

**PERANCANGAN METODOLOGI *TEXT MINING*
UNTUK MENENTUKAN *CRITICAL-TO-QUALITY*
PRODUK BERDASARKAN ULASAN KONSUMEN**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat guna mencapai gelar
Sarjana dalam bidang ilmu Teknik Industri

Disusun oleh:

Nama : Lusienna Hung

NPM : 6131901124



**PROGRAM STUDI SARJANA TEKNIK INDUSTRI
JURUSAN TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
BANDUNG
2023**

**PERANCANGAN METODOLOGI *TEXT MINING*
UNTUK MENENTUKAN *CRITICAL-TO-QUALITY*
PRODUK BERDASARKAN ULASAN KONSUMEN**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat guna mencapai gelar
Sarjana dalam bidang ilmu Teknik Industri

Disusun oleh:

Nama : Lusienna Hung

NPM : 6131901124



**PROGRAM STUDI SARJANA TEKNIK INDUSTRI
JURUSAN TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
BANDUNG
2023**

**FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
BANDUNG**



Nama : Lusienna Hung
NPM : 6131901124
Program Studi : Sarjana Teknik Industri
Judul Skripsi : PERANCANGAN METODOLOGI *TEXT MINING*
UNTUK MENENTUKAN *CRITICAL-TO-QUALITY*
PRODUK BERDASARKAN ULASAN KONSUMEN

TANDA PERSETUJUAN SKRIPSI

Bandung, 9 Februari 2023
**Ketua Program Studi Sarjana
Teknik Industri**

(Dr. Cecilia Tesavrita, S.T., M.T.)

Pembimbing Pertama

(Hanky Fransiscus, S.T., M.T.)

Pembimbing Kedua

(Dedy Suryadi, S.T., M.S., Ph.D.)

PERNYATAAN TIDAK MENCONTEK ATAU MELAKUKAN PLAGIAT

Saya yang bertanda tangan di bawah ini,

Nama : Lusienna Hung

NPM : 6131901124

dengan ini menyatakan bahwa Skripsi dengan Judul:
PERANCANGAN METODOLOGI *TEXT MINING* UNTUK MENENTUKAN
CRITICAL-TO-QUALITY PRODUK BERDASARKAN ULASAN KONSUMEN

adalah hasil pekerjaan saya dan seluruh ide, pendapat atau materi dari sumber lain telah dikutip dengan cara penulisan referensi yang sesuai.

Pernyataan ini saya buat dengan sebenar-benarnya dan jika pernyataan ini tidak sesuai dengan kenyataan, maka saya bersedia menanggung sanksi yang akan dikenakan kepada saya.

Bandung, 9 Februari 2023



Lusienna Hung

NPM : 6131901124

ABSTRAK

Six Sigma merupakan salah satu metode untuk mempertahankan dan memperbaiki mutu produk. Dalam penggunaan *Six Sigma*, terdapat *Critical-to-Quality* (CTQ), yaitu karakteristik penting dari suatu produk ataupun jasa yang dapat diukur dan menjamin kepuasan pelanggan. Saat ini sudah digunakan beberapa metode yang digunakan untuk menentukan CTQ, tapi metode-metode tersebut cenderung berbasis intuisi pembuat keputusan. Selain itu, data yang digunakan untuk penentuan CTQ diperoleh dengan wawancara dan kuesioner. Teknologi *text mining* memiliki potensi sebagai metode alternatif penentuan CTQ. Dengan adanya *text mining*, penentuan CTQ dapat menggunakan data *online review* yang sudah tersedia. Penentuan CTQ juga dapat dibuat lebih kuantitatif dengan adanya perhitungan koherensi pemodelan topik LDA dan nilai sentimen analisis sentimen yang tersedia dalam *text mining*. Pemilihan karakteristik produk yang penting memanfaatkan korelasi *rating* dan sentimen, sehingga penentuan kepentingan karakteristik tidak lagi mengandalkan intuisi. Hasil penelitian menemukan adanya karakteristik durasi penggunaan produk sebelum produk tidak berfungsi, besar suara yang diblokir oleh *noise cancellation*, ada atau tidaknya suara yang keluar dari *microphone* saat fitur *mute* menyala dan sebaliknya, besar suara yang ditangkap *microphone*, dan durasi penggunaan *microphone* sebelum *microphone* tidak berfungsi. Karakteristik-karakteristik tersebut dinilai sebagai karakteristik yang penting berdasarkan koefisien korelasi. Seluruh karakteristik tersebut memiliki koefisien korelasi dalam rentang 0,3 sampai 0,5. Oleh karena itu, disimpulkan bahwa *text mining* dapat digunakan untuk penentuan CTQ. Namun, perbaikan dan perkembangan masih bisa dilakukan untuk menghasilkan metodologi yang lebih baik.

ABSTRACT

Six Sigma is a method to maintain and improve product quality. In using this method, there are important, quantitative characteristics of a product or service that guarantee customer satisfaction. These characteristics are called Critical-to-Quality (CTQ). Nowadays, there are several methods to determine CTQ. However, these methods tend to depend on an intuition of a decision maker. Additionally, data used in determining CTQ is obtained through interview and questionnaire. By using text mining, available online review data is able to be used in determining CTQ. CTQ determination will also be more quantitative with computation of LDA topic modelling coherence and sentiment analysis score, which are both available in text mining. Rather than depending on intuition, choosing important product characteristics now utilizes correlation between rating and sentiment. Research found characteristics product usage duration before product stopped functioning, volume of sound that could be blocked by noise cancellation, existence of sound produced by microphone when mute feature is used and vice versa, volume of sound that could be picked up by microphone, and usage duration of microphone before microphone stopped functioning extracted from consumer's reviews. These characteristics are deemed important based on correlation coefficient with range of value between 0.3 to 0.5. In conclusion, the use of text mining in determining CTQ is possible. However, improvement is still needed for this methodology.

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa karena atas berkat-Nya laporan skripsi berjudul “Perancangan Metodologi *Text Mining* Untuk Menentukan *Critical-to-Quality* Produk Berdasarkan Ulasan Konsumen” diselesaikan tepat waktu. Laporan ini merupakan salah satu syarat untuk mencapai gelar Sarjana Teknik Industri di Universitas Katolik Parahyangan. Laporan ini tidak akan selesai tanpa bantuan banyak pihak. Penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya, terutama kepada pihak-pihak berikut ini.

1. Bapak Hanky Fransiscus, S.T., M.T., dan Bapak Dedy Suryadi, S.T., M.S., Ph.D. selaku dosen pembimbing penulis yang telah membimbing penulis selama melakukan penelitian.
2. Keluarga penulis yang telah memberikan dukungan dan bimbingan, tidak hanya dalam penyusunan skripsi ini, tapi juga dalam perkuliahan.
3. Teman-teman kelas A, khususnya Fiona Stella Pricilia, Patricia Delinda Hiunedia, Chaterine Nathallia, Caleen Belinda, Rio Hananto, Ignatius Ferdi Setyadi, Chairunissa Hasna Alifah, Soffy Nur Hanifah, Vari Noor Laksono, dan Hieronimus Matthew Ekaristiano yang telah bersama-sama menempuh perkuliahan dan praktikum selama 7 semester di Teknik Industri Universitas Katolik Parahyangan.
4. Livia Nathania dan Cornelius yang telah banyak memberikan bantuan kepada penulis selama menjalani perkuliahan.
5. Ainun Saliha Aldian, Iretta Rengganis Kamerlangga, Gabriella Angelica, dan Eugenia Mentari yang telah memberikan semangat dan dukungan untuk penulis dalam penyusunan skripsi.
6. Hans Adi Kurniawan yang telah memberikan dukungan dalam menjalani perkuliahan dan penyusunan skripsi.
7. Teman-teman Teknik Industri Angkatan 2019 atas kebersamaannya.
8. Pihak-pihak lain yang turut membantu penyusunan skripsi dan tidak dapat disebutkan satu per satu.

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada seluruh pihak tersebut atas doa, bimbingan, bantuan, dan dukungan yang telah diberikan. Tanpa dukungan dari pihak-pihak tersebut, sulit untuk menyelesaikan laporan skripsi ini dengan baik. Penulis menyadari bahwa skripsi ini masih memiliki banyak kekurangan, dan memohon maaf atas berbagai kekurangan dan kesalahan yang terdapat dalam laporan skripsi ini. Penulis berharap laporan skripsi ini dapat memberikan pengetahuan baru dan inspirasi bagi pembaca, terutama dalam bidang *text mining* dan *quality control*. Selamat membaca.

Bandung, 10 Januari 2023

Lusienna Hung

DAFTAR ISI

ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR TABEL	vii
DAFTAR GAMBAR	ix
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN	I-1
I.1 Latar Belakang Masalah.....	I-1
I.2 Identifikasi Masalah.....	I-3
I.3 Pembatasan Masalah dan Asumsi Penelitian.....	I-9
I.4 Tujuan Penelitian	I-9
I.5 Manfaat Penelitian	I-9
I.6 Metodologi Penelitian.....	I-10
I.7 Sistematika Penulisan.....	I-11
BAB II TINJAUAN PUSTAKA	II-1
II.1 <i>Text Mining</i>	II-1
II.2 <i>Voice of Customer</i>	II-2
II.3 <i>Critical to Quality</i>	II-3
II.4 Cacat Produk	II-4
II.5 <i>Latent Dirichlet Allocation</i>	II-4
II.6 Analisis Sentimen.....	II-6
II.7 Penelitian Candra (2021)	II-6
BAB III PENGUMPULAN DAN PENGOLAHAN DATA	III-1
III.1 Data	III-1
III.2 Pemodelan Topik LDA	III-1
III.3 Interpretasi Atribut.....	III-6
III.4 Analisis Sentimen.....	III-12
III.5 Uji Distribusi dan Uji Korelasi	III-15
III.6 Perolehan CTQ	III-17

BAB IV ANALISIS	IV-1
IV.1 Pra-Pemrosesan	IV-1
IV.2 Pemodelan Topik	IV-3
IV.3 Interpretasi Atribut	IV-4
IV.4 Analisis Sentimen dan Uji Korelasi	IV-5
IV.5 Aplikasi <i>Text Mining</i> CTQ	IV-6
IV.6 Analisis Kekurangan Penelitian	IV-7
BAB V KESIMPULAN DAN SARAN	V-1
V.1 Kesimpulan	V-1
V.2 Saran	V-1
DAFTAR PUSTAKA	
LAMPIRAN	
RIWAYAT HIDUP PENULIS	

DAFTAR TABEL

Tabel II.1 Cuplikan Data Hasil <i>Web Scraping</i> Candra (2021).....	II-8
Tabel II.2 Aspek Produk Candra (2021) Menggunakan LDA.....	II-8
Tabel II.3 Perbandingan Aspek Produk Candra (2021) dengan NMF dan LDA	II-9
Tabel III.1 20 Kata dengan Frekuensi Tertinggi	III-3
Tabel III.2 Nilai <i>Coherence</i> dengan Parameter <i>Chunksize</i> 1.000-25.000.....	III-5
Tabel III.3 Nilai <i>Coherence</i> dengan Parameter <i>Chunksize</i> 30.000-90.000	III-5
Tabel III.4 Topik dan Kata-kata per Topik	III-5
Tabel III.5 Kata Benda Setiap Topik	III-6
Tabel III.6 Contoh Kalimat Kata-kata Benda Umum dan Spesifik.....	III-7
Tabel III.7 Kata Umum dan Spesifik.....	III-8
Tabel III.8 Kata-kata Benda Kualitatif dan Kuantitatif	III-9
Tabel III.9 Contoh Kalimat dari Kata Kualitatif dan Kuantitatif	III-9
Tabel III.10 Contoh Kalimat dari Kata Bukan Produksi dan Produksi	III-10
Tabel III.11 Kata-kata Benda Bukan Produksi dan Produksi.....	III-11
Tabel III.12 Kata Kunci Atribut dan Atribut	III-11
Tabel III.13 Contoh Kalimat dari Kata Kunci Atribut	III-12
Tabel III.14 Pemecahan Komentar Menjadi Kalimat	III-13
Tabel III.15 Pemilihan Kalimat yang Mengandung Kata Atribut.....	III-13
Tabel III.16 Cuplikan Hasil Analisis Sentimen.....	III-14
Tabel III.17 Hasil Uji Shapiro-Wilk.....	III-16
Tabel III.18 Hasil Uji Spearman Ranked	III-16
Tabel III.19 CTQ yang Diperoleh	III-17

DAFTAR GAMBAR

Gambar I.1 Pengaruh dari Perbaikan Mutu.....	I-1
Gambar I.2 Metodologi Penelitian Shankar et al. (2022).....	I-4
Gambar I.3 Metodologi Penelitian Rajendran et al. (2022).....	I-6
Gambar I.4 Metodologi Penelitian Kim (2018).....	I-7
Gambar I.5 Metodologi Penelitian.....	I-10
Gambar II.1 Metodologi Penelitian Candra (2021)	II-7
Gambar III.1 Hasil Penghapusan Karakter.....	III-2
Gambar III.2 Hasil Pemecahan Kalimat dan Pengubahan Huruf.....	III-2
Gambar III.3 Hasil Penghapusan <i>Stop Words</i>	III-2
Gambar III.4 Hasil Lematisasi.....	III-3
Gambar III.5 Metodologi <i>Text Mining</i> Hasil Penelitian.....	III-18

DAFTAR LAMPIRAN

LAMPIRAN A : CONTOH KALIMAT KATA BENDA

LAMPIRAN B : *CODING* PEMODELAN TOPIK

LAMPIRAN C : *CODING* ANALISIS SENTIMEN, UJI DISTRIBUSI, DAN UJI
KORELASI

BAB I

PENDAHULUAN

Bab I membahas mengenai latar belakang dilakukannya penelitian dan masalah yang diangkat dalam penelitian ini. Pertimbangan pemilihan masalah juga dijelaskan dalam bab ini. Selain itu, bab ini menjelaskan mengenai asumsi dan batasan penelitian, rumusan masalah dari masalah yang teridentifikasi, tujuan penelitian, metodologi penelitian yang dilakukan, dan sistematika penulisan hasil penelitian.

I.1 Latar Belakang Masalah

Mutu produk atau jasa dapat dikatakan sebagai kecocokan atau *fitness* dari suatu produk maupun jasa terhadap spesifikasi penggunaan yang diinginkan pelanggan atau bahkan melebihi spesifikasi tersebut (Mitra, 2016). Mutu produk sendiri tidak hanya menjadi spesifikasi ataupun ukuran kecocokan sebuah produk, tapi juga memiliki pengaruh terhadap pasar dan perusahaan. Gambar I.1 menunjukkan pengaruh-pengaruh dari peningkatan mutu produk.



Gambar I.1 Pengaruh dari Perbaikan Mutu
(Sumber : (Mitra, 2016))

Perbaikan mutu memberikan dampak peningkatan *market share*, posisi kompetitif, dan *profit* bagi sebuah perusahaan (Mitra, 2016). Gambar 1 juga menunjukkan keuntungan lain dari perbaikan mutu. Secara langsung, perbaikan mutu akan mengurangi biaya, meningkatkan efisiensi, dan meningkatkan kepuasan pelanggan. Turunnya biaya yang diperlukan berdampak pada peningkatan rasio harga/biaya. Sedangkan meningkatnya efisiensi akan menyebabkan peningkatan kapasitas dan hasil produksi. Meningkatnya kepuasan pelanggan akan mengakibatkan peningkatan *market share* serta rasio harga/biaya. Kemudian, peningkatan rasio harga/biaya, kapasitas dan hasil produksi, serta *market share* akan meningkatkan posisi kompetitif. Posisi kompetitif sendiri adalah pembeda suatu bisnis, produk, ataupun jasa, dibandingkan dengan kompetitor (CB Insights, n.d.). Pembeda inilah yang menjadi alasan bagi pelanggan untuk lebih memilih bisnis, produk, ataupun jasa tertentu dibandingkan kompetitor lain di pasaran.

Dengan begitu, mutu produk menjadi variabel yang berpengaruh dalam menentukan kesuksesan suatu perusahaan. Mutu produk sendiri terdiri dari beberapa komponen. Garvin (1987) mengusulkan delapan dimensi mutu produk, yaitu kinerja (*performance*), fitur, realibilitas, kesesuaian (*conformance*), daya tahan, kemampuan untuk diperbaiki (*serviceability*), estetika, dan *perceived quality*.

Pertahanan dan perbaikan aspek-aspek mutu produk tersebut dapat dilakukan dengan berbagai cara, salah satunya adalah metode *Six Sigma*. *Six Sigma* DMAIC terdiri dari 5 tahap, yaitu *define*, *measure*, *analyze*, *improve*, dan *control*.

Critical-to-Quality (CTQ) digunakan dalam *Six Sigma* untuk menentukan ukuran dari kinerja proses produksi saat ini (Bertolaccini, Viti, & Terzi, 2015). CTQ adalah karakteristik kunci dari sebuah produk yang dapat diukur, dimana performansi karakteristik yang bersangkutan harus memenuhi kebutuhan pembeli (He, Tang, & Chang, 2010). Sedangkan menurut Roderick, Govindarajan, & Zrymiak (2015) CTQ adalah karakteristik sebuah produk maupun jasa yang krusial dalam menjamin kepuasan pelanggan.

Terdapat berbagai metode yang saat ini telah digunakan untuk memperoleh CTQ. Namun, seiring perkembangan teknologi dan internet, maka ada potensi untuk mengembangkan metode penentuan CTQ dengan

pemanfaatan teknologi *text mining*. *Text mining* adalah identifikasi serta ekstraksi informasi dari data yang tidak terstruktur (Kao & Poteet, 2007). Teknologi *text mining* didukung dengan ketersediaan data *review* yang dapat diakses dengan mudah di internet dapat menghasilkan sebuah metode penentuan CTQ alternatif.

I.2 Identifikasi Masalah

Subbab latar belakang telah menjelaskan pentingnya mutu, CTQ dan potensi implementasi *text mining*. Terdapat beberapa teknik pengumpulan data untuk penentuan CTQ. Namun, Kim (2018) mengemukakan bahwa metode penentuan CTQ pada umumnya masih berbasis intuisi dari pengalaman penentu CTQ.

Contohnya Yulianto & Faritsy (2015) yang menggunakan metode Kano untuk memperoleh penilaian produk dan penentuan CTQ dengan penyusunan CTQ *tree* bersama manajemen perusahaan. Pada penyusunan CTQ *tree* yang dilakukan dengan manajemen perusahaan, sehingga atribut produk yang dianggap penting berasal dari perspektif manajemen. Penentuan CTQ tersebut dilakukan berdasarkan intuisi manajemen perusahaan.

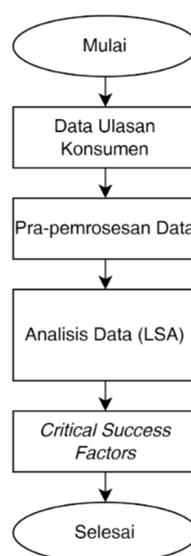
Tak hanya metode penentuan CTQ dari data yang sudah ada, cara perolehan informasi mengenai kebutuhan pelanggan juga tidak kalah penting. Informasi kebutuhan pelanggan penting karena tingkat mutu yang harus dipertahankan oleh perusahaan bergantung oleh kebutuhan dari pelanggan (Mitra, 2016). Kebutuhan pelanggan tersebut merupakan bagian dari *Voice of the Customer* (VOC), yaitu proses mendapatkan informasi-informasi yang berkaitan dengan pelanggan (Munro, Ramu, & Zrymiak, 2015). Selain menampung informasi yang berkaitan dengan pelanggan, VOC dibutuhkan dalam penentuan CTQ (Taghizadegan, 2006). Sayangnya, saat ini pengumpulan data pelanggan secara umum masih menggunakan wawancara dan kuesioner. Contohnya adalah Zhang (2017) yang menggunakan kuesioner dan wawancara berskala kecil untuk pengumpulan data yang setelahnya diubah menjadi CTQ. Adapun Hutabarat, Triswanti, & Septiari (2021) yang menggunakan kuesioner dalam penentuan CTQ minyak goreng.

Dalam implementasi *text mining*, VOC tidak lagi menggunakan data wawancara dan kuesioner, tapi memanfaatkan *online review* dari produk. *Review* tersebut dapat berasal dari berbagai *platform*, terutama *marketplace* seperti

Amazon, Tokopedia, dan Shopee. Pemanfaatan *online review* memiliki keunggulan tersendiri, yaitu data *review* yang tersedia secara terbuka dan dapat diakses dengan mudah. *Text mining* mampu memanfaatkan data yang sudah tersedia, yaitu *online review*, tanpa harus dilakukan pengumpulan data dari nol seperti pada kuesioner dan wawancara.

Keuntungan dari pemanfaatan *online review* sebagai VOC berdampak terhadap kemudahan pembaruan CTQ terhadap perubahan kebutuhan konsumen. Kano (2003) mengemukakan adanya siklus hidup dari kualitas atraktif. Seiring berjalannya waktu, atribut suatu produk yang dikategorikan sebagai atraktif oleh konsumen berubah menjadi atribut *must-be*. Tentunya perubahan atribut tersebut berdampak terhadap CTQ produk. Adanya *online review* menunjukkan secara langsung perubahan permintaan konsumen karena ulasan *online* dapat ditulis secara terbuka tanpa batasan waktu, berbeda dengan kuesioner atau wawancara yang dilakukan secara terbatas dalam satu periode tertentu.

Terdapat beberapa penelitian terdahulu lainnya yang juga menggunakan *text mining* untuk mengolah ulasan konsumen. Meskipun semua penelitian ini menggunakan *text mining* berdasarkan ulasan konsumen, terdapat perbedaan dari penelitian terdahulu dengan penelitian yang akan dilakukan. Perbedaan dapat berupa perbedaan dari segi alat atau metode *text mining*, jenis objek penelitian, dan hasil akhir dari *text mining*.



Gambar 1.2 Metodologi Penelitian Shankar et al. (2022)

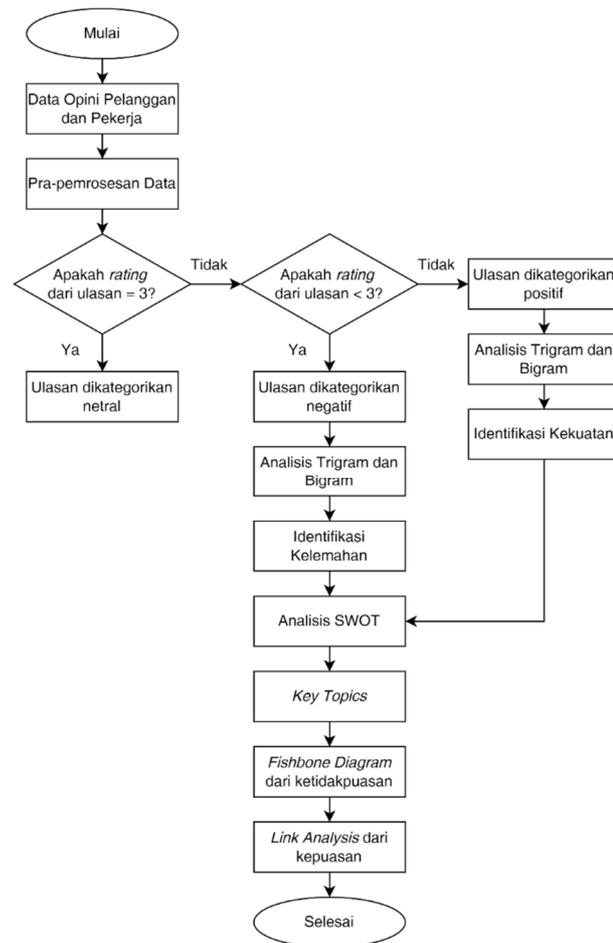
Gambar 1.2 menunjukkan metodologi pengolahan data *text mining* dari Shankar, Tiwari, & Gupta (2022). Penelitian tersebut bertujuan untuk mendapatkan faktor *critical success* dari aplikasi *mobile banking*. Setelah dilakukan *Latent Sentiment Analysis* (LSA), diperoleh faktor-faktor *critical success* berupa navigasi, *customer support*, privasi dan keamanan, efisiensi, dan *convenience*.

Penelitian Shankar et al. (2022) juga menyertakan kata-kata yang sering muncul dengan faktor-faktor *critical success* yang bersangkutan. Contohnya, kata *interface*, atraktif, dan ukuran memiliki asosiasi dengan faktor navigasi. Namun, hasil penelitian Shankar et al. (2022) berbeda dengan CTQ. CTQ adalah karakteristik bentuk kuantitatif, sedangkan faktor-faktor pada Shankar et al. (2022) masih bisa dirincikan menjadi CTQ.

Penelitian lain yang menggunakan ulasan konsumen dan *text mining* adalah penelitian Rajendran, Srinivas, & Pagel (2022). Penelitian ini ulasan dari 2 sumber, yakni konsumen dan pekerja perusahaan asuransi untuk memperoleh topik-topik kunci dari ulasan. Gambar 1.3 berisi metodologi yang digunakan oleh Rajendran et al. (2022).

Penelitian Rajendran et al. (2022) memisahkan data menjadi netral, negatif, dan positif sesuai *rating*. Topik kunci atau *key topic* diperoleh dari analisis trigram dan bigram dari data negatif dan positif. Topik kunci pada penelitian ini juga disertai dengan deskripsi dan kata-kata yang asosiasinya tinggi dengan topik tersebut berdasarkan data ulasan. Beberapa topik yang diperoleh adalah fitur, pembayaran, dan *client assistance*. Fitur kemudian dideskripsikan sebagai atribut yang menambah nilai kepada pelayanan asuransi. Kata-kata signifikan yang diasosiasikan dengan atribut ini adalah aplikasi *website* dan penggantian mobil. Setelah perolehan topik, deskripsi, dan kata-kata signifikan, atribut-atribut ini tidak diubah menjadi bentuk kuantitatif.

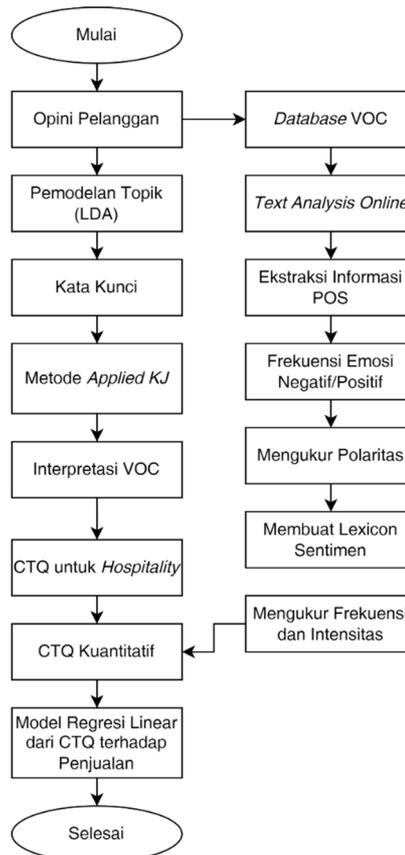
Penelitian Rajendran et al. (2022) berfokus pada topik atau karakteristik yang sering muncul dalam ulasan positif serta akar masalah dari ulasan negatif. Penelitian ini menganalisis akar masalah ketidakpuasan pelanggan dan pekerja berdasarkan topik kunci negatif dengan menggunakan diagram *fish bone*. Sedangkan topik kunci diamati frekuensinya di dalam ulasan yang dikategorikan positif, lalu dibuat menjadi *link analysis*.



Gambar I.3 Metodologi Penelitian Rajendran et al. (2022)

Penelitian terdahulu lain yang menggunakan *text mining* dan ulasan konsumen adalah Kim (2018). Hasil akhir yang ingin dicapai peneliti sama dengan penelitian Kim (2018), yaitu CTQ. Oleh karena itu, penelitian ini akan mengadopsi metode penelitian Kim (2018) karena hasil akhir yang ingin diperoleh mirip, tapi dengan beberapa perubahan. Meskipun tujuan penelitian sama-sama menentukan CTQ, penelitian Kim (2018) menggunakan *text mining* untuk memperoleh CTQ dari industri jasa berupa perhotelan, sedangkan penelitian ini berfokus pada CTQ produk.

Berbeda dengan dua penelitian sebelumnya, Kim (2018) mengolah kata kunci yang telah diperoleh lebih lanjut. Kata kunci yang didapat menggunakan pemodelan topik diinterpretasi menjadi CTQ kuantitatif. Gambar I.4 menunjukkan alur metode *text mining* Kim (2018) untuk menentukan CTQ.



Gambar I.4 Metodologi Penelitian Kim (2018)

Metode Kim (2018) ternyata memiliki sebuah kekurangan. Dalam implementasinya, ada kemungkinan terdapat data-data yang kurang sesuai dengan topik yang sudah ditentukan. Akibatnya, data-data ini dipaksa masuk ke salah satu topik yang tersedia karena pembuatan topik baru yang lebih sesuai tidak diterima di dalam penggunaan metode ini. Oleh karena itu, penelitian ini mempertimbangkan metode lain dalam penentuan CTQ, yaitu menentukan topik sesuai data yang tersedia.

Kim (2018) tidak menjelaskan secara rinci metodologi penentuan kata kunci setiap aspek mutu. Selain itu, Kim (2018) tidak mendeskripsikan cara interpretasi VOC menjadi CTQ. Ketidaklengkapan informasi menyebabkan penentuan CTQ dengan metodologi Kim (2018) membutuhkan penelitian tambahan.

Setelah dilakukan penentuan CTQ, Kim (2018) mengolah analisa sentimen menjadi regresi untuk mengetahui pengaruh sentimen terhadap penjualan. Regresi linear dibuat Kim (2018) untuk mengetahui dampak sentimen

aspek-aspek mutu terhadap penjualan. Namun, penelitian ini tidak meneliti pengaruh sentimen aspek-aspek mutu terhadap penjualan. Oleh karena itu, pengolahan analisa sentimen menjadi regresi tidak digunakan dalam penelitian ini.

Dapat disimpulkan bahwa Kim (2018), Rajendran et al. (2022), dan Shankar et al. (2022) sama-sama menggunakan ulasan *online* dengan *text mining* meskipun metode yang digunakan berbeda. Namun, penggunaan ulasan *online* dari ketiga penelitian terdahulu belum mempertimbangkan penggunaan *review* teks dan *rating* angka sekaligus. Penelitian Kim (2018) dan Shankar et al. (2022) memperoleh penilaian konsumen berdasarkan *review* teks konsumen. Meskipun Rajendran et al. (2022) menggunakan *rating* angka dan *review* teks, penelitian tersebut tidak mempertimbangkan besar korelasi *rating* angka dan *review* teks dan mengkategorikan ulasan hanya berdasarkan *rating*. Penggunaan *review* teks dan *rating* angka serta pertimbangan korelasi antara kedua bentuk ulasan menjadi salah satu pembeda penelitian ini dengan ketiga penelitian tersebut.

Penelitian ini mempertimbangkan penggunaan *review* teks dan *rating* angka berdasarkan penelitian McAuley & Leskovec (2013). Dibandingkan dengan pendekatan yang hanya menggunakan salah satu di antara *review* teks dan *rating* angka, prediksi *rating* yang dihasilkan akan lebih akurat dengan pendekatan yang menggunakan *review* teks dan *rating* angka sekaligus (McAuley & Leskovec, 2013). Selain itu, penelitian ini mempertimbangkan pula adanya *Text-Rating Review Discrepancy* (TRRD), yaitu inkonsistensi antara *review* teks dan *rating* angka (Almansour, Alotaibi, & Alharbi, 2022). *Review* teks yang diberikan oleh pembeli kemungkinan dapat berbeda dengan *rating* angka yang bersangkutan. Oleh karena itu, penilaian pelanggan terhadap suatu produk tidak cukup menggunakan salah satu di antara *review* teks atau *rating* angka saja, sehingga dilakukan pengkajian korelasi antara *review* teks dan *rating* angka.

Penelitian ini akan menggunakan data produk *headset* dari penelitian (Candra, 2021) sebagai contoh kasus untuk mengimplementasikan *text mining* dalam menentukan CTQ. Data dikumpulkan oleh Candra (2021) dengan proses *web scraping*. Namun, penelitian Candra (2021) bertujuan untuk memperoleh kebutuhan dan preferensi konsumen produk *headset*, sehingga hasil akhir yang diperoleh berbeda dengan penelitian ini.

Text mining diaplikasikan melalui modifikasi metode Kim (2018) yang telah dijelaskan pada paragraf-paragraf di atas. Berdasarkan masalah-masalah

yang telah diidentifikasi sebelumnya, rumusan masalah pada penelitian ini adalah bagaimana metodologi berdasarkan *text mining* yang dapat menentukan CTQ produk dari *online review*?

I.3 Pembatasan Masalah dan Asumsi Penelitian

Batasan masalah dibuat untuk memastikan penelitian terarah dan terfokus sesuai dengan permasalahan yang telah ditentukan. Berikut merupakan batasan masalah dari penelitian ini.

1. Data yang digunakan pada penelitian adalah *online review* dari *website* Amazon.
2. Sumber data yang digunakan adalah dalam Bahasa Inggris.

Asumsi penelitian adalah anggapan dasar yang digunakan dalam penelitian. Berikut adalah asumsi-asumsi yang digunakan selama penelitian ini dilakukan.

1. Tidak ada perubahan *review* atau komentar dari data yang digunakan selama penelitian.
2. Tidak ada perubahan spesifikasi produk selama penelitian berlangsung.

I.4 Tujuan Penelitian

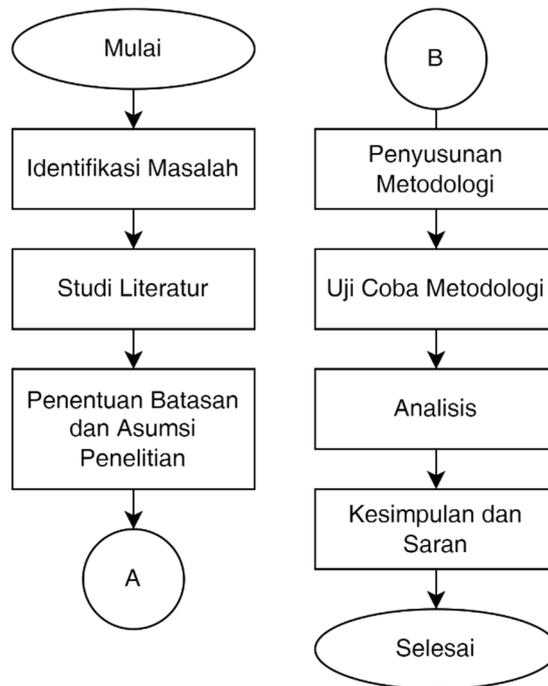
Pada penelitian ini, terdapat tujuan yang ingin dicapai. Tujuan ini akan menjawab rumusan masalah pada penelitian. Tujuan penelitian ini adalah menentukan metodologi berdasarkan *text mining* yang digunakan untuk menentukan CTQ produk.

I.5 Manfaat Penelitian

Sebuah penelitian selayaknya memberikan manfaat, baik bagi pemilik masalah, pembaca, maupun pihak lainnya. Manfaat penelitian yang diharapkan adalah penelitian dapat menjadi dasar dalam mengimplementasikan *text mining* dalam penentuan CTQ produk, terutama produk yang serupa dengan objek penelitian.

I.6 Metodologi Penelitian

Metodologi penelitian berisi tahap-tahap penelitian yang akan dilakukan. Metodologi pada Gambar I.5 dibuat dalam bentuk *flowchart*. Terdapat total 7 tahap yang dilakukan dalam penelitian ini.



Gambar I.5 Metodologi Penelitian

Tahapan metodologi penelitian berupa identifikasi masalah, studi literatur, penentuan batasan dan asumsi penelitian, penyusunan metodologi, uji coba metodologi, analisis, dan kesimpulan dan saran. Berikut adalah penjelasan dari masing-masing tahapan.

1. Identifikasi Masalah

Proses identifikasi masalah dilakukan untuk menghasilkan deskripsi dan rumusan masalah. Kemudian, masalah yang teridentifikasi menjadi dasar dari penyusunan tujuan penelitian. Pada proses ini, dijelaskan mengenai permasalahan yang ditemukan berdasarkan studi literatur terkait CTQ, yang menjadi landasan penentuan masalah yang diangkat dalam penelitian.

2. Studi Literatur

Studi literatur dilakukan dengan mempelajari literatur yang berkaitan dengan penelitian. Literatur yang digunakan terutama membahas tentang *text*

mining, pemodelan topik, analisis sentimen, VOC, dan CTQ. Literatur yang diperoleh digunakan sebagai referensi penelitian.

3. Penentuan Batasan dan Asumsi Penelitian

Dalam penelitian ini, ditentukan batasan dan asumsi penelitian. Batasan penelitian merupakan panduan penelitian agar penelitian tetap sesuai dengan ruang lingkup yang telah ditentukan. Asumsi masalah merupakan anggapan dasar yang digunakan dalam penelitian, dikarenakan kemungkinan adanya pengaruh dari faktor-faktor lain yang tidak termasuk ke dalam penelitian.

4. Penyusunan Metodologi

Berdasarkan studi literatur, disusun metodologi *text mining* untuk menentukan CTQ berdasarkan *online review*. Selain itu, penyusunan metodologi juga mengambil referensi dari beberapa penelitian yang telah dilakukan sebelumnya. Penelitian yang diambil terutama penelitian *text mining* yang mengolah *online review* seperti Candra (2021) dan Kim (2018).

5. Uji Coba Metodologi

Setelah disusun metodologi penentuan CTQ tersebut, maka selanjutnya metodologi yang bersangkutan dilakukan uji coba. Uji coba dilakukan dengan memasukkan data Candra (2021) sebagai contoh kasus. Uji coba dinyatakan berhasil apabila metodologi mampu menentukan CTQ dari data tersebut.

6. Analisis

Analisis dilakukan terhadap hasil dari tahap penelitian sebelumnya. Poin-poin utama yang dianalisis berupa pra-pemrosesan, pemodelan topik, interpretasi atribut, analisis sentimen, aplikasi metodologi, dan kekurangan penelitian.

7. Kesimpulan dan Saran

Tahap ini merupakan penyusunan kesimpulan dan saran. Kesimpulan dibuat berdasarkan tujuan penelitian dan bertujuan untuk menjawab rumusan masalah. Sedangkan saran diberikan untuk penelitian selanjutnya.

I.7 Sistematika Penulisan

Sistematika penulisan laporan dibagi menjadi 5 bab. Pada subbab ini, dijelaskan 5 bab tersebut secara lebih mendetail, termasuk pembahasan yang tercakup dalam masing-masing bab. Berikut adalah penjelasan sistematika penulisan laporan penelitian ini.

BAB I PENDAHULUAN

Bab pendahuluan mencakup latar belakang dari penelitian dan identifikasi masalah yang diangkat dalam penelitian ini. Setelah latar belakang dan identifikasi masalah ditentukan, terdapat pembatasan masalah, asumsi penelitian, tujuan penelitian, manfaat penelitian, manfaat penelitian, metodologi penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab tinjauan pustaka berisi teori-teori yang digunakan di dalam penelitian. Penelitian ini berkaitan dengan topik *text mining*, VOC, CTQ, cacat produk, *Latent Dirichlet Allocation* (LDA), analisis sentimen, dan rangkuman penelitian Candra (2021). Oleh karena itu, tinjauan pustaka akan membahas teori yang berkaitan dengan topik-topik tersebut.

BAB III PENGUMPULAN DAN PENGOLAHAN DATA

Bab pengumpulan dan pengolahan data menjelaskan mengenai proses pengumpulan dan pengolahan data dalam penelitian ini. Data yang digunakan adalah *online review* yang telah diperoleh oleh Candra (2021). Setelahnya, data ini akan dilakukan penentuan topik dengan LDA, interpretasi atribut, analisis sentimen, uji distribusi, dan uji korelasi. Tahap pengolahan data sesuai dengan metodologi penelitian pada subbab I.6.

BAB IV ANALISIS

Bab ini berisi analisis dari hasil penelitian. Analisis meliputi pra-pemrosesan, pemodelan topik, interpretasi atribut, analisis sentimen, dan aplikasi metodologi.

BAB V KESIMPULAN DAN SARAN

Bab kelima berisi kesimpulan dan saran. Kesimpulan menjawab rumusan masalah dan tujuan penelitian yang telah ditetapkan. Saran pada bab ini merupakan saran peneliti untuk penelitian-penelitian selanjutnya yang merujuk pada penelitian ini.