimat yang sering dibahas dalam *review* setiap hotel dan menyimpulkan ciri-khas masing-masing hotel. Berdasarkan langkah-langkah proses *text summarization* dengan metode *TF-IDF* dapat menghasilkan informasi dari dokumen *review* hotel secara otomatis.

Kata Kunci: *text summarization, TF-IDF, online review, hotel*

Dosen Pembimbing I : Prof. Ir. A. Djoko Budiyanto SHR, M.Eng., Ph.D.

Dosen Pembimbing II : Ike Devi Sulistyaningtyas, S.Sos., M.Si.

Jadwal Sidang Tugas Akhir : Jumat, 03 Juli 2020

# BAB I. PENDAHULUAN

## 1.1. Latar Belakang

Pada era globalisasi ini penggunaan internet semakin terus meningkat dengan pesat. Indonesia merupakan salah satu negara terbesar yang masyarakatnya menggunakan internet hingga 133 juta pengguna pada tahun 2016 [1]. Banyaknya pengguna internet di Indonesia dimanfaatkan dalam penggunaan sosial media. Banyak tindakan yang dilakukan dalam sosial media yang salah satunya melakukan ulasan terhadap suatu produk atau jasa. Ulasan secara online ini memiliki dampak yang cukup besar terhadap keputusan yang akan diambil oleh konsumen. *Online review* ini memiliki banyak pengaruh terhadap proses pembelian sebuah produk atau pemilihan jasa. Ulasan konsumen ini menjadi salah satu informasi yang lebih dihandalkan oleh konsumen pemula dibandingkan melihat ulasan atau saran yang diberikan oleh pihak penjual. Sehingga banyak yang memanfaatkan ulasan online sebagai media untuk dapat melakukan pemasaran sebuah produk atau jasa agar lebih meningkatkan minat konsumen.

Pada masa lalu konsumen dalam memesan suatu hotel akan mempertimbangkan hotel yang ada secara detail sebelum memesannya. Hotel yang akan dipesan akan dibandingkan dengan hotel yang ada disekitarnya berdasarkan pembicaraan mulut ke mulut. Namun dengan kemajuan teknologi yang semakin canggih pemilihan hotel saat ini banyak dipengaruhi terhadap *online review* suatu hotel yang dituliskan dalam media sosial. *Online review* pada saat ini banyak dimanfaatkan untuk mempromosikan penjualan dari sebuah produk yang ada [2]. Pentingnya suatu ulasan bagi konsumen terhadap pengambilan keputusan merupakan suatu proses yang dimulai dengan konsumen memperhatikan (*attention*) pada suatu barang atau jasa lalu akan menimbulkan seorang konsumen merasa tertarik (*interest*) untuk mencari informasi mengenai barang atau jasa tersebut. Setelah konsumen merasa semakin tertarik dengan barang atau jasa maka akan menjadi berminat (*desire*)

yang kemudian akan menimbulkan seorang konsumen memutuskan untuk membeli suatu barang atau jasa tersebut [3].

Penelitian ini akan dilakukan untuk menganalisa *online review* terhadap *company profile* pada suatu hotel di sosial media. Selain itu, penelitian yang berjudul “*Company Profile Sebuah Hotel Melalui Online Review Menggunakan Text Summarization*” bertujuan untuk mendapatkan sebuah informasi dari *online review* yang dilakukan oleh konsumen untuk mengetahui *company profile* sebuah hotel menggunakan *text summarization*. Perlunya mengetahui sikap konsumen terhadap kepuasan pelayanan sebuah hotel dapat digunakan menjadi strategi manajemen hotel untuk dapat meningkatkan kualitas dari pelayanan hotelnya. Dalam analisa ini data yang digunakan merupakan data *review* yang dilakukan oleh konsumen di media sosial *TripAdvisor*.

## **1.2. Rumusan Masalah**

Rumusan masalah yang dapat disimpulkan dalam penelitian ini berdasarkan latar belakang yang ada diatas yaitu bagaimana mendapatkan sebuah informasi dari *online review* yang dilakukan oleh konsumen untuk mengetahui *company profile* sebuah hotel menggunakan *text summarization*?

## **1.3. Batasan Masalah**

Berdasarkan latar belakang dan rumusan masalah yang ada maka batasan masalah dalam penelitian ini adalah sebagai berikut :

- a. Data yang akan digunakan dalam analisis ini adalah data *review* konsumen sampai bulan maret 2020 dalam Bahasa Indonesia di media sosial *TripAdvisor* dari konsumen Hotel Royal Ambarrukmo Yogyakarta, Hotel THE 101 Yogyakarta, Hotel Novotel Yogyakarta, Hotel Grand Tjokro Yogyakarta dan Hotel Tentrem.
- b. Metode yang digunakan untuk analisis data adalah *Extractive text summarization*.

#### 1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang ada, maka dapat dilihat tujuan yang diharapkan dari penulisan ini yaitu untuk mendapatkan sebuah informasi dari *online review* yang dilakukan oleh konsumen untuk mengetahui *company profile* sebuah hotel menggunakan *text summarization*.

#### 1.5. Metode Penelitian

Jenis penelitian yang digunakan dalam penelitian ini merupakan penelitian kuantitatif yang akan menekankan analisa pada data numerik yang diharapkan akan mendapatkan sebuah informasi yang baru berdasarkan angka dan data yang dikumpulkan serta dianalisa. Namun hasil dari penelitian ini tidak sepenuhnya mengacu pada angka namun juga mengacu pada pernyataan-pernyataan yang ada. Tahapan dari penelitian ini yaitu:

1. Studi Pustaka

Pada tahap pertama ini akan dicari dan dipelajari metode *TF-IDF* yang nantinya memiliki keterkaitan dengan masalah yang akan dibahas dalam penelitian ini. Sumber pustaka yang digunakan dapat diperoleh dari berbagai macam sumber yaitu *internet*, buku, *e-book*, dan jurnal yang bersifat nasional dan internasional.

2. Pengumpulan Dokumen Review Hotel

Pada tahap ini dilakukan pengumpulan data dokumen *review* hotel dalam Bahasa Indonesia, dimana data tersebut akan digunakan untuk melakukan pengujian *text summarization*. Dokumen *review* hotel yang digunakan bersumber dari *review* hotel dalam *TripAdvisor*.

3. *Preprocessing*

Pada tahap ini akan dilakukan ekstraksi informasi, di mana kata yang tidak penting akan dihapus sehingga didapatkan hasil akhir yang diperlukan dalam *text summarization*. Pada *preprocessing* ini dilakukan delapan urutan dengan tahapan proses yaitu:

- a. *Case Folding*
- b. *Remove Emoticon*

- c. *Remove* Tanda Baca
- d. Normalisasi Singkatan dan tanda baca titik-titik
- e. Menambah Spasi setelah tanda titik (.)
- f. *Remove* Angka
- g. *Stopword Removal*
- h. Menghapus Spasi Berlebih

#### 4. *Text Summarization*

Pada tahap ini akan dilakukan *text summarization* untuk menghasilkan informasi dari dokumen *review* hotel. Pada *text summarization* ini dilakukan dengan urutan tahapan proses yaitu:

- a. *Tokenize the Sentence*.
- b. Menghitung frekuensi kemunculan kata ( $t$ ) di setiap kalimat ( $D$ ).
- c. Menghitung *term frequency (TF)*.
- d. Menghitung frekuensi kemunculan kata ( $t$ ) dalam satu dokumen ( $DK$ ).
- e. Menghitung *inverse document frequency (IDF)*.
- f. Menghitung *TF-IDF* dan menghasilkan matrik.
- g. Menghitung *TF-IDF* untuk memberi bobot setiap kalimat ( $D$ ).
- h. Menghitung skor rata-rata kalimat.
- i. Menentukan hasil *summarization* dengan mengambil kalimat ( $D$ ) yang memiliki nilai lebih besar atau sama dengan nilai rata-rata.
- j. Melakukan peringkasan *text summarization* kedua dan ketiga kalinya dengan mengulang langkah a hingga i untuk mendapatkan hasil yang lebih ringkas.

#### 5. Evaluasi dan pembahasan

Pada tahap ini, evaluasi dan pembahasan dilakukan untuk mengetahui hasil dari *text summarization* dan dilakukannya penggabungan teori ilmu komunikasi yang sesuai dengan hasil dari *text summarization*.

### 1.6. Sistematika Penulisan

#### BAB I PENDAHULUAN

Pada bab ini akan berisi latar belakang, rumusan masalah, batasan masalah,

tujuan penelitian, metode penelitian dan sistematika penelitian.

## **BAB II TINJAUAN PUSTAKA**

Pada bab ini berisikan penelitian-penelitian terdahulu yang menyangkut dengan penelitian yang dilakukan. Terdapat tabel perbandingan antara penelitian yang dilakukan dengan penelitian-penelitian terdahulu.

## **BAB III LANDASAN TEORI**

Pada bab ini berisikan penjelasan mengenai teori-teori yang menyangkut dengan penelitian yang dilakukan.

## **BAB IV ANALISIS DAN PERANCANGAN EKSPERIMEN**

Pada bab ini berisikan deskripsi problem, analisis kebutuhan eksperimen, dan perancangan eksperimen.

## **BAB V HASIL EKSPERIMEN DAN PEMBAHASAN**

Pada bab ini berisikan deskripsi eksperimen, hasil eksperimen, dan pembahasan eksperimen.

## **BAB VI PENUTUP**

Pada bab ini berisikan kesimpulan dan saran yang didapatkan dari penelitian ini.

## BAB II. TINJAUAN PUSTAKA

Penulisan tugas akhir ini akan menggunakan beberapa referensi yang didapatkan dari penelitian-penelitian yang sebelumnya sudah dilakukan. Penelitian tersebut dijadikan referensi yang digunakan sebagai pedoman penulisan penelitian yang berjudul “*Company Profile Sebuah Hotel Melalui Online Review Menggunakan Text Summarization*”. Penelitian pertama yang dijadikan sebagai pembandingan adalah penelitian yang ditulis oleh Zhan dkk pada tahun 2008 yang berjudul “*Gather customer concerns from online product reviews – A text summarization approach*”. Tujuan dari penelitian ini yaitu mengusulkan pendekatan perangkuman otomatis berdasarkan analisis struktur topik internal artikel ulasan untuk mengumpulkan masalah pelanggan. Penelitian ini nantinya akan menemukan dan mengekstraksi topik-topik yang menonjol dari serangkaian tinjauan online dan selanjutnya merangking topik-topik tersebut. Penelitian ini menggunakan lima dataset dari corpus Hu dan tiga dataset dari Amazon.com dengan setiap dataset berukuran antara 40 hingga 100 dokumen. Berdasarkan pengujian yang telah dilakukan metode yang diusulkan berdasarkan pada struktur topikal berperforma jauh lebih baik dibandingkan dengan *Opinion mining* dan *clustering summarization* [4].

Penelitian selanjutnya yang dijadikan sebagai pembandingan adalah penelitian yang ditulis oleh Hu dkk pada tahun 2017 yang berjudul “*Opinion mining from online hotel reviews- A text summarization approach*”. Tujuan dari penelitian ini yaitu mengusulkan teknik perangkuman multi-teks novel untuk mengidentifikasi kalimat paling informatif dari ulasan hotel. Penelitian ini menggunakan data dari dua hotel yaitu Red Roof Inn dan Gansevoort Meatpacking hotel dari ulasan TripAdvisor.com. Data tersebut dikumpulkan dari 1 Januari 2012 hingga 31 Maret 2013 dengan bahasa yang dipilih yaitu bahasa Inggris dan ulasan yang diposting oleh anggota TripAdvisor. Dari data yang sudah dikumpulkan tersebut dipilih 50 ulasan secara acak untuk digunakan dalam percobaan. Kemudian data tersebut akan diproses dengan menggunakan



tiga metode pendekatan untuk mendapatkan top-k. Pendekatan pertama yaitu memperlakukan semua ulasan menjadi satu dokumen dan hanya skor penting dari kalimat peninjauan yang dipertimbangkan dan semua faktor lainnya diabaikan. Pendekatan kedua yaitu memperlakukan semua ulasan menjadi satu dokumen dan kalimat-kalimat dalam dokumen akan dipartisi kedalam k *cluster* dengan menggunakan algoritma k-medoids. Pendekatan ketiga yaitu mempertimbangkan seluruh aspek yang diusulkan menjadi metode. Sehingga didapatkan hasil bahwa metode ketiga atau metode yang diusulkan memiliki skor tertinggi dibandingkan dengan metode pertama dan kedua dengan nilai hotel Red Roof Inn memiliki skor masing-masing 2,8 dan 2,75 dalam k=5 dan k=10 dan hotel Gansevoort Meatpacking menunjukkan skor masing-masing 2,5 dan 2,75 dengan k=5 dan k=10 [5].

Penelitian selanjutnya yang dijadikan sebagai pembandingan adalah penelitian yang ditulis Alsaqer dan Sasi pada tahun 2017 yang berjudul “*Movie Review Summarization and Sentiment Analysis using RapidMiner*”. Tujuan dari penelitian ini yaitu peningkatan akurasi ringkasan dengan analisis sentimen pada ulasan film menggunakan RapidMiner. Penelitian ini menggunakan data sebanyak 100 file dari 27889 dataset movie-review dalam bentuk file html. Hasil dari penelitian ini yaitu mendapatkan model peringkasan pertama dengan menggunakan ekstensi analisis teks aylien dan model kedua menggunakan ekstensi text processing. Kemudian analisis sentimen dilakukan dengan menggunakan ekstensi analisis teks aylien. Dari dua model peringkasan diatas didapatkan bahwa keakuratan analisis sentimen menggunakan model pertama mencapai 90% dan 96% untuk model kedua [6].

Penelitian selanjutnya yang dijadikan sebagai pembandingan adalah penelitian yang ditulis oleh Yusron dkk pada tahun 2019 yang berjudul “*Peringkasan Review Konsumen Restoran Menggunakan Weighted Frequent Itemset Mining*”. Tujuan dari penelitian ini yaitu penulis ingin mengusulkan suatu pendekatan yang dapat digunakan untuk melakukan peringkasan dengan hasil yang didapatkan tidak hanya memilih kalimat ringkasan berdasarkan bobotnya saja namun juga diharapkan dapat memilih ringkasan yang memiliki kumpulan

kata tertentu. Data yang digunakan dalam penelitian ini merupakan data yang diambil dari *website* [www.tripadvisor.co.id](http://www.tripadvisor.co.id). Penelitian ini menggunakan data 20 restoran dengan setiap restorannya diambil data *review* konsumen sebanyak 50 data. Kemudian data-data tersebut akan dipisah menjadi kalimat-kalimat dan akan dilakukan pemberian skor oleh pakar agar data tersebut dapat digunakan dalam tahapan evaluasi. Berdasarkan data diatas akan dilakukan beberapa pengujian untuk mengetahui bagaimana pengaruh parameter yang terdapat dan digunakan pada metode *Weighted Frequent Itemset Mining*. Pengujian pertama yaitu pengujian terhadap pengaruh nilai *minimum support* terhadap jumlah kalimat hasil ekstraksi, jumlah *weighted frequent itemset* yang dihasilkan dan jumlah anggota *weighted frequent itemset* tertinggi. Pengujian ini menghasilkan rentang nilai *minimum support* antara 0.25 samapi 2 persen sebagai rentang nilai yang baik untuk menghasilkan jumlah kalimat pada setiap restoran. Pengujian kedua yaitu pengujian pada parameter yang menggunakan langkah *preprocessing* untuk penghapusan nilai *stopword*. Dari pengujian ini dapat diketahui bahwa penghapusan *stopword* akan mendapatkan nilai yang kurang baik dibandingkan tanpa melakukan penghapusan *stopword*. Pengujian yang ketiga yaitu pengaruh jumlah kalimat ringkasan. Dari pengujian ini diketahui bahwa jumlah 10 kalimat akan menghasilkan peringkasan terbaik yang sesuai dengan pakar [7].

Penelitian selanjutnya yang dijadikan sebagai pembanding adalah penelitian yang ditulis oleh Savanti dkk pada tahun 2018 yang berjudul “Peringkasan Teks Otomatis Secara Ekstraktif Pada Artikel Berita Kesehatan Berbahasa Indonesia Dengan Menggunakan Metode Latent Semantic Analysis”. Tujuan dari penelitian ini yaitu ingin membantu pengguna dalam mempermudah proses ekstraksi informasi yang terdapat dalam dokumen dengan waktu yang singkat namun mendapatkan hasil jelas. Penelitian ini menggunakan data yang diambil dari *website* berita *online* [kompas.com](http://kompas.com) yang akan disimpan pada 10 dokumen dengan jumlah 20 kalimat. Data tersebut akan terlebih dahulu dilakukan proses *preprocessing* yang meliputi tahapan *parsing*, *tokenisasi*, *filtering* dan *stemming*, lalu akan dilakukan tahapan perhitungan menggunakan

*TF-IDF* untuk mendapatkan nilai bobot dari setiap *term*. Kemudian akan dilakukan perhitungan hasil bobot yang didapatkan dengan menggunakan *latent semantic analysis (LSA)* dengan *singular value decomposition SVD* yang akan mendekomposisi matriks menjadi 3 dan akan diekstraksi hasil ringkasan menggunakan *cross method* untuk mendapatkan hasil dari ringkasan. Pengujian ini menggunakan *precision, recall, f-measure* dan nilai akurasi akan dihitung menggunakan nilai parameter *compression rate*. Dalam pengujian ini dilakukan percobaan sebanyak 6 kali dengan presentase nilai 10%, 20%, 30%, 40% dan 50% pada setiap data yang akan digunakan. Berdasarkan data dan pengujian yang dilakukan diketahui bahwa ringkasan terbaik didapatkan dengan nilai *compression rate* 40% yang diketahui nilai *precision, recall, f-measure* dan akurasi yaitu 0.75, 0.6, 0.667 dan 0.727 dan pada *compression rate* 50% dengan nilai 0.668, 0.743, 0.700 dan 0.690 [8].

Penelitian selanjutnya yang dijadikan sebagai pembanding adalah penelitian yang ditulis oleh Shimpikar dan Govilkar pada tahun 2017 yang berjudul “A Survey of Text Summarization Techniques for Indian Regional Languages”. Tujuan dari penelitian ini yaitu memberikan studi komparatif dari berbagai teknik peringkasan teks yang digunakan untuk bahasa daerah India dan juga membahas dua jenis teknik peringkasan teks ekstraktif dan abstrak. Penelitian ini menggunakan data inputan berbagai bahasa seperti bahasa Bengali, Malayalam, Hindi, Odia, Gujarati, Tamil, Telugu, dll. Penelitian ini menggunakan beberapa algoritma peringkasan teks yaitu *Rule based, Tree based, Ontology based, TFIDF and cluster based*. Hasil yang didapatkan dari melakukan tinjauan pada berbagai jenis pendekatan dan membandingkan metode yang ada berdasarkan berbagai parameter, dapat disimpulkan bahwa TFIDF, berbasis grafik sebagai metode peringkasan teks yang paling efektif untuk bahasa regional India [9].

Penelitian selanjutnya yang dijadikan sebagai pembanding adalah penelitian yang ditulis oleh Chen dkk pada tahun 2009 yang berjudul “Text Summarization for Oil and Gas News Article”. Tujuan dari penelitian ini yaitu menerapkan sistem peringkasan teks untuk menghasilkan ringkasan yang

berisi informasi utama artikel berita minyak dan gas untuk memantau perilaku pesaing mereka dalam merumuskan strategi bisnis. Penelitian ini menggunakan data artikel-artikel berita yang akan dirangkum kedalam empat tingkat kompresi yang berbeda yaitu 25%, 50%, 75% dan 100%. Berdasarkan data tersebut dilakukan pengujian dengan menggunakan pengukuran kinerja presisi dan rasio recall. Hasil yang didapatkan dari pengujian tersebut yaitu pada tingkat kompresi 25% sistem dapat menghasilkan ringkasan yang oheren pada panjang teks yang lebih pendek yaitu pada tingkat kompresi 25% [10].

Penelitian selanjutnya yang dijadikan sebagai pembanding adalah penelitian yang ditulis oleh Lloret dan Palomar pada tahun 2013 yang berjudul "*Towards automatic tweet generation: A comparative study from the text summarization perspective in the journalism genre*". Tujuan dari penelitian ini yaitu ingin menganalisis tugas pembuatan tweet otomatis dari perspektif peringkasan teks dalam konteks genre jurnalisme. Untuk mencapai tujuan tersebut berbagai *summarizer* canggih dipilih dan digunakan untuk menghasilkan *tweet* multi bahasa dalam dua bahasa yaitu inggris dan spanyol. Penelitian ini membuat corpus dari kumpulan dokumen berita baru dari situs surat kabar *online*, tetapi dengan mempertimbangkan bahwa surat kabar *online* tersebut memiliki akun twitter dan harus memungkinkan pembaca untuk secara langsung berbagi artikel berita melalui media sosial twitter. Data yang digunakan dalam penelitian ini yaitu 200 berita untuk setiap bahasa dalam periode waktu selama 10 hari. Hasil yang didapatkan yaitu dapat mengidentifikasi empat jenis tweet yaitu tweet yang dinilai baik sebagai indikasi mengenai topik dan juga menarik, tweet tidak menunjukkan atau menarik, tweet yang topiknya tidak jelas secara langsung, tetapi para pengguna menemukan mereka menarik dan menarik untuk dibaca lebih lanjut, dan tweet yang sangat indikatif, tetapi sebaliknya, tidak ada minat di sana. Kasus-kasus ini mempengaruhi baik tweet otomatis dan tweet asli yang langsung berasal dari berita utama yang dapat digunakan secara berbeda tergantung pada kebutuhan spesifiknya [11].

Penelitian selanjutnya yang dijadikan sebagai pembanding adalah penelitian yang ditulis oleh Kar dkk pada tahun 2015 yang berjudul "*Summarization of*

*changes in dynamic text collections using Latent Dirichlet Allocation model”.*

Tujuan dari penelitian ini yaitu untuk mendapatkan ringkasan yang menggambarkan perubahan paling signifikan yang dibuat untuk dokumen selama periode tertentu. Data yang digunakan dalam penelitian ini yaitu 54 studi kasus untuk 49 artikel wikipedia yang berbeda dalam periode waktu tertentu. Studi kasus yang dipilih berdasarkan kriteria persis ada satu perubahan signifikan yang dibuat pada artikel dalam periode waktu yang dipilih dan perubahan harus diketahui secara apriori. Dalam penelitian ini menggunakan empat pendekatan yang berbeda yang diusulkan untuk meringkas perubahan. Pendekatan pertama yaitu yang secara berkala memonitor koleksi web dalam mencari perubahan terbaru dan membuat ringkasannya sehubungan dengan topik tertentu. Dalam pendekatan kedua setiap aspek temporal istilah diselidiki dengan mempertimbangkan probabilitas gabungan dari kedua peristiwa penyisipan dan penghapusan selama satu set versi dokumen dalam periode yang diberikan. Pendekatan ketiga didasarkan pada model *Latent Dirichlet Allocation (LDA)* untuk menemukan struktur topik tersembunyi atau laten dari perubahan. Pendekatan keempat adalah kombinasi dari dua pendekatan sebelumnya di mana kalimat peringkat teratas yang dihasilkan dari pendekatan ketiga diurutkan ulang dengan skor gabungan dari pendekatan kedua dan ketiga. Seluruh pendekatan diatas akan dievaluasi menggunakan metrik ROUGE dengan membandingkan ringkasan yang dihasilkan sistem dan ringkasan referensi yang dibuat oleh manusia. Dari hasil tersebut didapatkan bahwa pendekatan berdasarkan model LDA mengungguli yang lain [12].

Penelitian selanjutnya yang dijadikan sebagai pembanding adalah penelitian yang ditulis oleh Rouane dkk pada tahun 2019 yang berjudul “*Combine clustering and frequent itemsets mining to enhance biomedical text summarization*”. Tujuan dari penelitian ini yaitu mengusulkan sistem ringkasan teks biomedis yang mengabungkan dua teknik *data mining* yaitu *clustering* dan *frequent itemset mining*. Penelitian ini akan menggunakan algoritma K-means untuk mengelompokkan kalimat yang serupa dan menggunakan algoritma apriori untuk menemukan item set yang sering

diantara kalimat-kalimat yang sudah dikelompokan. Data yang digunakan dalam penelitian ini yaitu 100 makalah biomedis dari database BioMed Central. Ringkasan yang didapatkan nantinya juga akan dibandingkan dengan algoritma peringkasan lainnya yaitu *TextRank*, *TextTeaser*, *SweSum*, *ItemSet Based Summarizer*, *Microsoft AutoSummarize*. Hasil yang didapatkan menunjukan bahwa kombinasi *clustering* dan *frequent itemset mining* berhasil meningkatkan kinerja *summarization*, dan sistem yang diusulkan mengungguli *summarizer* lain yang diuji [13].

Berdasarkan penelitian-penelitian yang sudah dilakukan, penulis ingin membuat analisis yang berkaitan dengan *online review* terhadap *company profile* sebuah hotel melalui *online review* menggunakan *text Summarization*. Penelitian ini dilakukan bertujuan untuk mendapatkan sebuah informasi dari *online review* yang dilakukan oleh konsumen untuk mengetahui *company profile* sebuah hotel menggunakan *text summarization*.

Tabel 1. Perbandingan Penelitian

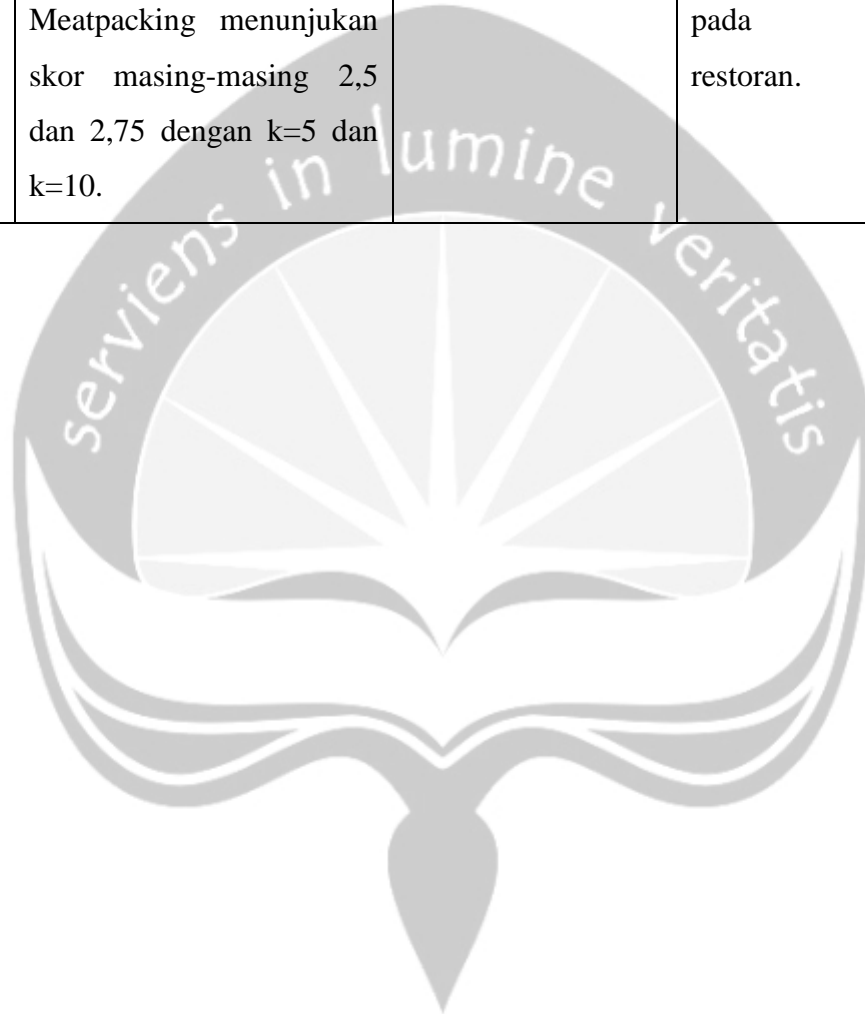
Penulis	Zhan dkk [4]	Hu dkk [5]	Alsaqer dan Sasi [6]	Yusron dkk [7]	Setiawan (2020)*
Judul	<i>Gather customer concerns from online product reviews – A text summarization approach.</i>	<i>Opinion mining from online hotel reviews- A text summarization approach.</i>	<i>Movie Review Summarization and Sentiment Analysis using RapidMiner.</i>	Peringkasan Review Konsumen Restoran Menggunakan <i>Weighted Frequent Itemset Mining.</i>	<i>Company Profile</i> Sebuah Hotel Melalui <i>Online Review</i> Menggunakan <i>Text Summarization.</i>
Tujuan	Mengusulkan pendekatan perangkuman otomatis berdasarkan analisis struktur topik internal artikel ulasan untuk mengumpulkan	Mengusulkan teknik perangkuman multi-teks novel untuk mengidentifikasi kalimat paling informatif dari ulasan hotel.	Peningkatan akurasi ringkasan dengan analisis sentimen pada ulasan film menggunakan RapidMiner.	Ingin mengusulkan suatu pendekatan yang dapat digunakan untuk melakukan peringkasan dengan hasil yang didapatkan tidak hanya memilih kalimat ringkasan	Untuk mendapatkan sebuah informasi dari <i>online review</i> yang dilakukan oleh konsumen untuk mengetahui <i>company profile</i> sebuah hotel menggunakan <i>text summarization.</i>

	masalah pelanggan.			berdasarkan bobotnya saja namun juga diharapkan dapat memilih ringkasan yang memiliki kumpulan kata tertentu.	
Metode	Struktur topical.	K-medoid.	Ekstensi teks aylien dan ekstensi text processing.	<i>Weighted Frequent Itemset Mining.</i>	<i>Extractive Text Summarization</i> menggunakan <i>TF-IDF</i> .
Hasil	Metode yang diusulkan berdasarkan pada struktur topical berperforma jauh lebih baik dibandingkan dengan <i>Opinion</i>	Metode yang diusulkan memiliki skor tertinggi dibandingkan dengan metode pertama dan kedua dengan nilai hotel Red Roof Inn memiliki skor masing-masing 2,8 dan 2,75 dalam k=5 dan k=10	Keakuratan analisis sentimen menggunakan model ekstensi analisis teks aylien mencapai 90% dan 96% untuk ekstensi text processing.	Menghasilkan rentang nilai <i>minimum support</i> antara 0.25 samapi 2 persen sebagai rentang nilai yang baik untuk menghasilkan	Penelitian ini menunjukkan bahwa nilai rata-rata yang didapatkan dalam perhitungan setiap hotel dapat mengambil kalimat yang sering dibahas dalam review setiap hotel dan menyimpulkan ciri-khas



	<i>mining dan clustering summarization..</i>	dan hotel Gansevoort Meatpacking menunjukan skor masing-masing 2,5 dan 2,75 dengan k=5 dan k=10.		jumlah kalimat pada setiap restoran.	masing-masing hotel.
--	--	--	--	--------------------------------------	----------------------

\*) Penelitian yang dilakukan



## BAB III. LANDASAN TEORI

### 3.1 *Text Summarization*

*Text summarization* merupakan proses pengurangan dokumen teks untuk mendapatkan sebuah ringkasan yang mempertahankan poin paling penting dari sebuah dokumen asli [14]. *Text summarization* adalah proses membuat versi pendek dari teks atau dokumen dengan memilih bagian paling penting dan menghasilkan ringkasan yang relevan [15]. Tujuan digunakannya *text summarization* yaitu untuk dapat merepresentasikan makna inti dari sebuah dokumen asli [6]. Secara umum *text summarization* diklasifikasikan menjadi dua kategori utama yaitu metode *extractive summary* dan metode *abstractive summary*. Metode *extractive summary* merupakan peringkasan dengan memilih kalimat yang representatif yang berbetuk kalimat dari dokumen aslinya. Teknik peringkasan *extractive* dapat dilakukan dengan menggunakan beberapa metode yaitu metode berbasis *cluster*, metode *Latent Sematic Analysis* (LSA), metode jaringan saraf tiruan dan lain sebagainya [6]. Dalam metode *extractive summary* kalimat yang akan diekstraksi dari dokumen asli akan digabungkan secara koheren untuk mendapatkan versi yang lebih pendek dari teks aslinya tanpa mengubah makna dari dokumen aslinya. Berikut ini merupakan tahapan proses *Text Summarization* yang akan dilakukan dalam penelitian ini yaitu:

#### a. *Preprocessing*

Pada tahap ini akan dilakukan ekstraksi informasi, di mana kata yang tidak penting akan dihapus sehingga didapatkan hasil akhir yang diperlukan dalam *text summarization*. Pada *preprocessing* ini dilakukan delapan urutan dengan tahapan proses yaitu:

##### 1. *Case Folding*

Proses ini bertujuan untuk mengubah teks menjadi huruf kecil semua agar dapat memudahkan proses tahapan berikutnya.

##### 2. *Remove Emoticon*

Proses ini bertujuan untuk menghapus semua emoticon.

### 3. *Remove* Tanda Baca

Pada tahap ini dilakukan untuk menghilangkan tanda baca selain tanda baca titik (.). Pada tahapan ini menggunakan fungsi *regular expression* (re.sub) untuk menghilangkan tanda baca seperti '?', '!', '@', '#', '%', dan '\$', meskipun tanda baca tersebut terdiri dari satu atau lebih.

### 4. Normalisasi Singkatan dan tanda baca titik-titik

Proses pada tahap ini dilakukan untuk menormalisasikan kata dari singkatan dan menghilangkan tanda baca titik-titik. Pada normalisasi ini menggunakan kamus atau daftar kata yang terdapat pada ('normalisasi.txt'), dimana akan dilakukan pengecekan tiap kata pada teks *review* hotel dengan kamus.

### 5. Menambah Spasi setelah tanda titik (.)

Pada tahapan ini digunakan untuk memberi spasi setelah tanda titik (.)

### 6. *Remove* Angka

Pada tahapan ini digunakan untuk menghapus angka.

### 7. *Stopword Removal*

Proses pada tahap ini dilakukan untuk menghilangkan kata-kata yang tidak diperlukan dalam *review* hotel. Pada *stopword removal* ini menggunakan kamus atau daftar kata yang terdapat pada ('stopword.txt'), dimana akan dilakukan pengecekan tiap kata pada teks *review* hotel dengan kamus yang berisi dengan kata yang tidak diperlukan dan akan menghapusnya.

### 8. Menghapus Spasi Berlebih

Pada tahapan ini digunakan untuk menghilangkan spasi berlebih.

## b. *Text Summarization*

Pada tahap ini akan dilakukan *text summarization* untuk menghasilkan informasi dari dokumen *review* hotel. Pada *text summarization* ini dilakukan dengan urutan tahapan proses yaitu:

### 1. *Tokenize the Sentence*

Membagi dokumen *review* hotel menjadi setiap kalimat (*D*).

2. Menghitung frekuensi kemunculan kata ( $t$ ) di setiap kalimat ( $D$ ).
3. Menghitung *term frequency* ( $TF$ ) yaitu jumlah frekuensi kemunculan kata ( $t$ ) pada kalimat ( $D$ ). Tahap ini akan menghitung *term frequency* ( $TF$ ) jumlah frekuensi kemunculan kata ( $t$ ) dan dibagi dengan total kata ( $t$ ) yang terdapat pada kalimat ( $D$ ).

$$\frac{t}{n(t)}$$

4. Menghitung frekuensi kemunculan kata ( $t$ ) dalam satu dokumen ( $DK$ ).
5. Menghitung *inverse document frequency* ( $IDF$ ) dengan cara log dari jumlah total kalimat ( $D$ ) dalam dokumen *review* hotel dibagi jumlah kalimat yang mengandung kata ( $t$ ) didalamnya dalam dokumen ( $DK$ ).

$$\log \frac{\text{jumlah total kalimat } (D) \text{ dalam dokumen}}{\text{jumlah kemunculan kata } (t) \text{ dalam dokumen}}$$

6. Menghitung  $TF-IDF$  dan menghasilkan matrik dengan mengalikan hasil dari  $TF$  dan  $IDF$ .

$$TF \times IDF$$

7. Menghitung  $TF-IDF$  untuk memberi bobot setiap kalimat ( $D$ ) dengan cara jumlah total skor seluruh kata ( $t$ ) dalam kalimat dibagi dengan jumlah total kata ( $t$ ) dalam kalimat ( $D$ ).

$$\frac{\text{jumlah total skor seluruh kata } (t) \text{ dalam kalimat } (D)}{\text{jumlah total kata } (t) \text{ dalam kalimat } (D)}$$

8. Menghitung skor rata-rata kalimat dengan cara membagi jumlah total nilai seluruh kalimat ( $D$ ) dengan jumlah kalimat ( $D$ ) untuk menentukan batas nilai minimal kalimat yang nantinya digunakan sebagai hasil *summarization*.

$$\frac{\text{jumlah total nilai seluruh kalimat } (D)}{\text{jumlah kalimat } (D)}$$

9. Menentukan hasil *summarization* dengan mengambil kalimat (*D*) yang memiliki nilai lebih besar atau sama dengan nilai rata-rata.
10. Melakukan peringkasan *text summarization* kedua dan ketiga kalinya dengan mengulang langkah 1 hingga 9 untuk mendapatkan hasil yang lebih ringkas.

### **3.2 Company Profile**

*Company profile* adalah sebuah aset dari suatu lembaga atau perusahaan yang dapat digunakan untuk meningkatkan suatu *image* atau citra dari perusahaan yang dapat menjalin kerja sama dengan lembaga, instansi atau perusahaan lainnya [16]. *Company profile* merupakan sebuah aset perusahaan yang dimiliki oleh setiap perusahaan atau lembaga. Selain digunakan sebagai aset perusahaan *company profile* juga berfungsi sebagai tanda atau identitas dari perusahaan dalam menjalin kerjasama yang dilakukan dengan lembaga, instansi atau perusahaan lainnya. *Company profile* adalah sebuah media promosi dan komunikasi perusahaan atau institusi untuk memperkenalkan produk dari perusahaan atau institusi dengan menggunakan media cetak maupun media audio-visual [17]. Kemudian, *Company profile* merupakan hal yang paling dicari oleh masyarakat untuk mengetahui informasi tentang gambaran umum sebuah perusahaan atau institusi. *Company profile* merupakan sebuah data atau *profile* tentang sebuah perusahaan atau sebuah instansi [18]. *Company profile* dibuat untuk ditujukan kepada calon nasabah (*client*) dalam memperkenalkan sebuah perusahaan untuk dapat meyakinkan calon nasabah (*client*) tersebut.

### **3.3 Hotel**

Menurut Peraturan Menteri Pariwisata dan Ekonomi Kreatif Republik Indonesia Nomor PM.106/PW.006/MPEK/2011 Tentang Sistem Manajemen Pengamanan Hotel dijelaskan bahwa hotel adalah penyediaan akomodasi secara harian berupa kamar-kamar di dalam 1 (satu) bangunan, yang dapat

dilengkapi dengan jasa pelayanan makanan dan minuman, kegiatan hiburan serta fasilitas lainnya [19]. Menurut Surat Keputusan Menteri Pariwisata, Pos dan Telekomunikasi No. KM. 94HK. 103MPPT – 87 Tentang Ketentuan Usaha dan Penggolongan Hotel dijelaskan bahwa hotel adalah salah satu jenis akomodasi yang mempergunakan sebagian atau seluruh bangunan untuk menyediakan jasa pelayanan penginapan, makanan dan minuman serta jasa lainnya bagi umum dan dikelola secara komersial [20]. Hotel adalah jenis akomodasi yang dikelola secara komersial dan profesional yang disediakan untuk setiap orang yang ingin mendapatkan suatu pelayanan penginapan (*rooms*), makanan dan minuman (*food and beverage*), hiburan seperti *night club*, *discotheque* dan pelayanan menarik lainnya yang tersedia [21]. Hotel juga dapat didefinisikan sebagai sejenis sarana akomodasi yang menyediakan pelayanan berupa penginapan, makan dan minum serta pelayanan lainnya yang dikelola secara profesional [22]. Hotel juga merupakan bidang usaha yang menggunakan suatu bangunan atau sebagian bangunan yang telah disediakan secara khusus untuk digunakan setiap orang dalam menginap, makan, memperoleh pelayanan dan juga menggunakan fasilitas lainnya dengan dikenakan pembayaran.

### **3.4 Online Review**

Ulasan konsumen merupakan sebuah ungkapan kata dari mulut ke mulut yang memiliki dampak besar pada pengambilan sebuah keputusan [23]. Ulasan yang dilakukan oleh seorang konsumen lebih dapat diandalkan dibandingkan informasi yang diberikan oleh seorang penjual karena mereka menawarkan saran yang dipersonalisasi. Selain itu *Online review* memiliki dampak yang signifikan bagi konsumen dan pengaruh terhadap pembelian karena mereka dapat berbagi informasi dengan orang banyak pada saat yang sama tanpa dibatasi oleh ruang dan waktu [24]. Konsumen pada era globalisasi ini pada umumnya akan meneliti *Online review* terlebih dahulu untuk dapat mengetahui kualitas dari sebuah produk sebelum melakukan pembelian. Selain itu konsumen juga akan mencari informasi melalui ulasan konsumen lain sebelum

memutuskan untuk membeli suatu produk. Sehingga ulasan konsumen ini merupakan sebuah informasi yang penting bagi seseorang yang mungkin meragukan kualitas barang yang akan dibelinya jika hanya melihat informasi produk yang dibuat oleh penjual [2].

Pemasaran dari mulut ke mulut berbasis internet atau *Electronic Word of Mouth* (eWOM) seperti *online review* dan pendapat seseorang sudah diakui sebagai salah satu saluran komunikasi yang paling berpengaruh di antara penyedia layanan dan konsumen itu sendiri. Seiring dengan berkembangnya teknologi dan juga sosial media, industri perhotelan mulai dipengaruhi oleh (eWOM) [5]. Sebagai informasi yang dibuat oleh seorang konsumen, *online review* cenderung lebih relevan bagi konsumen daripada informasi yang dibuat oleh seorang penjual. Informasi produk yang dibuat oleh seorang penjual lebih menggambarkan spesifikasi produk dan kinerja produk yang sesuai dengan standar teknis dari pihak penjual. Sebaliknya informasi produk yang diciptakan oleh konsumen didasarkan atas definisi yang menggambarkan atribut produk ketika dalam situasi penggunaan dan mengukur kinerja produk ketika digunakan [24].

## BAB VI. PENUTUP

### 6.1. Kesimpulan

Berdasarkan penelitian *text summarization* menggunakan metode *TF-IDF* pada *review* Bahasa Indonesia Hotel Royal Ambarrukmo Yogyakarta, Hotel THE 101 Yogyakarta, Hotel Novotel Yogyakarta, Hotel Grand Tjokro Yogyakarta dan Hotel Tentrem dapat menunjukkan bahwa nilai rata-rata yang didapatkan dalam perhitungan setiap hotel dapat mengambil kalimat yang sering dibahas dalam *review* setiap hotel dan menyimpulkan ciri-khas masing-masing hotel tersebut. Berdasarkan hasil informasi dari *text summarization*, ciri-khas masing-masing hotel tersebut akan dilakukan pembahasan untuk melakukan analisis *company profile* sebuah hotel yang dilakukan oleh seorang konsumen. Komentar-komentar pelanggan di ranah publik, akan memberikan gambaran mengenai *profiling* dari sebuah organisasi. Dengan demikian hal utama yang perlu diperhatikan dalam menciptakan profil organisasi yang positif adalah memahami pelanggan. Mencatat setiap komentar yang ada di ruang publik merupakan salah satu upaya memahami pelanggan. Selanjutnya, dari komentar yang pada akhirnya memberikan gambaran profil secara sangat spesifik, akhirnya akan membentuk *positioning*. Dengan demikian lahirlah ciri dari sebuah hotel, berdasarkan penilaian pelanggan.

### 6.2.Saran

Setelah dilakukan penelitian, penulis memiliki beberapa saran untuk penelitian berikutnya yang berkaitan dengan *Company Profile* dan *Text Summarization* yaitu:

1. Disarankan untuk mengembangkan formula metode *text summarization* yang lebih baik. Diharapkan dengan banyak metode *text summarization* untuk dokumen Bahasa Indonesia nantinya dapat dihasilkan fitur *text summarization* untuk dokumen Bahasa Indonesia yang lebih baik.
2. Disarankan untuk mengembangkan metode *Text Summarization* yang dapat menghasilkan informasi secara jelas untuk digunakan sebagai analisis *company profile*.



3. Penelitian *company profile* dapat dilakukan pada media sosial lainnya selain *TripAdvisor* dan perusahaan lainnya selain perhotelan.



## DAFTAR PUSTAKA

- [1] F. V. Sudjarmika, "Pengaruh Harga, Ulasan Produk, Kemudahan, dan Keamanan Terhadap Keputusan Pembelian Secara Online di Tokopedia.com," *Agora*, vol. 5, no. 1, pp. 1–7, 2017.
- [2] Y. Chen and J. Xie, "Online consumer review: Word-of-mouth as a new element of marketing communication mix," *Manage. Sci.*, vol. 54, no. 3, pp. 477–491, 2008.
- [3] M. A. Asri Nugrahani Ardianti, Dr. Widiartanto, "Pengaruh Online Customer Review dan Online Customer Rating terhadap Keputusan Pembelian melalui Marketplace Shopee .," *J. Ilmu Adm. Bisnis*, pp. 1–11, 2019.
- [4] J. Zhan, H. T. Loh, and Y. Liu, "Gather customer concerns from online product reviews - A text summarization approach," *Expert Syst. Appl.*, vol. 36, no. 2 PART 1, pp. 2107–2115, 2008.
- [5] Y. H. Hu, Y. L. Chen, and H. L. Chou, "Opinion mining from online hotel reviews – A text summarization approach," *Inf. Process. Manag.*, vol. 53, no. 2, pp. 436–449, 2017.
- [6] A. F. Alsaqer and S. Sasi, "Movie review summarization and sentiment analysis using rapidminer," *2017 Int. Conf. Networks Adv. Comput. Technol. NetACT 2017*, pp. 329–335, 2017.
- [7] M. I. Yusron, F. A. Bachtiar, and M. A. Fauzi, "Peringkasan Review Konsumen Restoran Menggunakan Weighted Frequent Itemset Mining," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 5, pp. 4961–4970, 2019.
- [8] N. Savanti, W. Gotami, and R. K. Dewi, "Peringkasan Teks Otomatis Secara Ekstraktif Pada Artikel Berita Kesehatan Berbahasa Indonesia Dengan Menggunakan Metode Latent Semantic Analysis," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 9, pp. 2821–2828, 2018.
- [9] S. Shimpikar and S. Govilkar, "A Survey of Text Summarization Techniques for Indian Regional Languages," *Int. J. Comput. Appl.*, vol. 165, no. 11, pp. 29–33, 2017.
- [10] Y. Y. Chen, O. M. Foong, S. P. Yong, and K. Iwan, "Text Summarization for Oil and Gas Drilling Topic," vol. 3, no. 5, pp. 31–34, 2009.
- [11] E. Lloret and M. Palomar, "Towards automatic tweet generation: A comparative study from the text summarization perspective in the journalism genre," *Expert Syst. Appl.*, vol. 40, no. 16, pp. 6624–6630, 2013.
- [12] M. Kar, S. Nunes, and C. Ribeiro, "Summarization of changes in dynamic text collections using Latent Dirichlet Allocation model," *Inf. Process. Manag.*, vol. 51, no. 6, pp. 809–833, 2015.
- [13] O. Rouane, H. Belhadef, and M. Bouakkaz, "Combine clustering and frequent itemsets mining to enhance biomedical text summarization," *Expert Syst. Appl.*, vol. 135, pp. 362–373, 2019.
- [14] N. ALAMI, M. MEKNASSI, and N. RAIS, "Automatic Texts Summarization: Current State of the Art," *J. Asian Sci. Res.*, vol. 5, no. 1, pp. 1–15, 2015.
- [15] M. A. Fattah and F. Ren, "GA, MR, FFNN, PNN and GMM based models

- for automatic text summarization,” *Comput. Speech Lang.*, vol. 23, no. 1, pp. 126–144, 2009.
- [16] Maimunah, L. Sunarya, and Nina Larasati, “MEDIA COMPANY PROFILE SEBAGAI SARANA PENUNJANG INFORMASI DAN PROMOSI,” vol. 5, no. 40, pp. 281–301, 2012.
- [17] Budiman, M. Adri, and D. Irfran, “Perancangan Media Presentasi Company Profile Universitas Negeri Padang Berbasis Multimedia Interaktif,” *Vokasional Tek. Elektron. Inform.*, vol. 4, no. 2, p. 50, 2016.
- [18] R. A. Rina and A. F. Nuryana, “Pembuatan Company Profil Berbasis Multimedia Flash PD BPR Bank Daerah Karanganyar,” *J. Speed – Sentra Penelit. Eng. dan Edukasi*, vol. 4, no. 2, pp. 8–13, 2012.
- [19] M. Pariwisata dan Ekonomi Kreatif RI, “Peraturan Menteri Pariwisata dan Ekonomi Kreatif Republik Indonesia Nomor PM.106PW.006MPEK2011 Tentang Sistem Manajemen Pengamanan Hotel,” *Sist. Manaj. Pengamanan Hotel*, vol. 2008, p. 8, 2011.
- [20] “Surat Keputusan Menteri Pariwisata, Pos dan Telekomunikasi No. KM. 94HK. 103MPPT– 87 Tentang Ketentuan Usaha dan Penggolongan Hotel.”
- [21] H. Hermawan, E. Brahmanto, and F. Hamzah, “Pengantar Manajemen Hospitality,” *Pengantar Manaj. Hosp.*, no. September, p. 205, 2018.
- [22] U. E. Wardhani, Viverawati, and Mustafa, *Usaha Jasa Pariwisata Jilid 1*. 2008.
- [23] S. G. Kim and J. Kang, “Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews,” *Inf. Process. Manag.*, vol. 54, no. 6, pp. 938–957, 2018.
- [24] B. Bickart and R. M. Schindler, “Internet Forums As Influential,” *J. Interact. Mark.*, vol. 15, no. 3, pp. 31–40, 2001.